3rd CEAS Air&Space Conference

21st AIDAA Congress

24/28 October 2011

San Giorgio Maggiore Island - Venice, Italy

CEAS 2011

AIDAA

proceedings sponsored by

agenzia spaziale italiana

esa

MONDAY OCTOBER 24th



CEAS 2011

Guidance, Navigation and Control

Design of an U.A.V. controller based on the open source platform Arduino

F. Rinaldi
Politecnico di Torino & Italian Air Force, Italy
F. Quagliotti
Politecnico di Torino, Italy
S. Chiesa
Politecnico di Torino, Italy

Keywords: U.A.V., Ardupilot, Arduino, autopilot, open source

Abstract

This paper focuses on the preliminary implementation and performances assessment of an unmanned aerial vehicles (U.A.V.) controller. Although not showing any innovative technical aspect of the automatic flight control, it underlines the simplicity of building a complete low cost unmanned aerial vehicle starting from a commercially available Ardupilot Mega board, telemetry modules and a G.P.S. sensor. In order to assess the reliability of this product some experimental flights have been performed and the relevant measured and processed data show an acceptable behaviour.
The open source nature of this product represents a revolutionary aspect of this technology allowing an impressively fast understanding and developing of these autonomous systems. This approach is especially applicable to research and academic environments.

1 Introduction

A growing area in aerospace engineering is the use and development of U.A.V.s for military and civilian applications. Currently there is a huge demand for U.A.V.s for real-time remote sensing missions. In particular U.A.V.s are very efficient in natural disasters monitoring and management. For instance in the occurrence of a forest fire, it is very difficult to have a precise situation awareness. In similar cases the deployment of U.A.V.s capable of flying at low altitudes and able to follow pre-assigned G.P.S. waypoints, equipped with suitable remote sensing devices, can drastically improve human situation awareness.

2 The Ardupilot Mega board

This paper concentrates on the performances assessment of an U.A.V. controller based on an *Ardupilot Mega board (A.M.b.)*. A.M.b. is a printed circuit board (P.C.B.) provided with an embedded processor and combined with circuitry to switch between the radio control (R.C.) and the autopilot control - see Fig. 1 -.
It is based on a 16 MHz Atmega 2560 processor and provided with a built-in hardware failsafe that uses a separate circuit (multiplexer chip and Atmega 328 processor) to transfer control from the R.C. system to the autopilot and back again.
The A.M.b. dual processor design allows 32 million instructions per second (MIPS) and supports up to 700 waypoints memorization, thanks to the 4 Kb EEPROM. Other memory embedded devices are a 128 Kb Flash Program Memory and a 8 Kb SRAM. The board is equipped with a 6-pin G.P.S. connector, 16 spare analog inputs - each one provided with an analog to

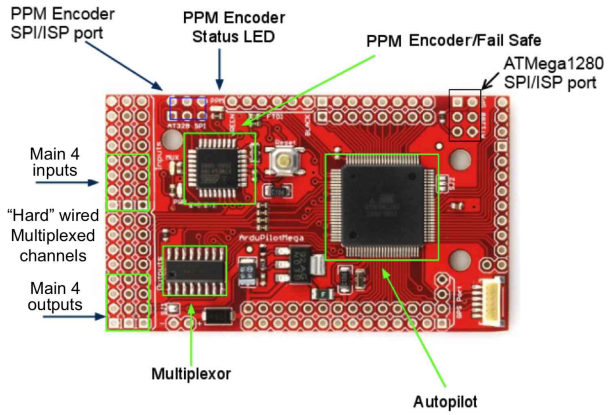


Figure 1: The Ardupilot Mega board

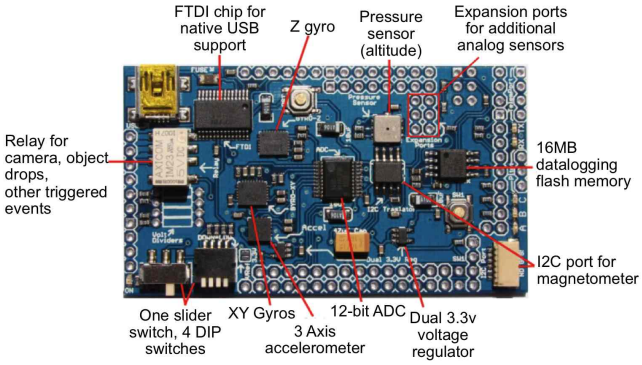


Figure 2: The OilPan Inertial Measurement Unit

digital converter (A.D.C.) - and 40 digital inputs/outputs for additional sensors. Four dedicated serial ports for two-way telemetry (using XBee modules) are available. Finally 8 R.C. channels, including the autopilot on/off one, can be processed by the autopilot. Sensors are mounted on a different P.C.B., the Inertial Measurement Unit (I.M.U.) board, commercially known as OilPan I.M.U. - see Fig. 2 -. This P.C.B. is equipped with

- a dual 3.3 volts voltage regulator,
- 3-axis accelerometers,
- 3-axis gyros,
- 12-bit A.D.C. for gyros,
- a built-in 16 MB data logger (*"the black box"*),
- a temperature sensor,
- a barometric pressure sensor for altitude sensing,
- a voltage sensor for battery status.

3 Telemetry and Ground Station

Wireless telemetry has been added using two 900 MHz XBee modules, the airborne and the ground transceiver. This solution has been chosen to avoid interferences with the 2.4 GHz R.C. gear and because of its guaranteed outdoor line-of-sight range of 10 Km.
The Ground Station (G.S.) is composed of a personal computer that, through the usage of a specific software, allows visualizing and recording several telemetry information like

- the U.A.V. attitude in terms of roll, pitch and yaw angles,
- the current three dimensional G.P.S. positioning and heading,
- the throttle level,
- the airborne lithium-polymer battery level,
- the airspeed (if airspeed sensors are connected),
- an aerial image of the overflown area, downloaded in real-time from the Google servers, through an on site web connection.

This is possible thanks to an open source available software, namely QGroundControl. QGroundControl is based on the open source conceived MAVLink Micro Aerial Vehicle Communication Protocol. The main features of this software are linked to the possibility to change real-time, in-flight parameters (previously stored in the EEPROM).
Several data can be modified from the G.S., for instance the controller gains. This allows for in-flight tuning of the automatic control law.
Another characteristic that needs to be pointed

out is the ability of changing and assigning waypoints directly from the G.S. while the U.A.V. is operating and flying.
Moreover the MAVLink Protocol supports up to 255 vehicles simultaneously, opening the panorama of *autonomous formation flying vehicles*.

4 The airframe

The chosen airframe is a Multiplex Easystar three channels powered glider. The three channels correspond to the motor, the rudder and the elevator servo; the aeromodel is not provided with ailerons and it is characterized by

- 1.37 m wingspan,
- 0.24 m^2 wing area,
- 3.28 Kg/m^2 wing loading,
- 0.86 m fuselage length,
- 0.68 Kg dry mass.

5 The controller code

The controller code can be loaded into the A.M.b. using the Arduino integrated development environment. This application is cross-platform, it is written in Java and based on Processing, avr-gcc and other freely available libraries. Being Arduino an open source platform it is fully programmable, allowing the users to prototype different control approaches.
The initial experimental flights have been conducted using a basic Proportional Integral Derivative (P.I.D.) control law. The autopilot handles both stabilization and navigation, eliminating the need for a separate stabilization system. In the tested configuration Ardupilot had a range of built-in flight modes, ranging from a simple flight stabilization system to a full autopilot able to follow pre-assigned G.P.S. waypoints.

6 Experimental flight tests

The following sections show the main experimental flight tests measurements, this in order to assess the controlled airplane flight performances.

6.1 Altitude control

Figure 3 shows the airplane behaviour in terms of a reference altitude tracking. The two vertical straight lines delimit the time interval during which the airplane was autonomously piloted. The controller objective was to track a certain given reference altitude.
The experimental measurements show an oscillating behaviour bounded in a range of $\pm$ 9 m around the objective altitude.
The nervous graph trend is symptom of a large control activity, this gives margin to further improvements.

6.2 Pitch, roll and yaw control

Figures 4, 5 and 6 show the aircraft attitude angles with respect to the starting position of the inertial platform. Of course, before taking off, the platform needs to be zeroed, i.e. aligned with the horizon.
In order to avoid stall conditions the airplane is constrained to not exceed the 15° pitch angle. As visible in Fig. 4, during the whole autonomously controlled time interval the airplane respects this condition. Some spikes are visible, probably due to strong ascensional currents; they are anyway well managed by the autopilot system that demonstrates its stability.
Roll angle, shown in Fig. 5, is successfully constrained by the autopilot to belong to the interval $[-45°, +45°]$.
Concerning the yaw angle no constraints are to be applied. In Fig. 6 some discontinuities can be observed, which are due to the way the yaw angle is measured. Essentially the yaw angle belongs to the interval $[0°, +180°[$ for rotations to the right and to the interval $]0°, -180°]$ for rotations to the left.
For completeness the ground speed, measured by means of a 10 Hz G.P.S. sensor, is shown in Fig. 7. This parameter is not constrained to track any reference.

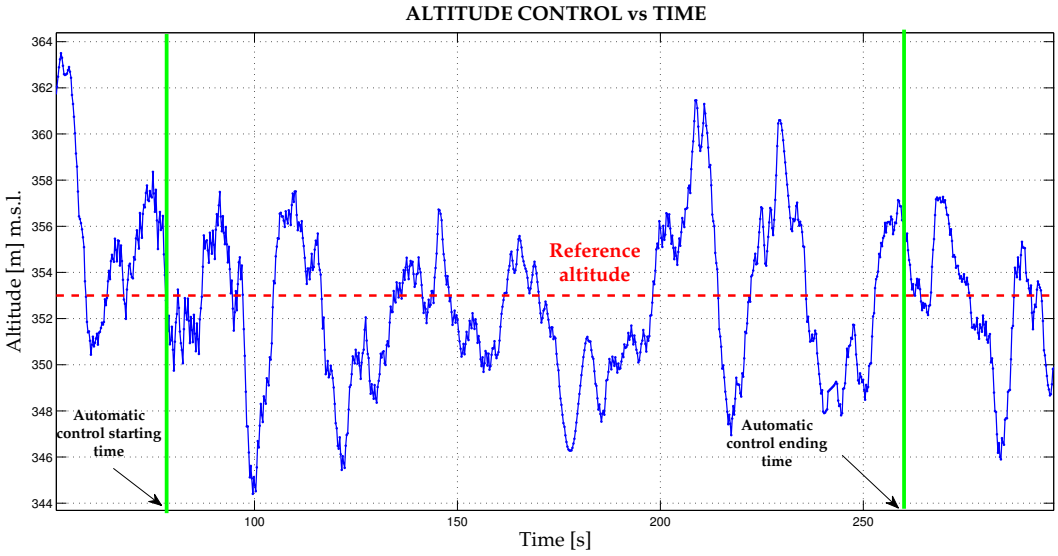


Figure 3: Altitude tracking

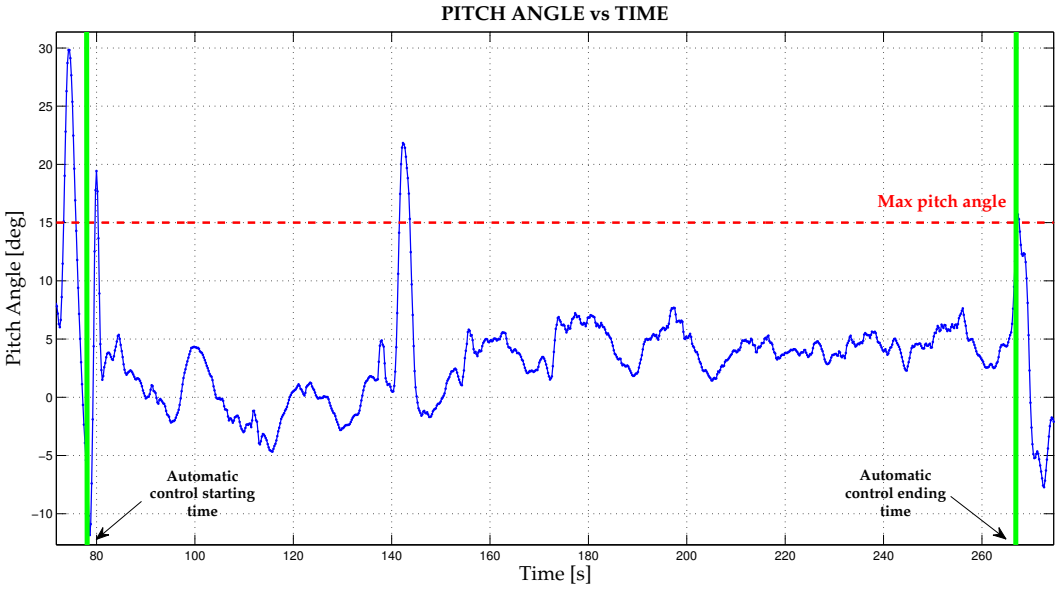


Figure 4: Pitch angle behaviour during the experimental test

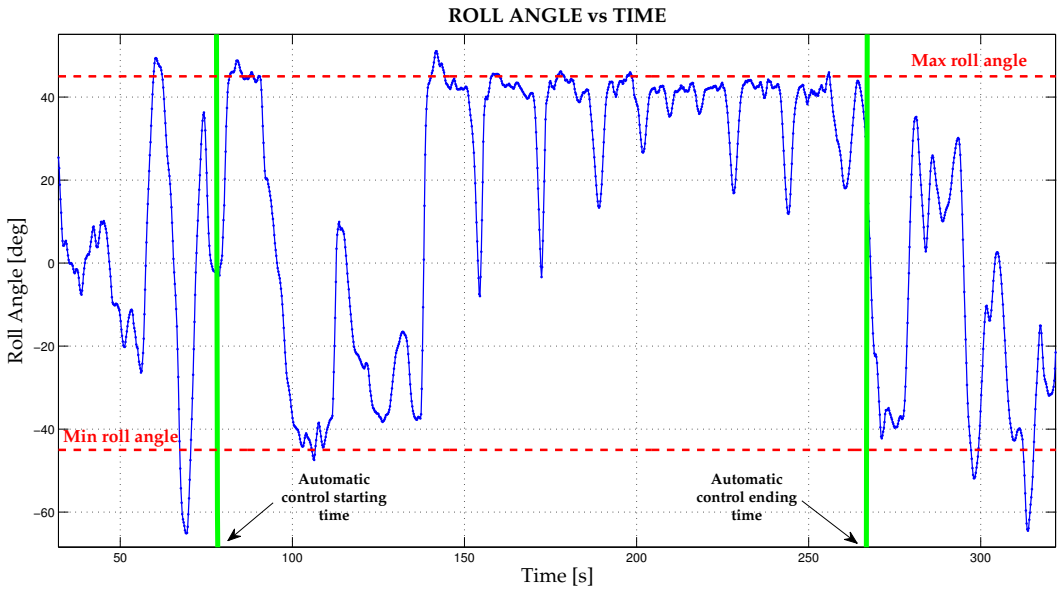


Figure 5: Roll angle behaviour during the experimental test

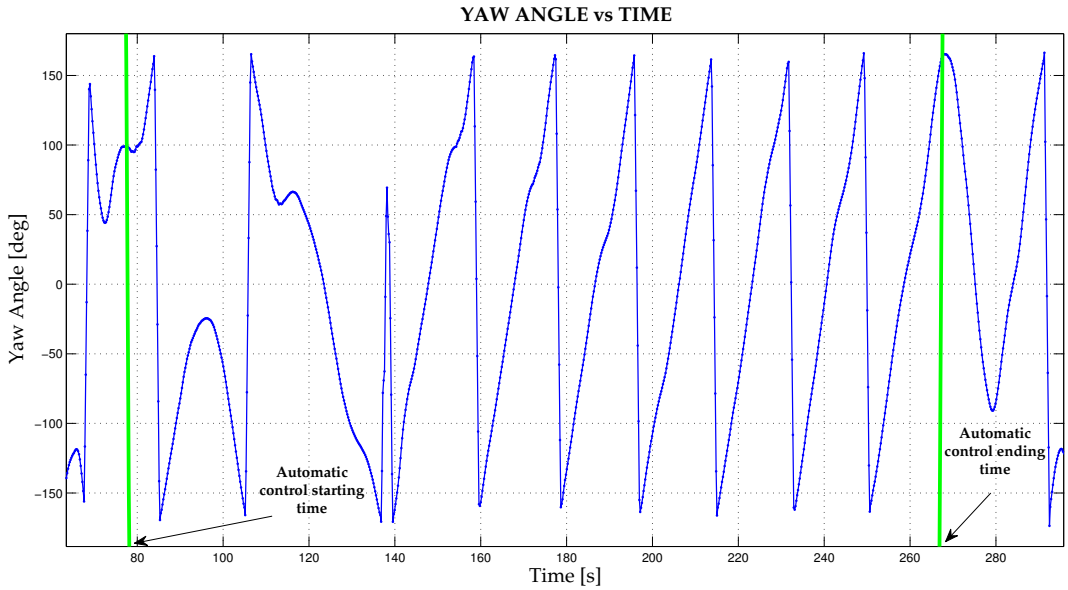


Figure 6: Yaw angle behaviour during the experimental test

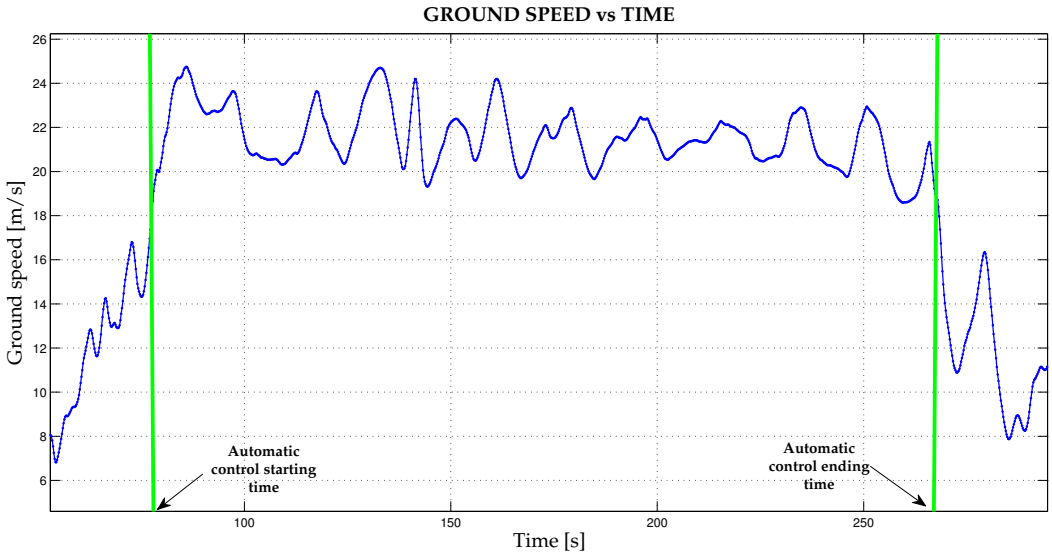


Figure 7: Ground speed behaviour during the experimental test

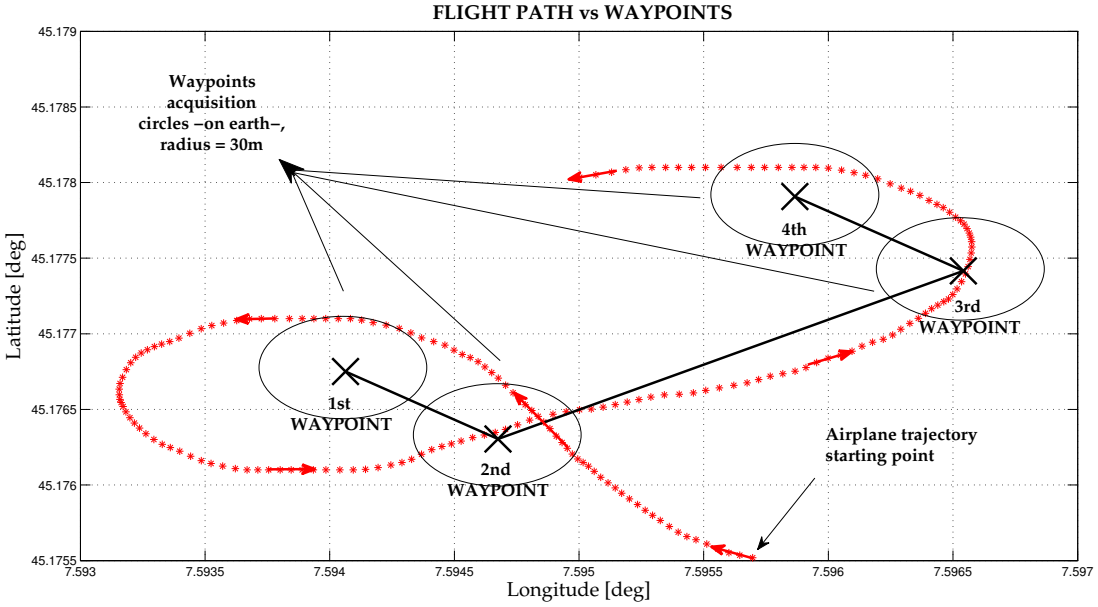


Figure 8: Experimental flight path through waypoints

6.3 Flight path through waypoints

The tested controller algorithm allows declaring a distance, from the waypoints, at which the target can be considered acquired. During the above described flight test, this parameter was set to 30 m.
Figure 8 shows the navigational behaviour of the controlled system. Crossed points represent the pre-assigned waypoints and the ellipses represent, on earth, circles of radius 30 m around them. As it can be observed in the figure, all the 4 targets are acquired by the test U.A.V. .

7 Competitiveness aspects

Two are the main competitiveness aspects linked to the described autopilot system.
The first one is the overall price of about 400 $ comprising the A.M.b, the OilPan I.M.U., the sensors and the telemetry modules. This price is far away from the other C.O.T.S. platforms.
The main factor is, of course, the open source nature of this autopilot system, since it is made of both open source software and hardware. This means that the hardware schematics, P.C.B. files, parts list are all freely available on the web, published under a Creative Commons license that allows free use and modifications as long as the resulting products retains the producer credit.
The software, both for the autopilot and the G.S., are also open source, published under a Lesser General Public License (L.G.P.L.) that allows free use and modification, as long as the resulting product is also open source and the producer attribution is retained.

8 Conclusions

This paper summarizes some experimental results obtained in order to test the basic performances of a complete open source autopilot.
In order to assess its behaviour a suitable data acquisition and elaboration system has been realized. This system, also partially based on an open source platform, allows recording and processing several flight parameters in order to understand the real behaviour of the flying platform.
Being the controller software completely open source, the future intention is to further modify it in order to prototype different control laws to be assessed with the developed data acquisition and elaboration system.

References

[1] Bin H. and Justice A., "The design of an unmanned aerial vehicle based on the Ardupilot", *Indian Journal of Science and Technology*, Vol. 2, No. 4, 2009, pp. 12-15.

[2] *http://code.google.com/p/ardupilot-mega/*

UAV Guidance Law for Trajectory Tracking of a Normalised Lemniscate

Niki Regina and Matteo Zanzi
University of Bologna, DEIS-ARCES,Italy

Keywords: UAV,Trajectory Tracking,Guidance law,Lyapunov.

Abstract

This paper derives a UAV guidance law to track a Lemniscate-like curve. A normalized Lemniscate is presented in order to create a constant speed reference path centered around a ground target position suitable to be overflown by an unmanned air vehicle. Tracking is achieved by introducing a virtual target that is moving along the Lemniscate according to the UAV actual position and velocity. The the guidance law together with its local stability is derived using Lyapunov theory.

1 Introduction

One of the main problems in unmanned aerial vehicle (UAV) research involves the determination of suitable autonomous guidance strategies that achieve specific tasks. For example, the ability to perform UAV surveillance operations or to track a specific target or to follow specific trajectories have been previously demonstrated importance in military context [3]. Similar types of surveillance and tracking tasks are also important in civil and commercial applications, such as: surveillance [17], maritime and rural search and rescue [15]. In these problems the ability of the UAV to track a pre-planned trajectory with high precision can be critical (especially in difficult urban contexts where the UAV flight path may be surrounded by buildings).

Different approaches have been proposed for UAV trajectory tracking applications. Among the different approaches existing in literature, two of the most important techniques are 1) path following and 2) virtual waypoint techniques. Path following technique states the objective of following onto a pre-defined path (freely in time) rather than following a sequence of trajectory points which must be reached at specific time instants. The most well known path following approach is the vector fields technique in which a field of vectors describe the heading angles required to guide the UAV onto the desired path (see[1] and [11]). An important practical property of these methods is that they exhibit some ability to reject disturbances, such as wind.

Alternatively, in a virtual waypoint trajectory tracking technique the UAV is directed towards a particular position on the desired path at a particular time instant. This point is termed "virtual" because the location of this reference point varies along the desired trajectory in time and in a manner that ensures it cannot be reached by the UAV. This virtual waypoint tracking approach has previously been used for tight tracking of curved and straight line tracking by UAVs [12], robotic [16] and marine vehicle [14] applications. The limits of the performances between path following and reference trajectory tracking is studied in [2] and [13].

Recently, we have investigated the design

of alternative trajectory tracking technique in which a non linear guidance law is able to track a moving or a fixed target (avoiding the need for either virtual waypoint or vector field framework) [9]. However, one clear disadvantage of this guidance approach is the inability of the UAV to overpass a fixed target more than once.

In this paper, a similar strategy is pursued but with some upgrades: the use of a Lyapunov technique in order to directly design a controller that is able to track a Lemniscate-like trajectory (which allows continuous surveillance and overfligth of a fixed point). One key aspect of this development is that it assures the fulfillments of almost constant air speed and the maximum turn rate constraints required by the air vehicle during the flight. The development of this nonlinear guidance law is accomplished by introducing a special sort of Frenet frame whose origin is defined in the closest point on the reference trajectory from the UAV actual position. Stability results are presented with establish that the controller will reduce both the distance and heading error to zero. The particular choice of the Lemniscate, called normalized Lemniscate, ensures a continuous overflight on a target placed at its origin. Consequently it seems to be perfect for surveillance or monitoring purposes.

This paper is organized as follows. Section 2 describes the tracking reference trajectory, i.e. the normalized Lemniscate and the kinematic model of the UAV. Section 3 derives a transformed kinematic model of the UAV necessary for the development of the guidance law. Section 4 provides a special reference system linked to the position of the "target" is presented. Section 5 develops the guidance law through the Lyapunov technique. Finally, section 6 gives some simulation results and comments are expressed in section 7.

2 Problem Mathematical Model

2.1 Tracking Reference System

A normalized Lemniscate has been chosen as the candidate reference trajectory for tracking. The Lemniscate is a plane curve with the following time description [8]:

$$\begin{cases} x_L(p) = \frac{a\cos(p)}{1+\sin^2(p)} \\ y_L(p) = \frac{a\sin(p)\cos(p)}{1+\sin^2(p)} \end{cases} \quad (1)$$

where a is the maximum abscissa value, as depicted in Fig. 1, and $p \in \mathbf{R}$ is the time varying curve parameter.

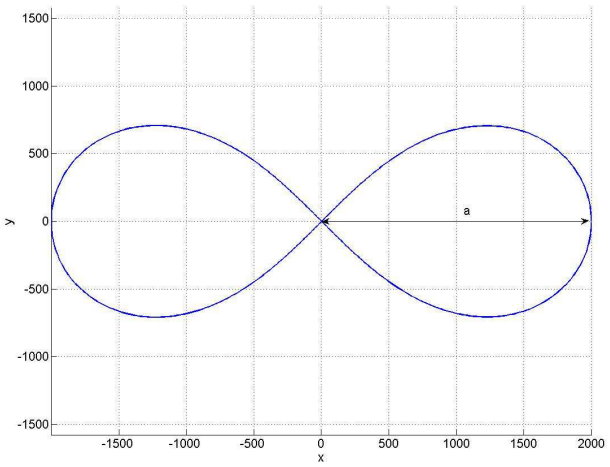


Figure 1: Geometry of the Lemniscate (with $a = 2000[m]$)

2.2 Kinematic Model of the UAV

The magnitude of the velocity (or speed) of the UAV is assumed constant to simplify presentation. Wind is assumed to be not present. This assumption is made also in other papers [1] and [9]. We highlight that the constant aircraft speed assumption is consistent with the usual tracking requirements because aircraft velocity changes are generally considered undesirable if altitude changes are to be avoided. A Cartesian frame description of UAV kinematic 2D motion is (see Fig. 2):

$$\begin{cases} \dot{x} = V\cos(\psi) \\ \dot{y} = V\sin(\psi) \\ \dot{\psi} = \frac{a_n}{V} \end{cases} \quad (2)$$

Here, vector $z = [x, y]^T$ denotes the position of the UAV in a fixed reference frame. The UAV is assumed to be equipped with an autopilot that implements the heading angle ψ hold mode. V represents the constant airspeed of the UAV

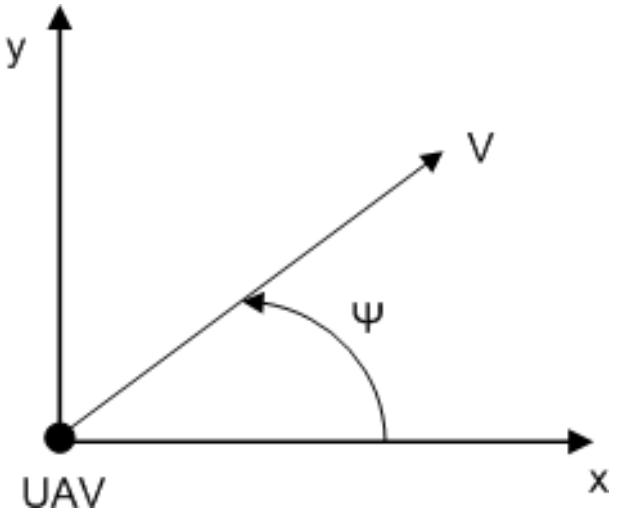


Figure 2: Velocity and angle definition

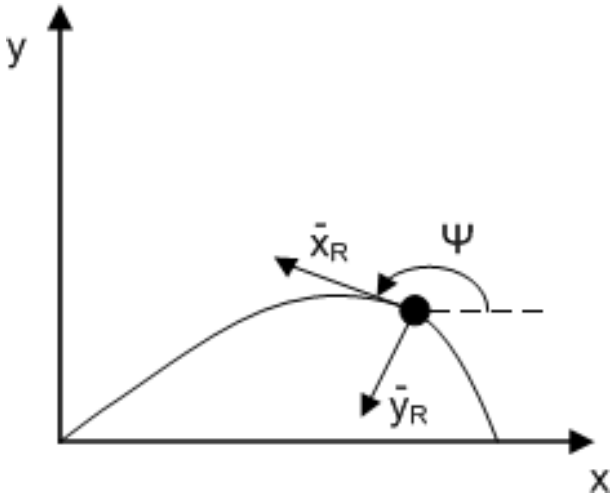


Figure 3: Ferret-Serret Reference Frame

while a_n is the commanded lateral acceleration from an outer guidance loop that controls the turn-rate.

A similar mathematical model holds for a reference point moving along the Lemniscate; in this case the addition of subscript R is made on the state variables.

Define the current reference tracking error as:

$$\phi = \psi - \psi_R \tag{3}$$

The objective of the guidance law a_n is to drive reference tracking errors to zero.

3 Transformed Mathematical Model of Tracking

In the previous section a fixed reference frame, that now we indicate with I, is considered. In order to track the Lemniscate trajectory, shown in Fig. 2.1, we introduce a second reference frame that is defined along the curve.

In fact, at any point on the Lemniscate a new reference system can be added with origin on this point with the x axis tangent to the curve (in forward heading direction) and y-axis at 90 degrees to the x-axis. The new reference system (which will be termed a Frenet - Serret frame F) is shown in Fig. 3.

Considering the vector z_r as the closest point between the position of the UAV and the Lemniscate, it is also possible to define the vector $\bar{z}_r$ as the position of the UAV with respect to the F frame and note that $z = z_r + \bar{z}_r$. Figure 4 describes the geometry of the vectors and the two different frames I and F.

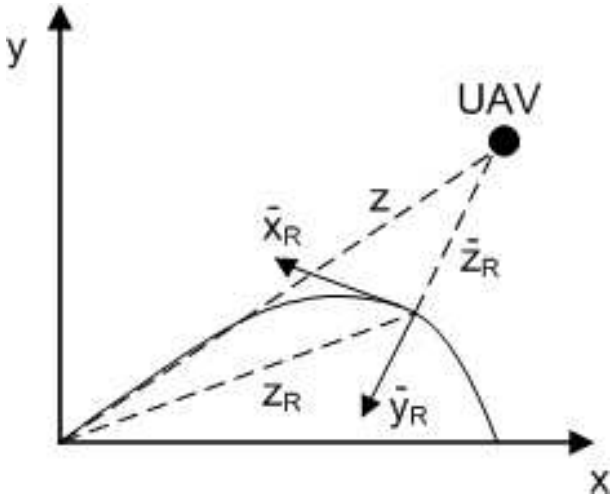


Figure 4: Lemniscate Vectors

Working with multiple frames, we will use $\left.\frac{d(\cdot)}{dt}\right|_I$ to denote the time derivative respect to the fixed frame and $\left.\frac{d(\cdot)}{dt}\right|_F$ respect to the new frame $\{F\}$.

In the following, $[x_r, y_r]^T$ indicates the coordinates of vector z_r in the $\{I\}$; moreover $[\bar{x}_r, \bar{y}_r]^T$

are the coordinates of $\bar{z}_r$ with respect to $\{F\}$.

Lemma

The velocity of the UAV in the {I} frame can be described by:

$$\left.\frac{dz}{dt}\right|_I = \left.\frac{dz_R}{dt}\right|_I + \left.\frac{d\bar{z}_R}{dt}\right|_F + \bar{g}_R \tag{4}$$

where $\bar{g}_R = \left[-\dot{\psi}_R \bar{y}_R \;\; \dot{\psi}_R \bar{x}_R\right]^T$

Proof

The Coriolis rule for a vector in a rotating frame is [10]:

$$\left.\frac{d\bar{z}_R}{dt}\right|_I = \left.\frac{d\bar{z}_R}{dt}\right|_F + \bar{g}_R$$

but

$$\left.\frac{d\bar{z}_R}{dt}\right|_I = \left.\frac{dz}{dt}\right|_I - \left.\frac{dz_R}{dt}\right|_I$$

The substitution of last equation into the first one gives (4).

We now derive some important relationships between information in the $\{I\}$ and $\{F\}$ frames. By indicating with R_I^F the rotation matrix between the two reference frames, it holds

$R_I^F = \begin{bmatrix} \cos(\psi_R) & \sin(\psi_R) \\ -\sin(\psi_R) & \cos(\psi_R) \end{bmatrix}$ (see [10]).

Moreover the following equation holds:

$$\bar{z}_R = R_I^F (z - z_r) \tag{5}$$

Hence, in order to express the velocity of the UAV in the $\{F\}$ frame we can write:

$$R_I^F \left.\frac{dz}{dt}\right|_I = R_I^F \left.\frac{dz_R}{dt}\right|_I + \left.\frac{d\bar{z}_R}{dt}\right|_F + \bar{g}_R \tag{6}$$

We also note that the first term on the right side of Eq. (6) can be also defined through the curvilinear abscissa s of the curve, which leads to the following expression:

$$\left.\frac{dz_R}{dt}\right|_F = [\dot{s}\;\; 0]^T \tag{7}$$

Recalling that $\left.\frac{dz}{dt}\right|_I = [\dot{x}\;\; \dot{y}]^T$ and $\left.\frac{d\bar{z}_R}{dt}\right|_F = [\dot{\bar{x}}_R \;\; \dot{\bar{y}}_R]^T$ and substituting these expressions and Eq. (7) into Eq. (6) it results:

$$\begin{cases} \dot{x}\cos(\psi_R) + \dot{y}\sin(\psi_R) = \dot{s} + \dot{\bar{x}}_R - \dot{\psi}_R \bar{y}_R \\ -\dot{x}\sin(\psi_R) + \dot{y}\cos(\psi_R) = \dot{\bar{y}}_R + \dot{\psi}_R \bar{x}_R \end{cases} \tag{8}$$

Define

$$\dot{\psi}_r = k\dot{s} \tag{9}$$

where k is the curvature of the Lemniscate equations; Eq. (8) becomes:

$$\dot{x}\cos(\psi_R) + \dot{y}\sin(\psi_R) = \dot{\bar{x}}_R + \dot{s}(1 - k\bar{y}_R) \tag{10}$$

By substituting the first two equations of (2) in the above equations gives:

$$\begin{cases} \dot{\bar{x}}_R = V\cos(\phi) - \dot{s}(1 - k\bar{y}_R) \\ \dot{\bar{y}}_R = V\sin(\phi) - \dot{s}k\bar{x}_R \end{cases} \tag{11}$$

4 Uav Motion in a special $\{FS\}$ Frame

We will now consider a special frame $\{FS\}$ in which our Frenet - Serret frame is defined by the closest point on the Lemniscate to the current UAV location as shown in Fig. 5.

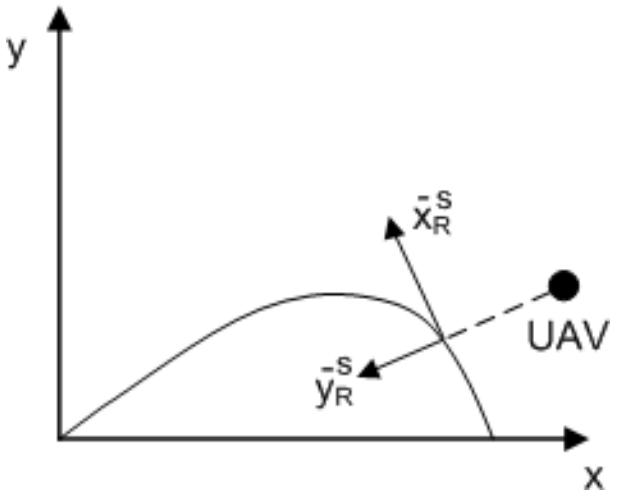


Figure 5: Special Frame

In this $\{FS\}$ frame we will have that $\bar{x}_R^s$ and $\dot{\bar{x}}_R^s$ are always equal to zero and consequently the system of Eq. (11) reduces to:

$$\begin{cases} \dot{\bar{x}}_R^s = 0 \\ \dot{\bar{y}}_R^s = V \sin(\phi^s) \end{cases} \tag{12}$$

where $\phi^s = \psi - \psi_R^s$.

Hence from Eq. (11) the variable $\dot{s}$ is given by

$$\dot{s} = \frac{V \cos(\phi^s)}{1 - k\bar{y}_R} \tag{13}$$

By multiplying both sides of Eq. (13) by curvature k and considering Eq. (9), it is possible to rewrite Eq. (12) as:

$$\begin{cases} \dot{\bar{\psi}}_R^s = \frac{kV\cos(\phi^s)}{1-k\bar{y}_r} \\ \dot{\bar{y}}_R^s = V \sin(\phi^s) \end{cases} \tag{14}$$

The first equation of 14 has a singularity in $\bar{y}_R = \frac{1}{k}$.

The curvature can be expressed as:

$$k = \left| \frac{\dot{x}_L \ddot{y}_L - \ddot{x}_L \dot{y}_L}{\left(\dot{x}_L^2 + \dot{y}_L^2\right)^{\frac{3}{2}}} \right| \tag{15}$$

Considering the particular geometric shape of the Lemniscate we can substitute Eq. (1) into Eq. (15). Figure 4 shows the curvature on the Lemniscate as a function of its parameter and, considering a Lemniscate with an abscissa axis a equal to 2000$[m]$, the maximum value of the curvature is on the most distant point on abscissa axis and its value is equal to $k_{max} = 0.0015$ [1/m]. Consequently, a tube around the Lemniscate, where $\bar{y}_R = 666.66$[m], can be considered as shown in Fig. 4. The two lines (inner and outer) to the Lemniscate represent the bounds within which Eq. (13) is always defined.

The initial position of the unmanned aerial vehicle must be inside this bound.

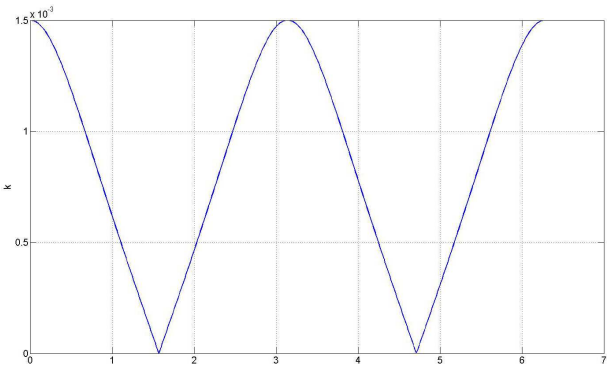


Figure 6: Curvature of the Lemniscate

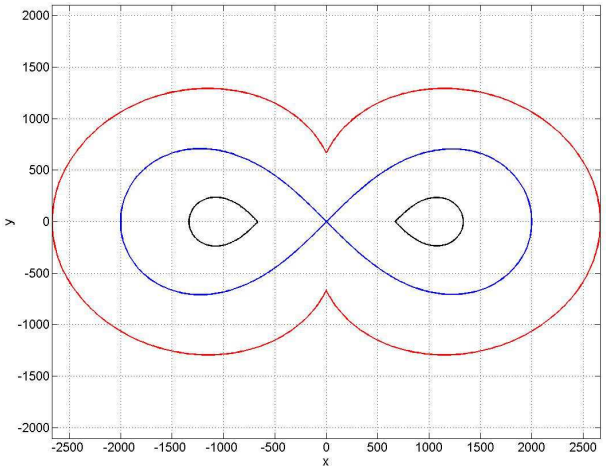


Figure 7: Bounds around the Lemniscate

5 Guidance Law

The aim of this paper is to develop a guidance law which stabilizes an UAV onto the normalized Lemniscate curve. In this section we derive a guidance law which ensures the tracking of the Lemniscate.

A suitable guidance law is provided by the following theorem.

Theorem:

Consider the intermediate control variable:

$$\begin{array}{l}\frac{d\phi^s}{dt} = \frac{df(\bar{y}_R^s)}{dt} - \\ - \frac{K_1 V \bar{y}_R^s (\sin(\phi^s) - \sin(f(\bar{y}_R^s)))}{(\phi^s - (f(\bar{y}_R^s)))} - \\ - \mathrm{K}_2 (\phi^s - f(\bar{y}_R^s))\end{array} \quad (16)$$

where K_1 and K_2 are positive gains and

$$f(y_R^s) = -\frac{\pi}{4}\tanh(K_3 \bar{y}_R^s) \quad (17)$$

with K_3 positive constant.

The guidance law:

$$a_n = V\left(\dot{\psi}_R^s + \dot{\phi}^s\right) \quad (18)$$

is locally asymptotic stable to $\bar{y}_R = 0$ and $\psi = \psi_R$.

Proof:

We will establish the result until Theorem 25 page 168 of [4]. Consider the candidate Lyapunov function:

$$V_L = \frac{1}{2}\left[(\bar{y}_R^s)^2 + \frac{1}{K_1}(\phi^s - \mathrm{f}(\bar{y}_R^s))^2\right]$$

and the equilibrium point $\bar{y}_R^s = 0, \phi^s = 0$. We first note that the function V_L is C^1 continuous and, from Eq. (3), (12), (17), positive definite.

Let us introduce variable $z^s = (\phi^s - f(y_R^s))$.

We now check if $-\dot{V}_L$ is locally positive definite.

By differentiating V_L we obtain:

$$\dot{V}_L = V\bar{y}_R^s \sin(\phi) + \frac{1}{K_1}(\phi^s - f(\bar{y}_R^s))\left(\frac{d\phi^s}{dt} - \frac{df(\bar{y}_R^s)}{dt}\right).$$

Let us add and subtract the same value $V\bar{y}_R^s \sin(f(\bar{y}_R^s))$ on the right hand side of this expression to give:

$$\begin{array}{l}\dot{V}_L = V\bar{y}_R^s \sin(\phi) + \frac{1}{K_1}(\phi^s - f(\bar{y}_R^s))\left(\frac{d\phi^s}{dt} - \frac{df(\bar{y}_R^s)}{dt}\right) + \\ V\bar{y}_R^s \sin(f(\bar{y}_R^s)) - V\bar{y}_R^s \sin(f(\bar{y}_R^s)).\end{array}$$

Rearranging the equation we obtain:

$$\begin{array}{l}\dot{V}_L = V\bar{y}_R^s \sin(f(\bar{y}_R)) + \\ \frac{1}{K_1}(\phi^s - f(\bar{y}_R^s))\left(\frac{d\phi^s}{dt} - \frac{df(\bar{y}_R^s)}{dt}\right) + \\ V\bar{y}_R^s(\sin(\phi^s) - \sin(f(\bar{y}_R^s)))\end{array}$$

Multiplying the last right term by:

$$\frac{K_1}{K_1}\frac{(\phi^s - f(\bar{y}_R^s))}{(\phi^s - f(\bar{y}_R^s))}$$

gives:

$$\begin{array}{l}\dot{V}_L = V\bar{y}_R^s \sin(f(\bar{y}_R^s)) + \\ \frac{1}{K_1}(\phi^s - f(\bar{y}_R^s))\left(\frac{d\phi^s}{dt} - \frac{df(\bar{y}_R^s)}{dt}\right) + \\ \frac{K_1}{K_1}\frac{(\phi^s - f(\bar{y}_R^s))}{(\phi^s - f(\bar{y}_R^s))} V\bar{y}_R^s(\sin(\phi^s) - \sin(f(\bar{y}_R^s)))\end{array}$$

Using Eq. (16) we obtain: $\dot{V}_L = \mathrm{V}\bar{y}_R^s \sin(f(\bar{y}_R^s)) - \frac{K_2}{K_1}(\phi - f(\bar{y}_R^s))^2$.

Our definition of $f(\bar{y}_R)$ in Eq. (17) can be used to show that $\dot{V}_L$ is locally definite negative except at the equilibrium point $\bar{y}_R^s = 0, \phi^s = 0$. From Eq. (17) we see that $\psi = \psi_R^s$.

All the required conditions have now been established and hence Theorem 4.1 page 114 [5] can now be applied to prove the theorem statement.

Theorem 6.1. shows that the provided guidance law drives the aircraft to the Lemniscate with the correct heading $\psi_{\mathrm{R}}^{\mathrm{s}}$.

One key difficulty that occurs in implementing the guidance law presented in the Theorem relates to how the special frame and hence ψ_R^s are determined. Due to the fact that these questions relate to the determination of the closest point on the Lemniscate, the solution can be determined numerically from the UAV actual position and heading and given the size of the Lemniscate.

6 Simulation Results

In this section we illustrate the performance of the non linear guidance law presented in Eq. (16) - (18) using simulation tests. In both the simulations the aircraft speed V = 15 [m/s]

and the parameter that gives the size of the Lemniscate a = 2000 [m]. It is important to highlight that in the guidance law give by Eq. (16) - (18) we consider the closest point on the Lemniscate according with the sign of the heading angle of the UAV. The initial heading angle of the UAV is directed towards the Lemniscate.

In the simulations the gains are $K_1 = 0.2$ and $K_2 = 0.002$.

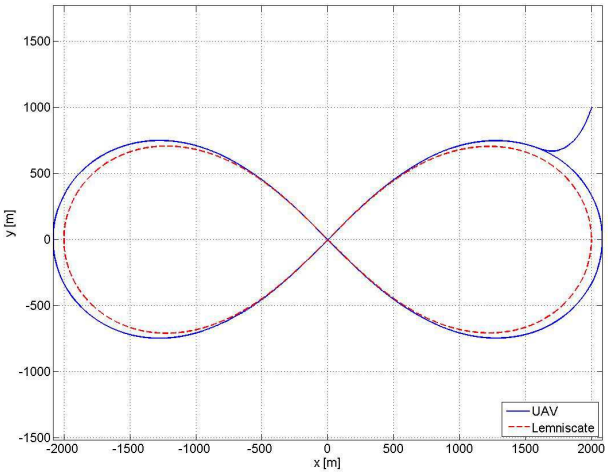


Figure 8: Trajectory of the UAV with the closest point technique

6 shows the trajectory of the UAV. The failure to track the Lemniscate is due to the tuning issue described at the end of section 6. Figure 6 shows that the maximum value of the lateral crossing error is less than 40 [m].

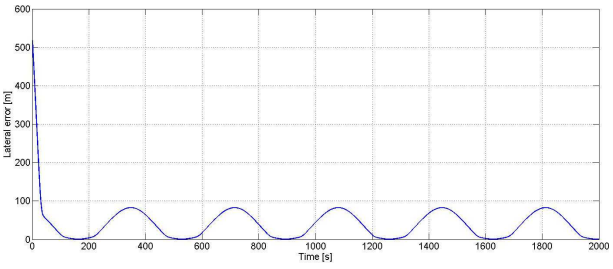


Figure 9: Lateral crossing error with algorithm technique

7 Conclusion

This paper presents an autonomous aerial vehicle non-linear guidance law for tracking a ground target located at the center of a Lemniscate. The basic idea is that of a virtual target which moves along the Lemniscate in a manner depending on the actual UAV position. The derivation of the guidance law, together with the proof of closed loop stability, is obtained by using the Lyapunov stability technique. Simulation results have shown good tracking performances that make possible the target continuous over-flight according flight constraints.

References

[1] Nelson D., Beard B., McLain T. and Beard W.,"Vector Field Path Following for Miniature AirVehicles",*IEEE Transaction on Robotics*, Vol. 23, No. 3, 2007, pp. 519-529.

[2] Aguiar I. P., Hespanha J., Kotovic P., "Path following for non minimum phase system removes performance limitations",*IEEE Transaction on Automatic Control*, Vol. 50, No. 2, 2005, pp. 234-239.

[3] ,"Trajectory Optimization with Risk Minimization for Military Aircraft",*Journal of Guidance and Control and Dynamics*, Vol. 12, No. 3, 1989, pp. 311-317.

[4] Vidyansagar M., *Nonlinear System Analysis*, 2nd ed., Prentice Hall, 1993.

[5] Khalil M., *Nonlinear System*, 3rd ed., Prentice Hall, 2002.

[6] Spivak M., *A Comprehensive Introduction to Differential Geometric*, 3rd ed., Perish, 1999.

[7] Krisitc M., Kannellakopoulus I. and Kotovic P., *Nonlinear and Adaptive Control Design*, John Willey and sons Inc., 1995.

[8] Battaglini G., *Giornale di Matematiche*, Napoli Benedetto Pellerano Editore, 1867.

[9] Regina N. and Zanzi M., "2D Tracking and Over-Flight of a Target by means of Non-Linear Guidance Law for UAV", *Aerospace Conference*, IEEE, Big Sky, 2009.

[10] Titterton D., Weston J. "StrapDown Inertial Navigation Technology", *Guidance and Control Conference*, AIAA, Reston, 2004.

[11] Lawreance D., Frew E. and Pisano W., "Lyapunov Vector Field for Autonomous UAV Flight Control", *Guidance and Control Conference*, AIAA, South California, 2007.

[12] Park S., Deyst J. and How J., "A New Nonlinear Guidance Logic for Trajectory Tracking", *Guidance and Control Conference*, AIAA, South Reston, 2004.

[13] Aguiar I. P., Dacic D., Hespanha J., Kotovic P., "Path following or reference tracking? An answer relaxing the limits to performance",*Symposium Intelligence Autonomous Vehicles*, IFAC/EURON, Lisbon, 2004.

[14] Encarnaa P., Pascoal A., "Combined trajectory Tracking and path following: an application to coordinated control of marine craft",*Conference of Decision and Control*, IEEE, Florida, 2001.

[15] Westall P., Oshea P., Ford J., "Improved Maritime Target Tracker using Colour Fusion",*Proceedings of HPCS 2009*, IEEE, Leipzig, 2009.

[16] Michalek M., Kozlowski K., "Trajectory Tracking for a Threecycle Mobile Robot: the Vector Field Orientation Approach",*Conference on Decision and Control*, IEEE, Sevilla, 2005.

[17] Frew E., Lawreance D., "Cooperative Stand off tracking of moving targets by a team of autonomous aircraft", *Guidance and Control Conference*, AIAA, South California, 2005.

Autonomous Take Off System: Development and Experimental Validation

N. Genito, E. De Lellis, V. Di Vito, L. Garbarino, C. Marrone
Italian Aerospace Researches Center (CIRA), ITALY

Keywords: *Autonomous take off, design and test approach, real-time hardware-in-the-loop validation, experimental data*

Abstract

This paper describes the design and validation process of an innovative Autonomous Take Off system, developed by the Italian Aerospace Research Center (CIRA) in the framework of the Italian national funded project TECVOL (Technologies for the Autonomous Flight). The autonomous take-off module is part of the autonomous Guidance, Navigation and Control prototype worked out by CIRA in the same project, where significant research effort has been devoted to achievement of high automation during all the flight phases, from take off to landing. The developed automated system allows take off, navigation through three-dimensional waypoints and landing of an aircraft without human intervention, also in presence of environmental disturbances and/or subsystem failures. In aerospace research and development activities not only functional requirements play an important role in the project, also process requirements and system engineering methods are fundamental for project success. In particular, the autonomous take off system development and validation process has been designed in order to be highly reliable but with a substantial reduction of needed time and costs. In the paper the process of design and validation applied to the proposed system development is examined in details, while providing also a description of the automatic take off system.

1 Introduction

In the aerospace industry, the research and development activities concerning the design of new automation logics and control algorithms and their integration in the avionic architecture typically requires a huge effort in terms of HW/SW design, implementation and testing. This is mainly due to current SW engineering process, which needs a tight monitoring of the SW development process for final product qualification. Moreover, a huge amount of validation tests have to be performed to obtain the final product acceptance for flight execution. This leads to the identification of two main requirements for Giudance, Navigation and Control (GNC) systems development: to perform a reliable development and validation process and to reduce development time and costs.

The achievement of the previous objectives may be too stringent or very time-consuming if the design process is approached using a traditional development cycle. Actually, this kind of approach allows validation and testing only during the final stages of the development process. If a design issue is detected during this phase, therefore, a revision of the project may be required, resulting in additional development time and costs.

The achievement of the previously indicated requirements, indeed, can be obtained by using techniques such as Control System Rapid Prototyping and real-time Hardware In the Loop

(HIL) simulation, together with Automatic Program Building. This approach allows the validation of the developed algorithms and systems by using real time laboratory tests instead of real world tests, so involving a significant reduction of both time and costs. These techniques, therefore, have been used in the TECVOL project, with particular reference to development of advanced guidance and control systems, from design to flight demonstration.

This paper shows the design and validation process applied to the development, in the TECVOL project framework, of an autonomous take off system. The paper is organized as follows. After a general description of the TECVOL project framework (section 1), the adopted design and test process is examined in section 2. The description of the proposed autonomous take off system is then reported in section 3, while section 4 focuses on the description of the validation facilities. Some relevant results obtained by the laboratory and real world validations, finally, are reported in sections 5 and 6, respectively.

2 The TECVOL project

Several research activities have been carried out in recent years in order to increase the autonomy features in UAVs [1], to expand the flight envelope [2] and to improve security levels of modern aircrafts, both manned and unmanned. In particular, a significant research effort has been devoted to achievement of high automation in the landing phase, so as to allow the landing of an aircraft without human intervention, also in presence of environmental disturbances and/or subsystems failures.

In the framework of the national-founded project TECVOL (Technologies for the Autonomous Flight), which continues and extends the previous CIRA project ATOL (Automatic Take-Off and Landing), successfully completed in 2004 [3], CIRA developed a complete autonomous mid-air flight, collision avoidance and landing system for fixed wing aircrafts.

In the ATOL project, CIRA developed algorithms for fixed path autonomous landing and successfully demonstrated their effectiveness by means of several in-flight experimentations, based on the use of a small scale fixed wing UAV. These algorithms still had some limitations, such as the ability to perform the autolanding manoeuvre only starting from a limited 3D region (fixed path autolanding), the use of GPS in RTK mode, the trajectory generation without considering the vehicle dynamic constraints (their satisfaction was demanded to the tracking algorithms) and the use of only basic recovery modes.

TECVOL project, therefore, aims to complete and extend the ATOL results, fully overcoming all the above mentioned limitations. This has been obtained by developing an autonomous mid-air flight system able to perform 3D waypoint following and an autonomous landing algorithm able to overcome the limitations of the autolanding system developed in the previous ATOL project.

In particular, the autonomous landing system designed in the TECVOL project is able to real time generate a trajectory compliant with the dynamic constraints acting on the vehicle, performing a fully autonomous landing of a fixed wing aircraft starting from any point of the three dimensional space and using the DGPS/AHRS technology. This algorithm has been successfully tested by means of both real time with hardware in the loop simulations [4] and flight tests [5].

For what concerns the take off functionalities, finally, in the ATOL project the Autonomous Take Off algorithm was developed only in a preliminary stage, while in the TECVOL project it has been completed and fully validated, as it will be shown in detail in the following.

3 Design and Test Process

The design and test process adopted in the TECVOL project is described here in the following. The development phases have been performed by using a top-down process, usually well known as V-cycle (Fig. 1).

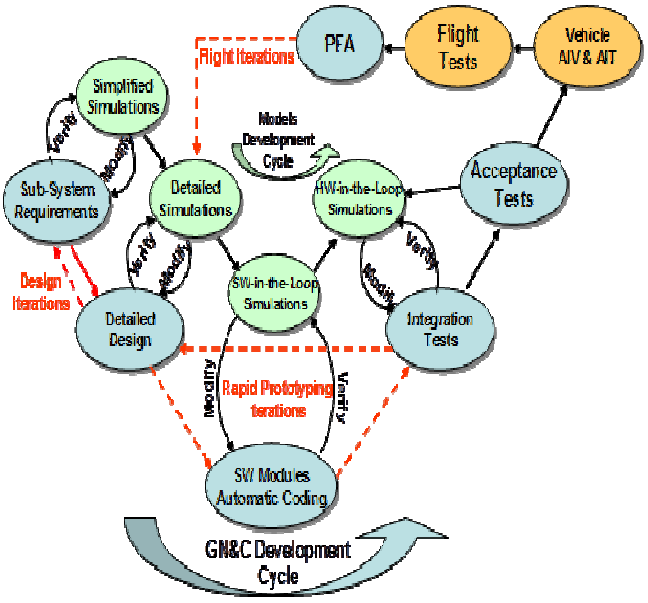


Fig. 1- V Cycle

Actually the whole design process can be divided into three main phases: requirement definition & system design, implementation & testing and integration & startup.

3.1 Requirement definition and system design

During this phase, system and user requirements are defined, system architecture is delineated and the design of the control system is carried out through the use of simulation models. The simulation models are developed in Matlab/Simulink/Stateflow environment. Each subsystem is validated through robustness and performance off-line analysis.

3.2 Implementation & testing

The developed control system is imported to an environment that allows the Automatic Program Building for a specific target machine. This approach has two main advantages:

- automation logics and control algorithms can be developed using a high level programming language;
- debugging can be easily done during both preliminary simulations, while defining the control strategy, and validation of the control system using HIL simulations.

During this phase the high level code is integrated with C/C++ code and downloaded to the target machine, which manages the resulting application according to its micro-kernel's primitives.

During this phase, HIL simulations are performed, allowing validation and testing of the control system, interacting with both the simulated environment and the real instrumentation (feedback sensors, real Human Machine Interfaces, Airborne Virtual Cockpits, Ground Control Stations).

An interesting feature of this technique is the capability to monitor and/or to modify (using SW tools) the system parameters and the control strategy. In this way, the validation process and the control system fine tuning become easy and quick.

3.3 Integration & startup

This is the final development phase. The GNC HW equipment is going to be integrated and GNC SW is going to be targeted and deployed in the host flight control computer. The GNC HW/SW equipment, finally, is going to be integrated in the aircraft and the system accepted after successful acceptance test sessions.

This development cycle has remarkable advantages, especially on the quality of the final product, including:

- close correlation between control system specifications, SW implementation and related documentation;
- reduction of the code generation time;
- "strong" control over implementation and/or specification errors.

4 Autonomous Take Off Algorithm Description

The Autonomous Take Off procedure is divided into four phases, each one corresponding to a specific automation logic state.

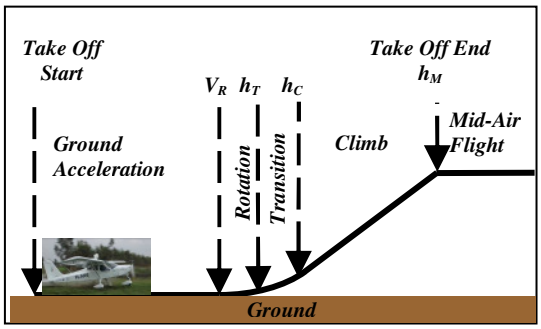


Fig. 2 Autonomous Take Off phases

Fig. 2 shows the indicated phases and the main variables generating the transition events between them.

The first take off phase is the ground acceleration. During this segment, the aircraft accelerates up to the specified rotation speed (V_R). The control action implemented by the automated system during this phase aims increasing the aircraft velocity, while at the same time maintaining the vehicle on the runway centerline.

The second take off phase (rotation) starts when the aircraft reaches the suitable rotation speed V_R (about $1.2 \cdot V_S$, where V_S the stall speed of the aircraft). During this phase, the control system commands a pull up, in order to increase the angle of attack so allowing the aircraft lift off.

When the aircraft reaches an height of 0.5 m (h_T), the rotation phase ends and the transition segment starts. In this take off segment, the control system action on the latero-directional channel changes from ground control to yaw damper control. Furthermore, during both the rotation and the transition phases, the control system manages to maintain the wings leveled, in order to prevent accidental impact of the wings with the ground.

When the aircraft reaches the safety altitude of 3 m (h_C), the last phase (climb) starts. During this phase, both the aircraft speed and altitude are increased.

The autonomous take off manoeuvre, finally, ends when the aircraft reaches the altitude of 35 m (h_M) above the ground level. At this point, the autonomous mission management system switches from take off macro-state to the mid air flight macro-state. Correspondingly, the aircraft control is passed from the autonomous take off system to the autonomous mid air flight system, which automatically generates the trajectory up to the first way point [6].

5 Validation Facilities Description

The TECVOL experimental set-up is mainly constituted by a flying demonstrator, named FLARE (Flying Laboratory for Aerospace Researches), and a Ground Control Station (GCS). The flying platform is shown in Fig. 3. It is an experimental ultra light aircraft with designed empty weight of 281 Kg, max take-off weight of 450 Kg, max speed s/l about 218 km/h, cruising speed about 190 km/h, wing area of 13.2 m^2, wing span of 9.6 m and maximum engine power of 100 hp.

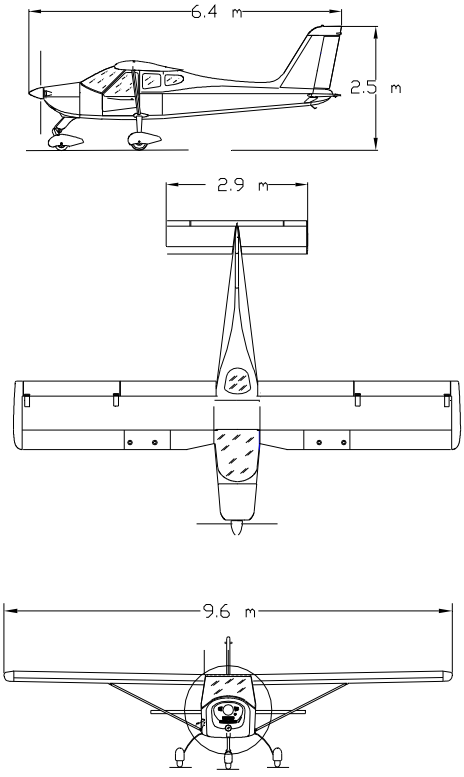


Fig. 3Aircraft 3-view

The FLARE vehicle is equipped with the avionic system used to implement the developed algorithms, whereas the ground control station is devoted, and suitably equipped, to allow the supervision and on-line setting of the in-flight experiments.

Both the on-board and the on-ground segments architectures will be described in the

next sections, in order to show how the laboratory validation test-rig is built starting by the real-set-up. As evident from the descriptions reported in the following paragraphs, the proposed architectural solution is suitably open and flexible to allow any update that may improve the system development.

All the critical components of the set-up are Commercial-Off-The-Shelf (COTS) elements. They have been previously selected and successfully validated on the platform FSSD (Flight Small Scale Demonstrator), used for the in-flight demonstrations in the ATOL [9 TBD] project framework.

The high level functional description of both the on board and ground segments is shown in Fig. 4.

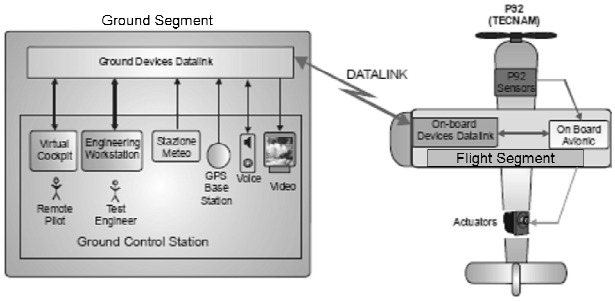


Fig. 4- High level system architecture

5.1 On board segment architecture

The FLARE high-level on board segment architecture is shown in Fig. 5.

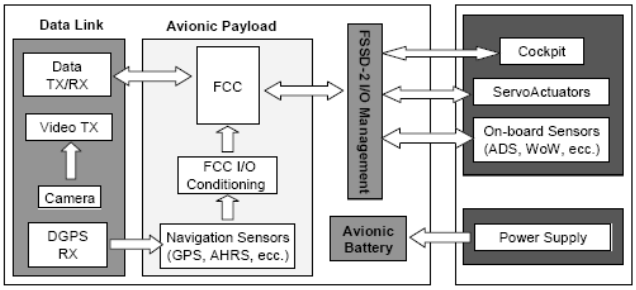


Fig. 5- FLARE on board segment architecture

The avionic architecture is based on the interconnection of the elements here listed:

- Flight Control Computer (FCC),
- GPS1, configured in DGPS mode,
- GPS2, configured in RTK mode,
- laser-altimeter,
- radar-altimeter,
- Air Data Computer (ADC),
- Attitude and Heading Reference System (AHRS),
- alpha and beta sensors,
- surface position sensors,
- actuators with own internal position sensors,
- downlink radio-modem,
- uplink radio-modem,
- GPS differential correction radio-modem,
- GPS1 avionic antenna,
- GPS2 avionic antenna.

Two GPS units are installed on board. They are differently configured, in order to have a meter comparison between the aircraft position and velocity obtained by the sensors fusion algorithm implemented in the FCC and the ones obtained by the RTK configured GPS. The GPS radio-modem output is connected to both the GPS units, in such a way each sensor receives and selects the RTCM messages required by the RTK and DGPS algorithms.

5.2 Ground segment architecture

The ground segment architecture is described in Fig. 6.

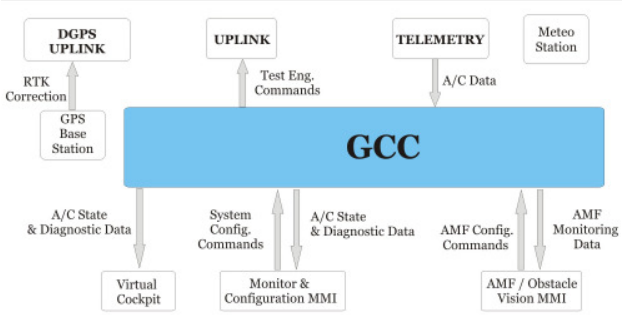


Fig. 6 - Ground segment architecture

This architecture is based on the interconnection of the elements listed in the next:

- Ground Control Computer (GCC),
- downlink radio-modem,
- uplink radio-modem,
- engineering workstation computer,
- virtual cockpit computer,
- Autonomous Mid-Air Flight Management Computer,
- GPS base station,

- GPS differential correction radio-modem.

The GCC is based on PC104 form-factor and is the core of the whole ground segment. It is connected, by an Ethernet field bus, with all the computers in the ground segment and, by dedicated point-to-point RS232 connection, with the uplink and downlink radio-modem.

The GCC also is equipped with a hard real time operating system.

5.3 Laboratory set-up architecture

The aim of this section is the description of the laboratory set-up where the on board software algorithms have been tested before the flight validation. The laboratory set-up is made up by the following components:

- a copy of the FCC installed on the vehicle,
- a functional realization of the GCS,
- a functional realization of pilot's cockpit and commands,
- simulator computer.

To execute real time pilot-in-the-loop simulations of the real mission, an important requirement is that the laboratory test-rig has to be able to acquire the pilot direct link commands. To achieve this purpose the laboratory test rig has also to replicate the particular front-end placed on the real pilot cockpit of the flight demonstrator.

In other words, all the aspects related to the management of real mission must necessarily be emulated in the laboratory set-up for all mission phases. The laboratory set-up functional architecture, designed based on the previous considerations, is shown in Fig. 7.

As shown in Fig. 7, the aircraft and sensors simulator is interconnected with both the FCC and the GCC. This choice has made because in the laboratory set-up the GCC manages also the communications with the pilot's cockpit simulation computer. This connection, indeed, is not present in the real experimental set-up, but it is necessary to simulate the direct connection between the pilot's cockpit and the aircraft.

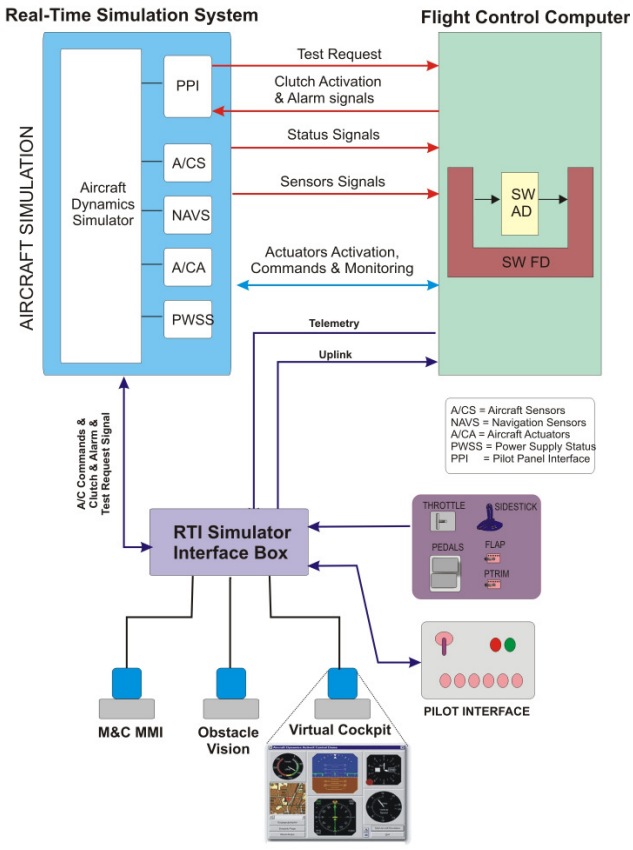


Fig. 7 - Laboratory set-up architecture

In particular, the ground control computer is used for the acquisition of pilot direct-link commands, for the management of the pilot cockpit emulator and for the management of data communications between aircraft simulator and ground segment.

The core of the laboratory set-up is the simulator computer. The main functionalities implemented in the simulator computer are:

- Interactive atmosphere model, which allows selecting wind direction, intensity, turbulence.
  It is used to build the predicted scenarios for testing on board algorithms robustness.
- Sensors model, like inertial unit, radar and laser altimeter, air data boom, surfaces position sensors, etc.
  During the run time test, the operator can modify the sensors behavior and insert sensor failures, with the aim of testing on board diagnostic algorithms and flight control law robustness.

- Actuators model, based on the open loop electrical actuators used in the experimental on-board set-up.
  It allows selecting different types of actuators and inserting the related failures.
- Vehicle dynamic model, a very high fidelity model of FLARE.
  It implements non-linear six degrees of freedom dynamic of a rigid body and accurate aerodynamic forces model.

For what concerns in particular the vehicle dynamic model architecture, it is shown in Fig. 8. The vehicle dynamic model validation has been carried out by means of several identification techniques [7]-[8]).

The model outputs have been compared with the results of proper identification flight tests and provided excellent matching. As output of the identification phase a trim procedure and a linearized model have been developed for offline test, needed as preliminary validation before implementing the real-time non linear model.

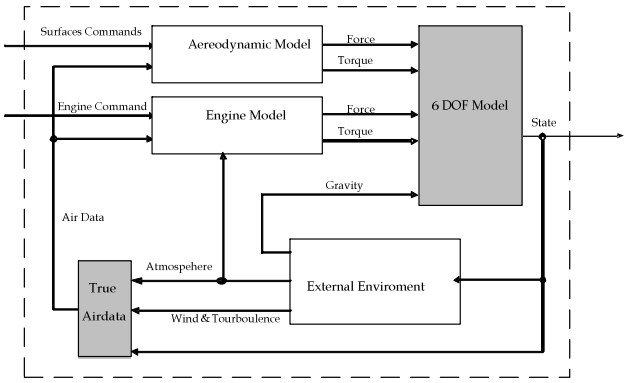


Fig. 8 – Vehicle Dynamic Model

Moreover, it has been modeled also a tricycle nose wheel steering landing gear. The landing gear model has been used both for the design and tuning of the autonomous take off algorithm and for its validation. The main assumptions adopted for the landing gear model design are:

- the wheel mass has been neglected;
- the dynamics of friction coefficient have been ignored.

The first assumption implies that the suspension has one degree of freedom. The second assumption, then, implies that the friction coefficients are constant.

The landing gear model inputs are: runway elevation; aircraft inertial state variables (position, velocity, attitude); nose gear steering angle; brake command; forces and moments in body axes. The outputs are: gear forces and moments in body axes; a Boolean signal simulating the Weight on Wheels (WoW) sensors.

6 Autonomous Take Off Validation

The validation process is composed of two kinds of test:

- laboratory tests;
- experimental tests.

These tests aim to verify the compliance of the developed system with the assigned requirements.

In the following paragraphs, the results are reported of the laboratory and experimental tests.

6.1 Autonomous Take Off Laboratory Validation

The laboratory tests are performed using two different simulation environments: a fast-time simulator and a real-time simulator.

The use of the fast time simulator allows evaluating:

- the performances of the autonomous take off system;
- the robustness of the autonomous take off system in presence environmental disturbances and uncertainties on the vehicle model parameters.

The full synthetic environment, including the landing gear model, has been used to perform several Monte Carlo simulation campaigns. They allowed testing the performances and the robustness of the whole autonomous GNC software in all flight phases, including on-ground segments. In the Monte Carlo simulations about thirty parameters have been considered as uncertain, so they have been characterized through suitable distributions. These uncertain parameters involved aerodynamics, propulsion, landing gear and inertial characteristics. Furthermore, in the

Monte Carlo simulations suitable variations have been considered of turbulence intensity, wind intensity and direction, discrete gust intensity and position.

The runway used for the test has the following characteristics and dimension:

- dry grass field,
- 433 m length,
- 20 m width.

The runway reference system is shown in Fig. 9.

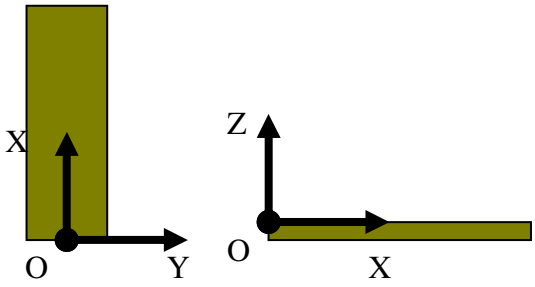


Fig. 9 Runway reference system

Summary results of one considered Monte Carlo simulation campaign are reported in Table 1. The reported data refer to the end of the rotation phase and have been obtained through a Monte Carlo simulation involving 1000 simulation trials. These results show that only 5 take-off are performed outside the runway.

The analysis of the worst cases showed that the take-off distance increase is due to the very particular simulation scenarios considered. Indeed, these cases refer to extreme conditions, namely:

- maximum rear wind;
- maximum longitudinal friction coefficient;
- maximum take off weight;
- maximum aerodynamic drag coefficient;
- minimum lift coefficient.

At the beginning of the rotation phase, moreover, the aircraft was swept by a rear gust.

These results show the good performances of the autonomous take off system and its robustness with respect to parameters uncertainties and external disturbances.

Before the execution of the in flight validation, then, the autonomous GNC system has been tested using real-time laboratory setup architecture.

Table 1 Take-off performances at the end of the rotation phase

Main longitudinal Statistics				
VAR	**Mean**	**Std**	**Max**	**Min**
X	235.98	68.77	476.02	95.88
V_Z	1.14	0.16	1.67	0.55
V_X	21.70	3.46	28.98	12.89
PITCH	9.44	0.52	11.13	8.02
ALPHA	7.04	0.75	9.69	4.83
Main Lateral Statistics				
Y	-3.12	1.71	1.37	-9.16
V_Y	-0.85	0.87	1.55	-3.23

This real-time hardware in the loop testing phase aimed to:

- verify the correct integration between the autonomous GNC system and the GCC;
- confirm the GNC system performances evaluation in a relevant laboratory environment using the same hardware used on board the vehicle FLARE (more precisely, using a copy of the FCC installed onboard);
- verify the correct real-time operation of the autonomous GNC system.

Fig. 10 shows the response of the control system to a step of 2 m in the Y direction. The test was performed to evaluate the on ground control system dynamics. The analysis of this test shows that the aircraft is able to perform the variation of lateral position with a small rudder input and with a very small overshoot. The step response shows the following features:

- raise time of about 5 s,
- steady-state error close to zero,
- overshoot of about 10%.

Fig. 11 shows a test performed to evaluate the aircraft trajectory during the first three phases of the autonomous take-off maneuver. The test was performed in presence of lateral wind and gust during the acceleration phase. The figure shows that the aircraft performs the take-off while almost remaining on the runway centerline. Only during the rotation and transition phases the aircraft puts the nose in the wind and gets away from the centerline of about 2 m.

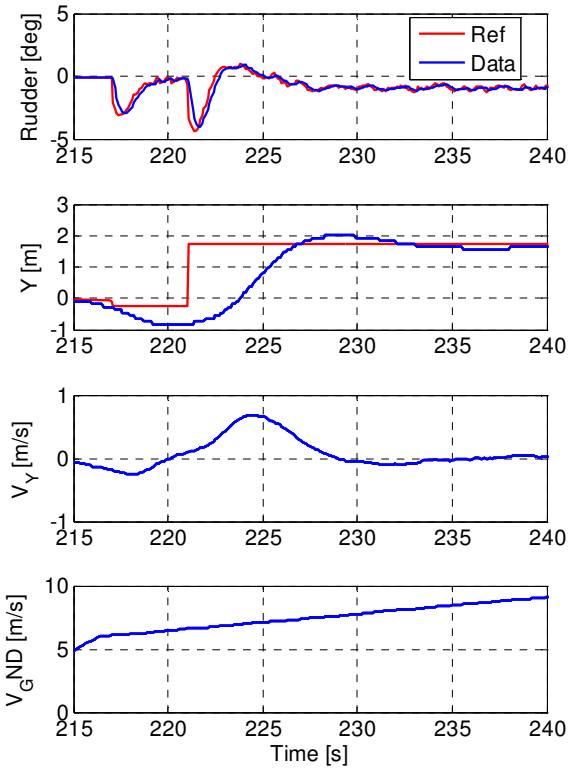


Fig. 10 - Step response in the Y direction (real-time simulation)

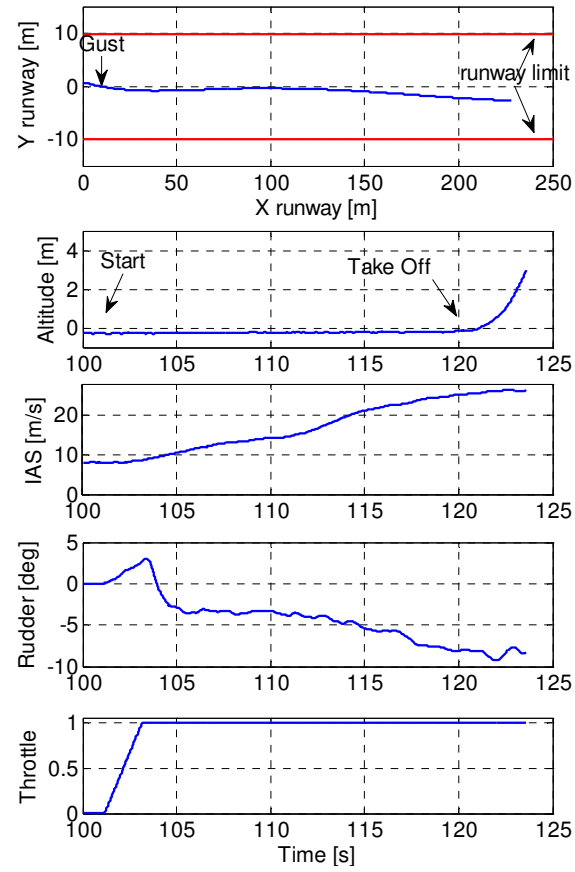


Fig. 11 – Aircraft trajectory during ground acceleration, rotation and transition phases (real-time simulation)

6.2 Autonomous Take Off Experimental Validation

The real world tests are necessary to verify the reliability of the adopted development process and of the implemented simulation environment.

Fig. 12 shows the same kind of test performed in the real-time simulation above cited: the response of the control system to a step of 2 m in the Y direction. The objective of this test was to confirm in the real environment the design characteristics of the control system. The step response shows the following features:

- raise time of about 6 s,
- steady-state error close to zero,
- overshoot of about 5%

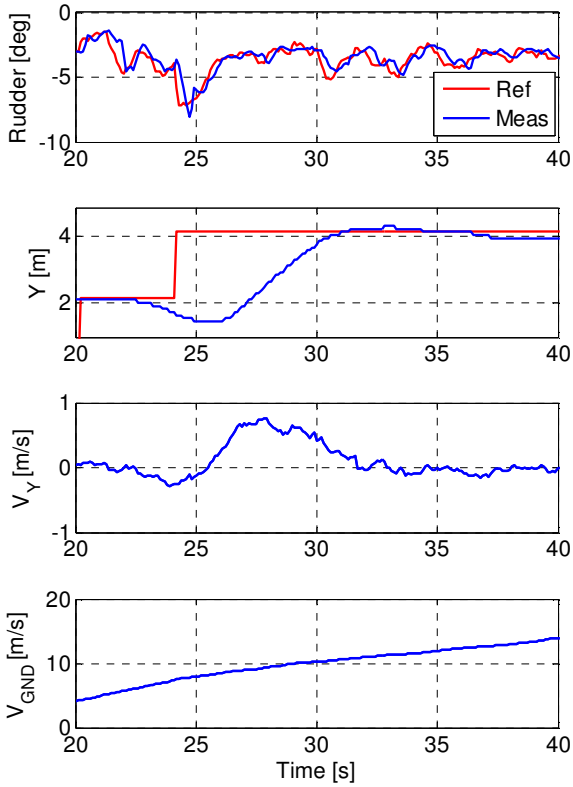


Fig. 12 - Step response in the Y direction (real environment)

Fig. 13 shows the first phase of the autonomous take off experimentation in a real environment. In particular the figure shows the performances of the aircraft during the acceleration phase.

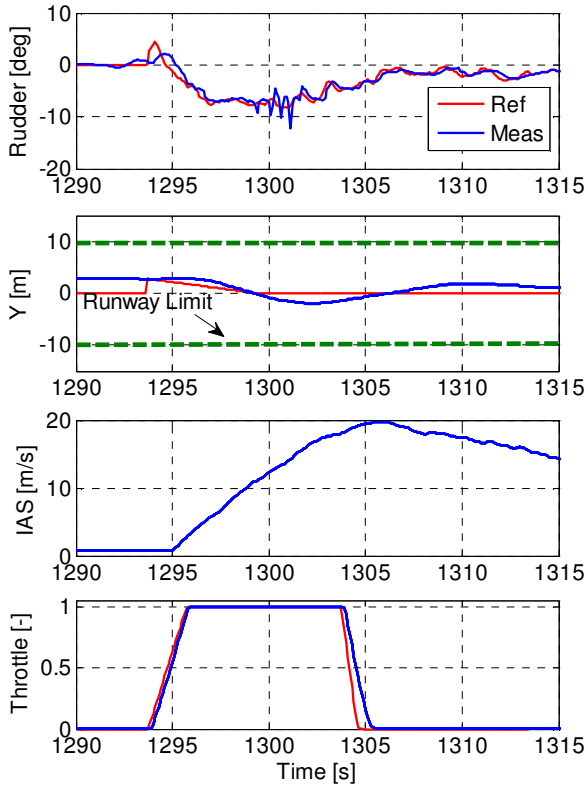


Fig. 13 - Aircraft trajectory during acceleration phase (real environment)

These results show the good performances of the autonomous take off system also in a real environment.

7 Conclusions

The paper described the development process and experimental validation of the autonomous take-off system recently developed by the Italian Aerospace Research center (CIRA) in the framework of the TECVOL project. It reported an insight about all the design phases, with particular emphasis on the validation by means of both laboratory on ground-validation test rig and real world experimentations.

The laboratory test rig has been described in details, with particular emphasis on its relevant peculiarity of including in the real-time with hardware-in-the-loop simulation also the human-machine interfaces, in addition to the aircraft, the external world simulator and the Flight Control Computer.

In the paper, furthermore, some laboratory real-time with hardware in the loop simulations have been reported, showing the flexibility of the proposed test rig and its usefulness in testing the algorithm in different operational conditions. Also results from real world validation have been described, showing the good performances of the autonomous take off system in a real operative environment and the consistency of the real world results with the ones obtained by using the laboratory simulation environment.

References

[1] Pachter M., "Challenges of Autonomous Control", IEEE Control Systems, August 1998

[2] Malaek S. M. B., Izadi H. A., M. Pakmehr, "Flight Envelope Expansion in Landing Phase Using Classic, Intelligent and Adaptive Controllers", Journal of Aircraft, vol. 43, No. 1, January-February 2006

[3] Accardo D., Cimmino G., Ciniglio U., Corraro F., Esposito F., Moccia A., "Integration of Advanced Altimetric Systems for UAV Vertical Navigation During Landing Manouevres", AIAA, Unmanned Unlimited Conference Workshop and Exhibit, (paper no. AIAA 2004-6318), 20 - 23 September 2004, Chicago, Illinois,USA.

[4] V. Di Vito, E. De Lellis, C. Marrone, U. Ciniglio, F. Corraro, "UAV Free Path Safe DGPS/AHRS Approach and Landing System with Dynamic and Performance Constraints", UAV Systems 2007 International Conference & Exhibition [CD-ROM], Paris, France, 12-14 June 2007

[5] V. Di Vito, E. De Lellis, C. Marrone, Genito N., U. Ciniglio, F. Corraro, "UAV Free Path Safe DGPS/AHRS Autolanding: algorithms and flight tests", submitted at UAS 2008 International Conference & Exhibition, Paris, France, 10-12 June 2008

[6] G. Ambrosino, M. Ariola, F. Corraro, E. De Lellis, et alii, "Path Generation and Tracking in 3-D for UAVs", IEEE Transactions On Control Systems Technology, VOL. 17, NO. 4, July 2009

[7] N. Genito, C. Marrone, E. De Lellis, and A. Giovannini, "Real-Time Landing Gear Simulation Model For A Very Light Aircraft: Development And Experimental Validation" In "Proceedings of AIDAA Conference", Milano, 2009.

[8] A.Vitale, N.Genito, L.Garbarino, U. Ciniglio and F. Corraro, "A hybrid frequency and time domains aerodynamic model identification from flight data", The Aeronautical Journal, vol.114, No.1156, June 2010, pp.377-385

Analysis of IMU error compensation based on rotating fiber strap-down inertial navigation system

Zhang ling，Liu jianye
Navigation Research Center, Nanjing University of Aeronautics and Astronautics, China

Lai jizhou，Lvpin
Navigation Research Center, Nanjing University of Aeronautics and Astronautics, China

Keywords: *environment function, bias-estimation, segmented calibration, error compensation*

Abstract:
With the development of Strapdown Inertial Navigation System (SINS)with rotating structure, the error output forms of inertial measurement unit(IMU) is changed, the system error is modulated, and the system performance is better than general SINS. For the error unsolved by rotating SINS, the impact of scale factors and installation errors on SINS is analyzed, and the corresponding error models are built. Based on the in-out IMU error characteristics, a method called environment function is proposed to identify and separate the IMU error coefficients. Besides, the positive-negative asymmetry of scale factor and singularity of environment function matrix have impact on system accuracy. In order to resolve the problem, a bias-estimation and segmented calibration are proposed to further improve the performance of rotating SINS. The experiment results indicate that the methods of environment function, bias-estimation and segmented calibration are effective, and the system property is improved.

1 Introduction

As an autonomous navigation method, SINS has been paid much attention and studied widely. Fiber-SINS(FSINS) is a kind of SINS, and has become the research focus of the world. The characteristics of FSINS show that when FSINS works for a long time, the navigation error accumulates over time, which limits long-time application of the FSINS system. Rotating modulation method is a valuable modulation means. It changes traditional FSINS structure, adds controllable rotating equipment, and sets IMU on the rotating equipment. The breakout on FSINS navigation performance testifies that the rotating modulation method plays an important role on the improvement of FSINS, and can provide a valuable orientation for autonomous navigation technique.

However, due to complex rotating movement, the present IMU error models of FSINS can not meet the accuracy demand of RFSINS. Take scale factor for example, due to the periodical changes of direction, scale factor errors include asymmetrical and symmetrical scale factor, where the symmetrical scale factor and some installation error are not modulated. As the impact of un-modulated error is relevant to velocity, the error may be complicated and calculated in light of the disturbed rotating

velocity.

In order to resolve the problem, an IMU error identification method based on environment function is modified .The method is used to identify and separate the IMU error of RFSINS. Besides, the advanced calculation methods are proposed such as bias-estimation and segmented calibration, for further enhancement of the error coefficients separation and improvement of the rotating SINS performance.

2 Principle of rotation modulation

According to traditional SINS, IMU is associated with vehicle, and susceptible to vehicle movement, environment and its own characteristics. In the rotation modulation design, IMU is fixed on the rotating structure of the vehicle, by which it can rotate relatively to the vehicle at a constant speed.

IMU with rotating structure is shown in Fig 1. The gyros and accelerometers are orthogonally installed. As the bridge of IMU and vehicle, the rotating structure drives IMU and does the synchronous movement with the vehicle. Meanwhile, the structure rotates regularly and modulates with IMU error. The rotation mode is shown in Fig 2[1].

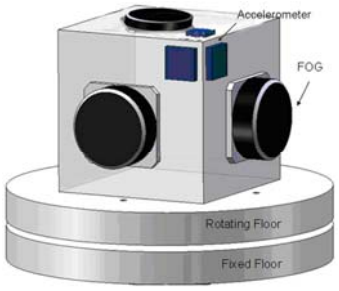


Fig 1. IMU with rotating structure

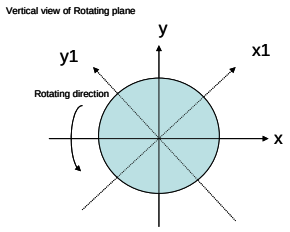


Fig 2. two-dimension rotation

Assume that the constant drift of the gyro is ε, after a t period, the angle error is calculated as

$$\Delta\theta = \varepsilon t \qquad (1)$$

With the help of rotating technology, when the structure rotates at a speed of Ω in a T period, the gyro drift is modulated and the angle error is calculated as

$$\Delta\theta = \int_{t0}^{t0+T} (\varepsilon \cos(\Omega t))tdt = 0 \qquad (2)$$

When the structure continually rotates, angle error caused by constant drift can not accumulate over time, which improves the system performance in a certain extent.

The principle of rotating SINS is elaborated as follows. At the beginning, the rotating coordinate(s) is coincident with the body coordinate(b), and ox_s represents the real x axis of gyro and accelerometer; oy_s represents the real y axis of gyro and accelerometer; oz_s is coincident with z axis of b coordinate, and vertical with ox_s 、 oy_s . The effective IMU outputs are received from the real IMU outputs in the process of coordinate conversion. The rotating angle is

$$\alpha = \Omega \cdot t \qquad (3)$$

The coordinate transformation matrix from s to b is

$$C_s^b = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(\Omega t) & -\sin(\Omega t) & 0 \\ \sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (4)$$

Then, the gyro and accelerometer error are

$$\begin{aligned}\Delta\omega^b &= C_s^b \Delta\omega^s \\ &= \begin{bmatrix} \cos(\Omega t) & -\sin(\Omega t) & 0 \\ \sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta\omega_x \\ \Delta\omega_y \\ \Delta\omega_z \end{bmatrix} \\ &= \begin{bmatrix} \Delta\omega_x \cos(\Omega t) - \Delta\omega_y \sin(\Omega t) \\ \Delta\omega_x \sin(\Omega t) + \Delta\omega_y \cos(\Omega t) \\ \Delta\omega_z \end{bmatrix}\end{aligned} \quad (5)$$

$$\begin{aligned}\Delta f^b &= C_s^b \Delta f^s \\ &= \begin{bmatrix} \cos(\Omega t) & -\sin(\Omega t) & 0 \\ \sin(\Omega t) & \cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \end{bmatrix} \\ &= \begin{bmatrix} \Delta f_x \cos(\Omega t) - \Delta f_y \sin(\Omega t) \\ \Delta f_x \sin(\Omega t) + \Delta f_y \cos(\Omega t) \\ \Delta f_z \end{bmatrix}\end{aligned} \quad (6)$$

Where, b represents body coordinate, s represents rotating coordinate, C_a^b represents coordinate transformation matrix from a to b, $\Delta\omega$ represents gyro error in s coordinate, Δf represents accelerometer error in s coordinate, Ω represents rotating speed.

In Eq.(5) and (6) of error modulation, IMU errors vertical to rotating coordinate(x axis and y axis) are modulated by rotation, and error forms are changed to regular cosine forms.

3 Analysis of rotating SINS

3.1 error model

The impact of random errors are ignored in SINS, constant error, scale factor error and installation error are considered as the composition of the IMU error. Ignore more than one- order small amount, and build the three-axis accelerometer model[2]:

$$\begin{aligned}{}_2\mathbf{a}_m &= (\mathbf{I}+\mathbf{K}_a)(\mathbf{I}+\boldsymbol{\theta}_a)\mathbf{a} + \nabla\mathbf{a} \\ &\approx (\mathbf{I}+\mathbf{K}_a+\boldsymbol{\theta}_a)\mathbf{a} + \nabla\mathbf{a}\end{aligned} \quad (7)$$

Then, the accelerometer error model is

$$\Delta\mathbf{a} = \nabla\mathbf{a} + \mathbf{K}_a\mathbf{a} + \boldsymbol{\theta}_a\mathbf{a} \quad (8)$$

Similarly, the gyro error model is

$$\Delta\boldsymbol{\omega} = \nabla\boldsymbol{\omega} + \mathbf{K}_\omega\boldsymbol{\omega} + \boldsymbol{\theta}_\omega\boldsymbol{\omega} \quad (9)$$

Where, $\nabla\boldsymbol{\omega}$ and $\nabla\mathbf{a}$ are IMU constant drift, $\mathbf{K}_\omega$ and $\mathbf{K}_a$ are the scale factor error matrix, $\boldsymbol{\theta}_\omega$ and $\boldsymbol{\theta}_a$ are the alignment matrix.

According to the error model, the compensated accelerometer value is

$$\mathbf{a}_1 \approx (\mathbf{I} - \mathbf{K}_a - \boldsymbol{\theta}_a)(\mathbf{a}_m - \nabla\mathbf{a}) \quad (10)$$

Likewise, the compensated gyro value is

$$\boldsymbol{\omega}_1 \approx (\mathbf{I} - \mathbf{K}_\omega - \boldsymbol{\theta}_\omega)(\boldsymbol{\omega}_m - \nabla\boldsymbol{\omega}) \quad (11)$$

3.2 Analysis of scale factor asymmetry

Due to processing technique and in-out relationship, the scale factor asymmetry is inevitable[3]. The input angle speed is set as ω_1、ω_2、ω_3,

Define that

$$sign(\omega) = \begin{cases} +1 \ if\ (\omega > 0) \\ -1 \ if\ (\omega < 0) \end{cases} \quad (12)$$

The gyro scale factor error matrix is

$$K_g = \begin{bmatrix} k_{g1}^+ + k_{g1}^- \cdot sign(\omega_1) & 0 & 0 \\ 0 & k_{g2}^+ + k_{g2}^- \cdot sign(\omega_2) & 0 \\ 0 & 0 & k_{g3}^+ + k_{g3}^- \cdot sign(\omega_3) \end{bmatrix} \quad (13)$$

Where, k_{gi}^{+} is the gyro-i scale factor error (i=1,2,3); k_{gi}^{-} is gyro-i asymmetric scale factor error. The navigation coordinate is set as east-north-up coordinate, and rotates along the vertical direction at ω, so the input angular velocity is

$$\begin{cases} \omega_1 = \omega_{ieN}\sin(\omega t) \\ \omega_2 = \omega_{ieN}\cos(\omega t) \\ \omega_3 = \omega + \omega_{ieU} \end{cases} \qquad (14)$$

Where, ω_{ieN} is northward component of earth's rotation rate, and ω_{ieU} is the upward component.

At t moment, the angular velocity error caused by gyro scale factor is

$$\delta\omega_{ib}^{b} = \begin{pmatrix} k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1) & 0 & 0 \\ 0 & k_{g2}^{+} + k_{g2}^{-}\cdot sign(\omega_2) & 0 \\ 0 & 0 & k_{g3}^{+} + k_{g3}^{-}\cdot sign(\omega_3) \end{pmatrix} \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \qquad (15)$$

Accordingly, the error model in n coordinate is

$$\begin{aligned} \varepsilon_E^n &= C_b^n \delta\omega_{ib}^b \\ &= \begin{pmatrix} \cos(\Omega) & -\sin(\Omega) & 0 \\ \sin(\Omega) & \cos(\Omega) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1) & 0 & 0 \\ 0 & k_{g2}^{+} + k_{g2}^{-}\cdot sign(\omega_2) & 0 \\ 0 & 0 & k_{g3}^{+} + k_{g3}^{-}\cdot sign(\omega_3) \end{pmatrix} \cdot \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \\ &= \begin{pmatrix} \omega_1\cos(\Omega)(k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1)) - \omega_2\sin(\Omega)(k_{g2}^{+} + k_{g2}^{-}\cdot sign(\omega_2)) \\ \omega_1\sin(\Omega)(k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1)) + \omega_2\cos(\Omega)(k_{g2}^{+} + k_{g2}^{-}\cdot sign(\omega_2)) \\ \omega_3(k_{g3}^{+} + k_{g3}^{-}\cdot sign(\omega_3)) \end{pmatrix} \\ &= \begin{pmatrix} 0.5\omega_{ieN}\sin(2\Omega)(k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1) - k_{g2}^{+} - k_{g2}^{-}\cdot sign(\omega_2)) \\ 0.5\omega_{ieN}(1+\cos(2\Omega))(k_{g1}^{+} + k_{g1}^{-}\cdot sign(\omega_1) + k_{g2}^{+} + k_{g2}^{-}\cdot sign(\omega_2)) \\ (\omega+\omega_{ieU})(k_{g3}^{+} + k_{g3}^{-}\cdot sign(\omega_3)) \end{pmatrix} \end{aligned} \qquad (16)$$

The integrated angle error over the T period is

$$\int_0^T \varepsilon^n \cdot dt = \begin{pmatrix} 0 \\ 0.5(k_{g1}^{+} + k_{g2}^{+})\omega_{ieN}T \\ (\omega + \omega_{ieU})(k_{g3}^{+} + k_{g3}^{-}\cdot sign(\omega_3))T \end{pmatrix} \qquad (17)$$

In Eq. (17), the impact of symmetrical scale factor error is not modulated.

3.3 Analysis of installation error influence

The installation error reflects the transformation relationship between s coordinate and non-orthogonal installation coordinate (g coordinate), and the matrix is[4]

$$C_g^s = \begin{bmatrix} 1 & -\tau_{yz} & \tau_{zy} \\ \tau_{xz} & 1 & -\tau_{zx} \\ -\tau_{xy} & \tau_{yx} & 1 \end{bmatrix} \qquad (18)$$

Take the gyroscope for example, the angular velocity error of rotating SINS caused by installation error can be expressed as:

$$\Delta\omega_{ib}^{b} = \begin{bmatrix} -\tau_{xz}\sin\Omega t & -\tau_{yz}\cos\Omega t & \tau_{zy}\cos\Omega t + \tau_{zx}\sin\Omega t \\ \tau_{xz}\cos\Omega t & -\tau_{yz}\sin\Omega t & \tau_{zy}\sin\Omega t - \tau_{zx}\cos\Omega t \\ -\tau_{xy} & \tau_{yx} & 0 \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \qquad (19)$$

Where, $\begin{bmatrix} \omega_x & \omega_y & \omega_z \end{bmatrix}^T$ is the actual output of the gyro, $\Delta\omega_{ib}^{b}$ is error expression. Due to the greater value of the turntable speed, bigger error may occur in $\tau_{zy}\cos\Omega t + \tau_{zx}\sin\Omega t$、$\tau_{zy}\sin\Omega t - \tau_{zx}\cos\Omega t$.

4 Analysis of Environment function

➢ *Error separation*

Part of IMU error can not be modulated, and the system performance is impacted greatly. So, an error separation method is proposed for identification and compensation[5]. The IMU error of SINS is set as observations, the projection of IMU error coefficients to IMU error are set as environment function matrix, and the IMU error coefficients are dependent variable. The corresponding vector-matrix equation is

$$\delta = \mathrm{S} \times D_{12\times 1} \qquad (20)$$

Where, $D=[g_1^{+}, g_1^{-}, \tau_{yx}, \tau_{zx}, \nabla\omega_x, g_2^{+}, g_2^{-}, \tau_{zy}, \tau_{xy}, \nabla\omega_y, g_3^{+}, g_3^{-}, \tau_{xz}, \tau_{yz}, \nabla\omega_z]^T$, g_i (i-1,2,3) represents scale factor, τ_{ab} (a=x,y,z; b=x,y,z) is alignment coefficient, $\nabla\omega$ is constant drift, δ is IMU error, S is environment function matrix establishing the relation between IMU error and error coefficients.

Random error and kinds of outliers exist in

the measurement, which lead to some large error. Instead of system error, IMU error is used as observations for identification. Better than general identification method, it can avoid the discordance between theoretic navigation performance and real IMU accuracy.

The theoretic angular velocity output ω_{ib}^{b} means the projection of the angular velocity on the body coordinate and the velocity is that of the body coordinate relative to the inertial coordinate. According to the navigation principle, ω_{ib}^{b} can be inverse solved by attitude information:

$$\begin{cases} \omega_{nb}^{b} = \begin{bmatrix} \cos\gamma & 0 & \sin\gamma\cos\theta \\ 0 & 1 & -\sin\theta \\ \sin\gamma & 0 & -\cos\gamma\cos\theta \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ \dot{\gamma} \\ \dot{\psi} \end{bmatrix} \\ \omega_{ib}^{b} = \omega_{nb}^{b} + (\omega_{ie}^{b} + \omega_{en}^{b}) \end{cases} \quad (21)$$

Where, ω_{ib}^{b} represents theoretic gyro value, ω_{nb}^{b} is body-coordinate projection of body coordinate relative to navigation coordinate, ω_{en}^{b} is body-coordinate projection of navigation coordinate relative to earth coordinate, ω_{ie}^{b} is body-coordinate projection of earth coordinate relative to inertial coordinate, and calculated by vehicle velocity, latitude, height and earth's rotation rate, $\begin{bmatrix} \dot{\theta} & \dot{\gamma} & \dot{\psi} \end{bmatrix}^{T}$ are attitude rate received by attitude.

➢ *Matrix solution*

Generally, parameters are estimated by least square estimation:

$$D = (S^{T} \cdot S)^{-1} \cdot S^{T} \cdot \delta \quad (22)$$

Parameters have dependence, s matrix is singular, and the regression model is sick. In order to solve the problem, biased estimation is introduced into the regression analysis. On the condition of keeping the $(S^{T} \cdot S)$ invariant, kI matrix is used to modify the Eigen values and the sickness of $(S^{T} \cdot S)$ matrix, and then D is

$$D = (S^{T} \cdot S + kI)^{-1} \cdot S^{T} \cdot \delta \quad (23)$$

Where k is set as a small quantity according to gyro precision.

The flow chart of error separation is shown as Fig 3:

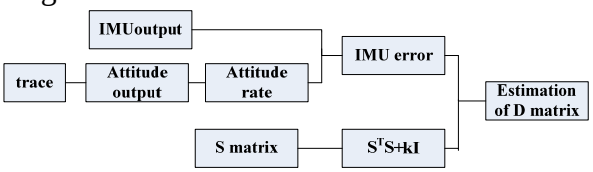


Fig 3 Flow chart of error separation

➢ *Segmented calibration*

The IMU error characteristics of non-linear and asymmetric are analyzed, segmented calibration is studied to truly reflect the IMU error coefficients. Meanwhile, the too large number of segmentation is not suitable because too many velocity dots may lead to larger amount of calculation.

For m sections, (m-1) dots are required, the issue is to be resolved: how to choose the segmented point to the best effect.

In this paper, m is set 4.

Considering that scale factor has positive-negative asymmetric features, zero point is selected as one dot to calibrate the positive and negative factor respectively.

According to negative gyro coefficients, negative velocity point is $\omega_{in(j)}$, and gyro output is $\omega_{on(j)}$,(j=1,2,…m), then the four polynomial of in-out put relationship is fit as:

$$\begin{bmatrix} \omega_{on(1)} \\ \omega_{on(2)} \\ \vdots \\ \omega_{on(m)} \end{bmatrix} = \begin{bmatrix} 1 & \omega_{in(1)} & \omega_{in(1)}^{2} & \omega_{in(1)}^{3} & \omega_{in(1)}^{4} \\ 1 & \omega_{in(2)} & \omega_{in(2)}^{2} & \omega_{in(2)}^{3} & \omega_{in(2)}^{4} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \omega_{in(m)} & \omega_{in(m)}^{2} & \omega_{in(m)}^{3} & \omega_{in(m)}^{4} \end{bmatrix} \begin{bmatrix} k_{n0} \\ k_{n1} \\ k_{n2} \\ k_{n3} \\ k_{n4} \end{bmatrix} \quad (24)$$

By least square solution, $k_{n} = (\omega_{in}^{T}\omega_{in})^{-1}\omega_{in}^{T}\omega_{on}$. Than the fit gyro polynomial is

$$\omega_{on} = k_{n0} + k_{n1}\omega_{in} + k_{n2}\omega_{in}^2 + k_{n3}\omega_{in}^3 + k_{n4}\omega_{in}^4 \qquad (25)$$

The gyro constant drift can be equal to k_{n0}. In calibration process, sub-point of great-change factor may have best effect, so the fast-changing first derivative sub-point is selected as one point, which is equal to the point that second derivative is extreme point, and the third derivative is zero. Then, the optimal negative sub-point is

$$\omega_{inp} = -k_{n3} / 4k_{n4} \qquad (26)$$

5 Simulation

5.1 Experiment of error separation and compensation

Based on error separation theory, IMU error coefficients are separated into constant drift, positive-negative scale factor, and installation error coefficients by environment function method. The real gyro outputs are compensated, and the position performance after compensation is compared with that before compensation. Here, gyro is VG951, and accelerometer is JN-06A. Through 180-seconds separation, the samples are calculated, and the corresponding error coefficients are received by least square algorithm in table 1.

Table 1. Gyro error coefficients

gyro		x	y	z
Constant drift		$-3.69*10^{-5}$	-0.074	$-2.56*10^{-5}$
scale factor	k_g^+	$1.60*10^{-9}$	-0.076	$-5.78*10^{-8}$
	k_g^-	$-3.19*10^{-9}$	-0.024	$-4.34*10^{-8}$
Alignment factor	τ_{xy}	$-2.46*10^{-4}$		
	τ_{xz}	$4.17*10^{-8}$		
	τ_{yx}		$-8.42*10^{-7}$	
	τ_{yz}		$-8.42*10^{-5}$	
	τ_{zx}			$1.73*10^{-9}$
	τ_{zy}			$7.67*10^{-4}$

The contrast curve of the position error before and after compensation is shown in Fig 4, which is navigated after 600 seconds.

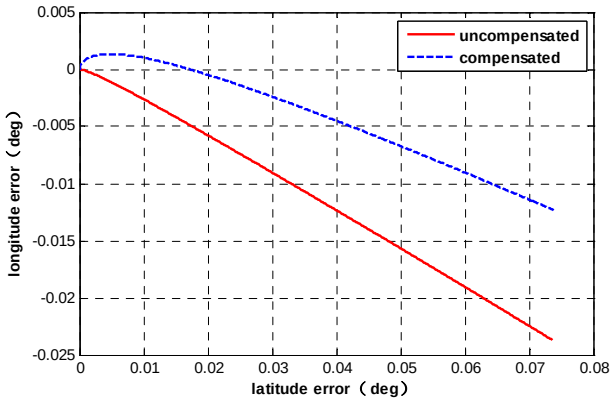


Fig 4. Position error comparison

In Fig 4, the blue line represents the results by compensation. The plot indicates that the position error is improved by compensation.

5.2 Experiment of error fractionation and Calibration

Based on the error separation results, segmented points are computed, the error coefficients in different segments are received and compensated. Through 600-seconds navigation, the results show that, the attitude accuracy is improved in Fig. 5 and Fig. 6. As IMU precision is high, and the vehicle is static, the improvement of position performance is not clear.

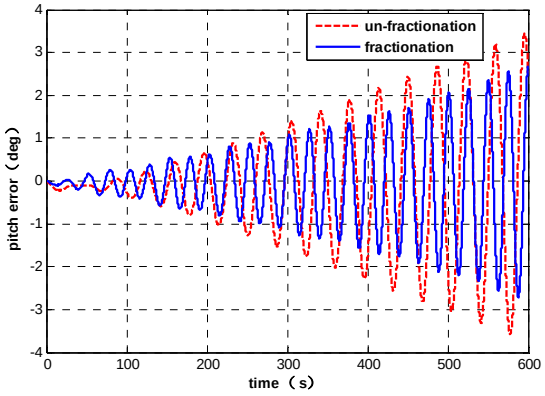


Fig5. Comparison of roll error

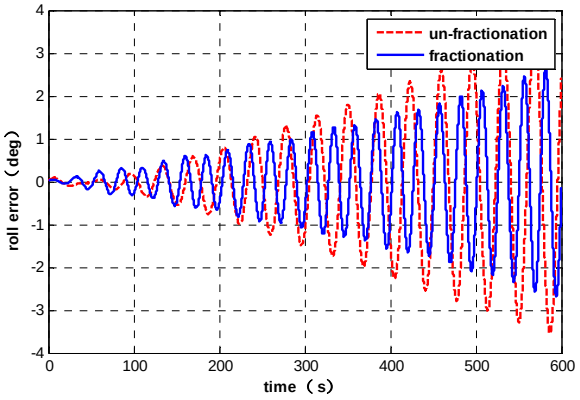


Fig 6. Comparison of pitch error

Fig. 5 represents roll error results, and Fig. 6 represents pitch error comparison. IMU rotates round the up-axis at the speed of 10deg/s. At 588.3 second, the value in blue line is -1.651, and that in red line is -3.407. From these pictures, they indicate that segmentation is effective, and the improvement after segmentation is better than that without segmentation.

From the above simulation analysis, it shows that:

(1) Through bias-estimation treatment, the environment function matrix is improved, and the singularity of $S^T S$ matrix is greatly improved;

(2) When the bias-estimation value increases, the condition for $S^T S$ matrix is reduced;

(3) The error separation effect is getting better when K is increasing; if k is bigger to a certain degree, the effect is worse. When k is small, the calculate error is dominant; when k is further bigger, the departure between bias-estimation and the real parameters is dominant.

(4) Through segmented calibration, the error message in each section is adequately reflected in the separation process, the error coefficients are more close to the real error situation, and the compensation effect is better than that without segmentation.

(5) If the segments are more, the information and the navigation is more complex. According to high-precision system, the effect is not clear with much more segments.

6 Conclusion

With the help of rotating structure, part of IMU error vertical to rotating axis is modulated, and the impact on SINS is partly reduced. However, the un-modulated IMU error is impacted by rotating velocity error, which brings about worse system performance. So, the positive-negative scale factor error and installation error are studied, the system impact is analyzed and the error modals are built in detail. Accordingly, a method called environment function is proposed to separate IMU error coefficients, the bias-estimation is put forward to deal with environment function matrix as well as improve its nonsingularity, and the segmented calibration is raised to discretely calibrate and compensate error coefficients. Experiments show that the proposed methods such as environment function based on rotating SINS, bias-estimation and segmented calibration are effective for IMU error coefficients identification, separation and compensation, which is of great theoretic and application value.

Reference

[1] Jizhou Lai, Jianye Liu, Yongrong Sun, Ling Zhang, “The Research on Errors-Modulated and Modeling Technique in RFSINS”, The Royal Navigation Conference, London, England, 2007, 1-7.

[2] Liu Jian-ye, Zhao Wei, Xiong Zhi, etc. Principle and Application of Navigation System1st ed., Northwestern Polytechnic University Press, Xi’an, Mar 21, 2010, Chaps. 2, 4.

[3] Yuan Bao-lun, “Research of Rotating Navigation System Based on Four Mode Differential Laser Gyro,” Ph.D.[D]. Dissertation, The college of opto-electric science and engineering,University of Defense Technology, China, 2008.

[4] Lai jizhou, Liu jianye, Lv pin, Zhang ling, “Analysis of error characteristic of the single axis rotating SINS”, IGNSS Symposium, Holiday Inn Surfers Paradise, Qld, Australia, 2009, PP.1-9.

[5] Sha lei, Zeng qinghua, Liu jianye, Huang lei, “Study on SINS Guidance Instrument Error Identification”, Inertial Technology Conference, Jiangsu, China, 2009, PP.1-5.

CEAS 2011

Aviation Operations

Environmental impact assessment of aircraft operation: A key for greening the aviation sector

Hannes Krieg and Robert Ilg
Dept. Life Cycle Engineering, Chair for Building Physics, University of Stuttgart, Germany

Keywords: *LCA, aircraft operation, environmental impacts, GWP*

Abstract

In this paper, the ecological impacts of the use phase of passenger aircrafts are assessed. The three modules aircraft operation, aircraft maintenance and infrastructure operation are analysed through an LCA.
Based on the data collection, an inventory model was created using the GaBi software. The functional unit of this study is the transport of one passenger over the entire flight distance, here for the case study of a medium haul flight over a distance of 2,500 km.
It was found that aircraft operation is responsible for the greatest share of impacts, causing more than 90 % of total impacts in each of the analysed impact categories. Impacts of infrastructure operation and aircraft maintenance are each responsible for 1 to 4 % of all impacts in the respective impact categories.

1 Background

Due to globalization and the related growth of international interconnectedness of people and businesses, transportation of persons and cargo is gaining more and more importance. The aerospace industry is an important element of transportation and a driving force for economic growth due to its abilities to cover long distances in short time. The importance of the sector can also be seen in its expected growth: In 2020, 300 million passengers are expected in Germany, corresponding to a growth of 60 % compared to 2009 [25]. Other forms of transportation are expected to grow slower, with an expected increase of 20 % for road and 22.5 % for rail transportation [1]. Subsequently, air transport will be more significant in comparison to other means of transportation than it is already today.

Apart from the economic relevance of the sector, aviation impacts the environment and therefore is subject to criticism [3]. Due to the expected above-average growth, those impacts will also increase. At the same time, public concern about the environmental impacts of products and processes increases. Impacts and their objective documentation are well-established nowadays as differentiation factors for products or services and sales arguments for consumers [23], also in the aerospace industry. This can be seen by the increasing popularity of emission calculators, as well as in the growing efforts of airlines to document and improve their environmental impacts (see e.g. [22]). Those impacts are also dealt with in politics, with air travel being subject to the European Union Emission Trading System (EU ETS) starting in 2012 [9]. As a result, emissions will become a cost factor and thereby a non-negligible element for planning of involved companies. Therefore, the assessment of environmental impacts will gain more importance, on the one hand to quantify the emissions and thereby resulting

costs for emission certificates, on the other hand to satisfy the public demand for information concerning the environmental impacts of air traffic. Detailed knowledge of environmental impacts will be of great strategic relevance.

To provide necessary information on environmental impacts, a Life Cycle Assessment (LCA) according to ISO 14040[18] and ISO 14044 [19] of the use phase of a passenger aircraft is conducted. This study covers operation of the aircraft, aircraft maintenance and infrastructure operation.

2 Methodology

2.1 Goal and Scope

The goal of this study is to assess the environmental impacts of the use phase of an aircraft system. An LCA is conducted for the exemplary case study of a medium haul flight. A complete flight cycle, including operation and maintenance of the aircraft as well as infrastructure operation, is analyzed.

Aircraft operation covers the flight cycle of the aircraft and provision of fuel. Aircraft maintenance includes service and overhaul. Infrastructure operation covers operation of the airport, maintenance of buildings, runways and taxiways, disposal of waste as well as the fuel consumption of airport service vehicles. Production and end of life of aircraft and infrastructure are not subjects of this study.

2.2 System Boundary and Data Origin

2.2.1 Functional Unit

This study focuses on passenger transportation; impacts are related to passenger kilometers. The functional unit is a full flight cycle, consisting of the transportation of one passenger over the entire flight distance as well as the impacts of aircraft maintenance and infrastructure use per passenger.

2.2.2 System Boundary

As mentioned, only the use phase of air travel is assessed. The system boundary for this study is shown in Fig. 1.

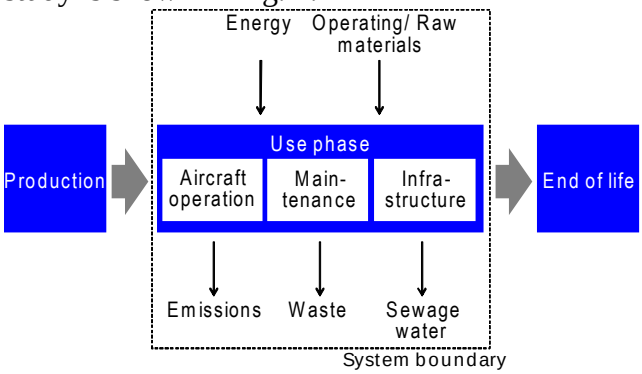


Fig. 1 System Boundaries for the use phase of an aircraft [20]

For this study, a medium range aircraft with a capacity of 150 passengers and a maximum range of 6.150 km is used [2]. For such an aircraft, an average 97 % of payload is due to passengers and their luggage, while the remaining 3 % is from cargo. Consequently, 97 % of the impacts during aircraft operation are allocated to passenger transportation, as well as 97 % of maintenance impacts [5].

The seat load factor is of great importance, as it determines the number of passengers in the aircraft, among which the environmental impacts are allocated. The average seat load factor was 68.1 % for European flights in 2008 [22]. Flight distance is assumed to be 2,500 km, which is about the distance from Oslo, Norway to Madrid, Spain. Also, it is assumed that kerosene is used as fuel.

Depending on flight distance, the flight pattern varies strongly, as shown schematically in Fig. 2. Here a medium haul flight (shown in black) was assessed. In general, the flight is separated into two main phases, the landing and take-off cycle (LTO), which includes all operations below an attitude of 1,000 m, and the cruise phase, which describes all operations above this limit. Engine performance differs drastically between the flight phases: while it is low during taxi or descent, it increases up to 80 % during climb-out and up to 100 % during take-off [7]. The calculation of fuel consumption is based on information taken from the Emission Inventory Guidebook (EIG) of the European Environment Agency [10].

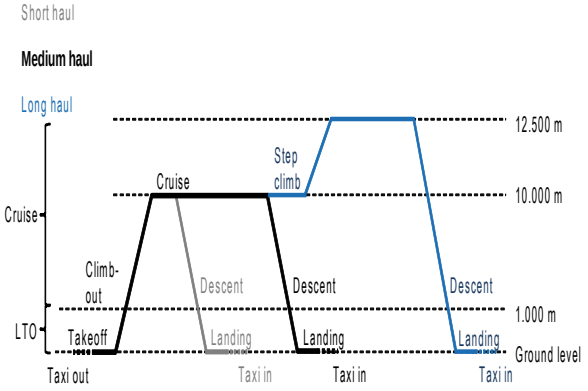


Fig. 2 Schematical flight patterns for the use phase of an aircraft [20]

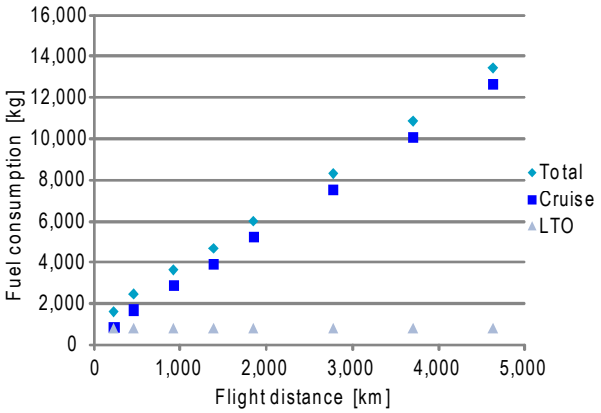


Fig. 3 Fuel consumption in relation to distance travelled

The take-off phase and the following climb consume much fuel, however, while cruising at a constant altitude the fuel consumption is relatively low. There is no correlation between fuel consumption during LTO and the flight distance; however its proportion of the total consumption decreases as the flight distance increases. As the share of fuel consumption during LTO decreases, the total fuel consumption per km decreases. Furthermore, the fuel consumption in high altitudes is lower due to the reduced air resistance. This is the reason why the flight altitude at long-haul flights is gradually increased along with decreasing weight, as the step climb in Fig. 2 indicates [16].

The fuel consumption of the assessed aircraft is shown in Fig. 3. The triangles describe the fuel consumption during landing and take-off (LTO), which is constant regardless of distance travelled [10]. The squares show the fuel consumption for cruise operation in relation to distance travelled. The diamonds show the sum of consumption during LTO and cruise. Fuel consumption per km is higher for short distances, as LTO has a larger share of total consumption. This share decreases when the distances increase, thereby reducing the consumption per km travelled.

Emission factors are partly based on the values measured in the EIG. Further data is based on statements provided by the German Federal Environment Agency on the emissions for the combustion of kerosene [4].

While the Radiative Forcing Index (RFI) is used by some emission calculators [3], the Intergovernmental Panel on Climate Change (IPCC) stated in 2007 that it "should not be used as an emission metric since it does not account for the different residence times of different forcing agents" [14]. Therefore, this study does not use the RFI, but the Global Warming Potential for the assessment of climate impacts.

It is assumed that impacts caused by airport operations, like the use of terminals or car parking and the respective maintenance, are exclusively caused by passenger services. For maintenance of runways and taxiways, the share of cargo flights and non-commercial flights has to be excluded. To analyse the impacts of the infrastructure operation, Hamburg Airport is used as the reference airport. In 2007, it had a passenger volume of about 12.7 million, which is close to the average value of international airports in Germany of 10.7 million passengers [24]. In the first half of 2008 there were 102,148 flight operations, 88,860 of those being passenger flights, corresponding to a share of 86.8 %[1]. Since all aircraft use runways and taxiways, 86.8 % of the maintenance is assigned to passenger transportation, assuming that every flight operation has the same impact on the infrastructure. Energy consumption of Terminal buildings and fuel consumption of service

[1] Ulbrich (Hamburg Airport) 2009, personal communication.

vehicles are fully allocated to passenger transport.

Data concerning the airport operation is based on the environmental reports of Hamburg Airport [11],[12],[13] and data provided on demand[2]. There was no possibility to obtain data on airport maintenance. Based on information on construction materials and life expectancy, an annual infrastructure maintenance effort of 5 % of construction efforts was assumed as suggested by Chester [5]. For each flight cycle, two airport operations are assessed, one for take-off, and one for landing.

Due to the nondisclosure of aircraft manufacturers and operators and the complexity of systems, no data are available concerning the maintenance of aircraft, neither for life expectancy of parts, nor for the production of spare parts. Due to this lack of reliable data, results from a previous study conducted by Chester [5] are used.

The structure of the underlying model in the GaBi software [17] is shown in Fig. 4.

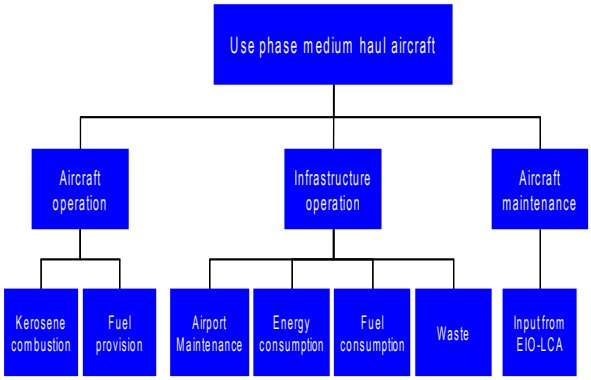
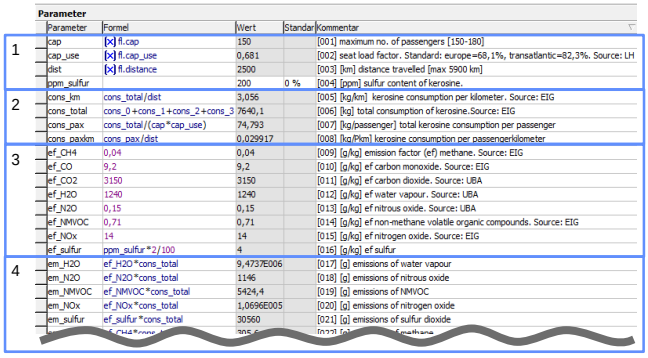


Fig. 4 Life Cycle Model structure of the aircraft use phase

While environmental impacts of aircraft and infrastructure operation are calculated with a process based LCA, information on aircraft maintenance is taken from an economic input-output LCA due to the lack of sufficient data for a process-based LCA. This hybrid approach was chosen in order to estimate the respective impact share of aircraft operation, infrastructure operation and maintenance, in order to determine further need for research.

An exemplary overview over a process modelled in the GaBi software is given in Fig. 5. It shows the parameters that define the aircraft model.


Parameter

	Parameter	Formel	Wert	Standar	Kommentar
1	cap	[x] fl.cap	150		[001] maximum no. of passengers [150-180]
	cap_use	[x] fl.cap_use	0,681		[002] seat load factor. Standard: europe=68,1%, transatlantic=82,3%. Source: LH
	dist	[x] fl.distance	2500		[003] [km] distance travelled [max 5900 km]
	ppm_sulfur		200	0 %	[004] [ppm] sulfur content of kerosine.
2	cons_km	cons_total/dist	3,056		[005] [kg/km] kerosine consumption per kilometer. Source: EIG
	cons_total	cons_0+cons_1+cons_2+cons_3	7640,1		[006] [kg] total consumption of kerosine.Source: EIG
	cons_pax	cons_total/(cap*cap_use)	74,793		[007] [kg/passenger] total kerosine consumption per passenger
	cons_paxkm	cons_pax/dist	0,029917		[008] [kg/Pkm] kerosine consumption per passenger/dimeter
3	ef_CH4	0,04	0,04		[009] [g/kg] emission factor (ef) methane. Source: EIG
	ef_CO	9,2	9,2		[010] [g/kg] ef carbon monoxide. Source: EIG
	ef_CO2	3150	3150		[011] [g/kg] ef carbon dioxide. Source: UBA
	ef_H2O	1240	1240		[012] [g/kg] ef water vapour. Source: UBA
	ef_N2O	0,15	0,15		[013] [g/kg] ef nitrous oxide. Source: UBA
	ef_NMVOC	0,71	0,71		[014] [g/kg] ef non-methane volatile organic compounds. Source: EIG
	ef_NOx	14	14		[015] [g/kg] ef nitrogen oxide. Source: EIG
	ef_sulfur	ppm_sulfur*2/100	4		[016] [g/kg] ef sulfur
4	em_H2O	ef_H2O*cons_total	9,4737E006		[017] [g] emissions of water vapour
	em_N2O	ef_N2O*cons_total	1146		[018] [g] emissions of nitrous oxide
	em_NMVOC	ef_NMVOC*cons_total	5424,4		[019] [g] emissions of NMVOC
	em_NOx	ef_NOx*cons_total	1,0696E005		[020] [g] emissions of nitrogen oxide
	em_sulfur	ef_sulfur*cons_total	30560		[021] [g] emissions of sulfur dioxide

Fig. 5 Parameters of the aircraft model

The figure is subdivided in four parts: field 1 shows the free parameters which can be adjusted by the user according to the subject of the study, such as capacity of the aircraft, seat load factor, flight distance and the sulfur content of the fuel. Field 2 contains the formulas that calculate the fuel consumption for the aircraft and per passenger kilometer, based on the flight distance, seat load factor, and capacity as well as background information on measured fuel consumption. In field 3, emission factors for the combustion of kerosene in aircraft engines are given. These factors are used in field 4 to calculate the total emissions per flight.

Processes are defined by parameters as shown above. Processes are then combined on plans in the GaBi software in order to construct a product model, as shown in Fig. 6.

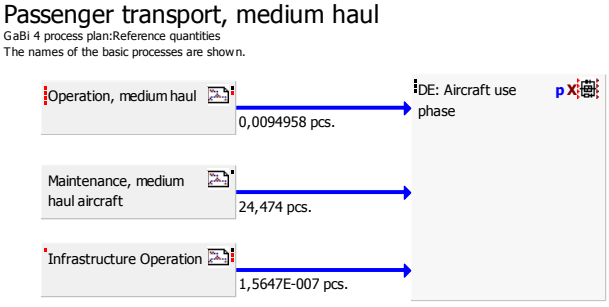


Fig. 6 GaBi plan of the transport model

Plans can again be sorted hierarchically to model complex product systems.

[2] Ulbrich (Hamburg Airport), 2009.

2.3 LCIA

The CML2001 indicators have been chosen as they are based on the most scientific and fully developed calculation methods of the several indicators. The study focuses on the following environmental indicators and impacts:

- Primary energy use – (PE) in [MJ],
- Global warming potential – (GWP) in [kg CO_2 equivalents],
- Ozone depletion potential – (ODP) in [kg R11 equivalents],
- Acidification potential – (AP) in [kg SO_2 equivalents],
- Eutrophication Potential – (EP) in [kg PO_4^{3-} – equivalents],
- Photochemical ozone creation potential (POCP) in [kg C_2H_4-equivalents].

3 Results

An overview on the environmental impacts of the assessed medium haul flight with a travel distance of 2,500 km is given in Fig**. 7**. For the sake of a better overview, the bars of the aircraft use phase are broken. As it is shown, the flight phase clearly dominates all impact categories. This is due to the relatively long distance covered and the correspondingly high fuel consumption of 7,640 kg of kerosene. With an assumed seat load factor of 68.1 % and an aircraft capacity of 150 passengers, an average 102 passengers are aboard the aircraft, resulting in an average fuel consumption of 74.8 kg per passenger. Total emissions are shown on top of the bars. The flight phase causes between 92 % and 98 % of all impacts in the respective impact categories. Infrastructure operation is responsible for over 4 % of PE, also causing 4 % of all greenhouse gas emissions. For the other impact categories, the contributions of infrastructure operation are around 2 %. Aircraft maintenance contributes around 2 % to PE, GWP and POCP, while the contributions to AP and EP are both around 1.5 %. Infrastructure operation contributes 2–3 % to all impact categories apart from EP, where the contribution is below 1 %.

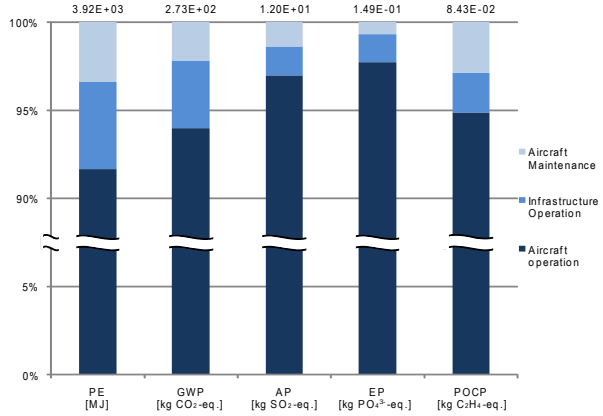


Fig. 7 Environmental indicators and impacts of a medium haul flight

It is apparent that the emissions of the flight phase dominate all impact categories, causing at least 92 % of the total impact in the analyzed impact categories. Nevertheless, both aircraft maintenance and infrastructure operation are still relevant contributors to the total impacts.

4 Conclusion

It becomes apparent that the flight phase dominates the environmental impact in all impact categories. Fuel consumption per passenger kilometer is comparable to other modes of transport, but distances are typically much greater for air travel than for ground travel. Air travel infrastructure is only used during landing and take-off, while ground travel infrastructure is used continuously during the entire journey. Due to this large share, the best leverage to reduce total environmental impacts is the flight phase. Currently, several projects and efforts aimed at reducing this impact are being conducted. Within the CleanSky project of the EU, engines with lower fuel consumption and lower emissions are investigated, as well as lightweight materials [6]. Another European project, Single European Sky, aims to improve cooperation of air traffic controllers in Europe. It is expected that by reducing holding patterns and improving flight routes, fuel consumption can be reduced [8].

Furthermore, an increased seat load factor would reduce the fuel consumption per passenger, allowing to decrease the total number of flights necessary to satisfy the demand for

transportation. For example, an increase of the seat load factor to 80 % would reduce the total use phase impacts per passenger by about 15 %.

Another approach to reduce PE and contributions to GWP is the use of biofuels. First tests of airlines are already being conducted. The fuels used are so called drop-in fuels that can be combusted in regular aircraft engines [15],[26]. While this can help in reducing PE and GWP, there is a number of other issues connected with the use of biofuels (see e.g. [21]).

The hybrid approach in this study was chosen due to the lack of data for a process-based LCA of aircraft maintenance. It provides an estimate for the environmental relevance of aircraft maintenance relative to the other modules. Given that the identified impact of aircraft maintenance is rather small, it seems justified to not elaborate further on the matter. Due to the small share of aircraft maintenance from total environmental impacts, the huge effort of data collection and process-based modeling does not seem to be justified for the expected gain of information at this time. The hybrid approach offers the possibility to fill this data gap and analyze overall use phase impacts on a sound basis.

It is obvious that the awareness for the importance of reducing environmental impacts and at the same time reducing cost and fuel consumption in the aerospace industry has been growing during the last years. Further studies should aim at quantifying the environmental impact of aircraft production, maintenance and end of life, as well as the impacts of the use of biofuels in aerospace applications. This paper provides an overview on the environmental impacts of the use phase and aims at providing a basis for further research in this sector.

References

[1] Acatech (2008) Dossier Mobilität. http://www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Aktuelles___Presse/Dossiers/Dossier_Mobilitaet/Dossier_Mobilitaet.pdf – Accessed 2011-02-22

[2] Airliners (2009) Flugzeuglexikon. http://www.airliners.de/magazin/flugzeuglexikon/flugzeugdetails/typ/a320/id/9/mid/4 - accessed 2011-02-24

[3] Atmosfair (2009) Einfluss des Flugverkehrs auf das Klima. http://www.atmosfair.de/index.php?id=60&L=3index.php%3Fid…in%2Fzero_vote%2Fevilx%3F%2Finclude%2Fprint_category.php%3Fsetup%5buse_category%5d%3D1 – Accessed 2011-02-22

[4] AvioPlan (1999) Modellsystem zur routinemäßigen Ermittlung umweltoptimierter Flugstrecken als Beitrag zum Schutz des Klimas. UBA-Forschungsbericht 29641838, März 1999

[5] Chester, M. V. (2008) Life-cycle Environmental Inventory of Passenger Transportation in the United States. Dissertation, University of California, Berkley

[6] Clean Sky (2009) What is Clean Sky JTIP. URL: http://www.cleansky.eu/index.php?arbo_id=35 – Accessed 2011-02-22

[7] DLR (2000) Aviation Emissions and Evaluation of Reduction Options (AERO). Main Report. Part I: Description of the AERO Modelling System

[8] European Comission (2009) Single European Sky. http://ec.europa.eu/transport/air/single_european_sky/single_european_sky_en.htm – Accessed 2011-02-22

[9] European Commission (2008) DIRECTIVE 2008/101/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 19 November 2008 amending Directive 2003/87/EC so as to include aviation activities in the scheme for greenhouse gas emission allowance trading within the Community. http://www.eea.europa.eu/policy-documents/directive-2008-101-ec – Accessed 2011-02-22

[10] European Environment Agency (2007) Joint EMEP/CORINAIR Atmospheric Emission Inventory Guidebook. Third Edition

[11] Flughafen Hamburg (2008) Umwelterklärung 2008

[12] Flughafen Hamburg (2009a) Der Neue Hamburg Airport stellt sich vor http://www.ham.airport.de/de/u_ham21_airport_stellt_sich_vor.html – Accessed 2011-02-22

[13] Flughafen Hamburg (2009b) Flughafengelände. http://www.ham.airport.de/de/u_daten_flughafengelaende.html – Accessed 2011-02-22

[14] Forster, P.; Ramaswamy, V., et al. (2007) Changes in Atmospheric Constituents and in Radiative Forcing. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change

[15] Green Energy News (2008) Virgin Atlantic. First Airline to Fly on Biofuel Blend. http://www.green-energy-news.com/nwslnks/clips208/feb08023.html – Accessed 2011-02-22

[16] Heinlin, H. (1995-2009) Flight planning and performance. URL: http://aerowinx.de/downloads/PlanPerf.pdf – Accessed 2011-02-22

[17] PE, LBP (1992-2011) GaBi 4 Software-systems and Databases for Life Cycle Engineering

[18] ISO 14040:2006 Environmental management – Life cycle assessment – Principles and framework

[19] ISO 14044:2006 Environmental management – Life cycle assessment – Requirements and guidelines

[20] Krieg, H; Ilg, R.; Sedlbauer, K. (2009) Ökobilanzielle Betrachtung der Nutzungsphase im Luftfahrtsektor. Diploma Thesis, Universität Stuttgart

[21] Lindner, JP (2008) Development of an indicator for biodiversity consideration in life cycle assessment. Master's thesis, Universität Stuttgart

[22] Lufthansa (2008) Balance. Das Wichtigste zum Thema Nachhaltigkeit

[23] Meffert, H.; Kirchgeorg, M. (1998) Marktorientiertes Umweltmanagement. Konzeption, Strategie, Implementierung. Schäffer-Poeschel, Stuttgart

[24] Statistisches Bundesamt (2009) Luftverkehr auf ausgewählten Flughäfen. Fachserie 8, Reihe 6.1

[25] Welt online (2009) Luftverkehr: Passagierzahlen sollen trotz Finanzkrise steigen. http://www.welt.de/reise/article3614109/Passagierzahlen-sollen-trotz-Finanzkrise-steigen.html – Accessed 2011-02-11

[26] Wired (2010) Lufthansa Plans Biofuel-Powered Passenger Flights. http://www.wired.com/autopia/2010/11/lufthansa-to-begin-scheduled-biofuel-passenger-flights/ - Accessed 2011-02-15

Analysis of Rates of Aircraft Operational Effectiveness

J. Żurek and M. Zieja
Air Force Institute of Technology, Poland

Keywords: *aircraft, effectiveness, availability, readiness, maintainability*

Abstract

The technical readiness of the aircraft for flying is measured with the probability of its staying, at any time t, in the sub-set of states that offer capabilities of being used and accomplishing the intended task. The technical readiness is a conditional rate of the aircraft's availability to perform the task in time t, on condition that other requirements decisive of aircraft's capabilities to perform its actions are satisfied. The rate of the object's readiness found with the method described in the paper expresses the potential to effectively perform its functions. The comparison between this rate and the aircraft fill rate (the degree of having available seats filled with passengers) will give a general rate of effectiveness that describes the aircraft availability and market-defined demand for a passenger aircraft.

1 Introduction

Any aircraft operated within some logistics system shows, as any engineered object, the following general features:

- it serves a particular purpose (a set of tasks to be accomplished);
- it shows some limited usability (some limited potential for operation);
- it needs maintenance, materials, power supplies, information, etc.;
- it is defined with some limited lifetime.

The process of aircraft operation consists in using its functioning capability created in the course of manufacture, and in the periodically performed restoration of this capability to keep the aircraft capable of further operational use.

Among features most essential to the process of aircraft operation one should mention safety of operational use and effectiveness. The complexity of the safety and effectiveness dedicated issues and a great variety of definitions in use need any individual type of a system to be approached separately, and objectives for which the effectiveness measures are determined – to be well defined.

The aircraft safety depends, among other things, on the intensity of destruction processes that affect particular structural components, functional systems and systems that generate failures/damages, including the frequency with which a system/component fails, i.e. the failure rate.

Factors that prove decisive of the aircraft operational effectiveness within some logistics system may be as follows:

- readiness to perform air tasks,
- cost of maintenance and repairs,
- degree of the aircraft systems' engagement while performing the tasks.

2 Effectiveness of the Process of Aircraft Operational Use

2.1 Analysis of Aircraft Readiness

The technical readiness of the aircraft for flying is measured with the probability of its staying, at any time *t*, in the sub-set of states that offer capabilities of being used and accomplishing the intended task. It may also be measured with the expected value of time of the aircraft's staying in some selected states, and with other measures, depending on the objective of the analysis of the process of aircraft operation. The technical readiness is a conditional rate of the aircraft's availability to perform the task in time *t*, on condition that other requirements decisive of aircraft's capabilities to perform its actions are satisfied (e.g. those regarding fuel and gases supplies, furnishing the aircraft with actuators/start units, etc.).

Determination of the probability of an aircraft dwell in the availability state for a 4-state operation process model can serve as an example how the analysis method can be useful:

H_1 – the operation status (flight),
H_2 – the stand-by status (maintaining a steady readiness),
H_3 – the status of prophylactic and scheduled repairs,
H_4 – the status of random repairs.

Time periods of aircraft dwelling at individual operational statuses are random variables T_{H_i} of known distributions:

- the cumulative distribution function of the duration time of the *k-th* flight *k* = 1,2,3,...;

$$F_{kH_1}(t) = P\{T_{kH_1} < t\} \qquad (1)$$

- the cumulative distribution function of the time of the *k-th* stand-by *k* = 1,2,3,...;

$$F_{kH_2}(t) = P\{T_{kH_2} < t\} \qquad (2)$$

- the cumulative distribution function of the time of scheduled services and repairs;

$$F_{kH_3}(t) = P\{T_{kH_3} < t\} \qquad (3)$$

- the cumulative distribution function of random repairs.

$$F_{kH_4}(t) = P\{T_{kH_4} < t\} \qquad (4)$$

For an object realizing a multiple process of operation the readiness to take on and carry out tasks can be determined by the method of the semi-Markov's analysis of a finished number of phase conditions progressing within a continuous time $t \geq 0$.

In order to be able to use the semi-Markovs model reflecting the process of using an object the conditions as follows must be met:

- all transitions of an object from an operational status $i \in E$ to an operational status $j \in E$ occur in a sudden form.
- the dwell time of an objest in an „*i*" state prior to transition into an „*j*" state is a random variable T_{ij} taking on values independent of the calender time.
- random variables T_{ij} are independent for every $i, j = 1, 2, 3, \ldots, r$.

Lets discuss an aircraft operation process composed of a finished set of operational statuses $E = \{1,2,...,r\}$ represented by a random process $\{Z_t: t \geq 0\}$ determined in a following way: $Z_t = i$, if at a time *t* the process is at the *i* state. $Z_t = j$, if the process at a time *t* is in the *j* state.

The dwell time of the process in the *i* state prior to transition to the *j* state is a random variable of a probability distribution function $F_{ij}(t)$, i.e. that probability function determines the said process dwell time in the *i* state prior to

transition to the j state etc. Probabilities of a transition from the state i to the subsequent state j is given by a matrix $r \cdot r$ $\boldsymbol{P} = [P_{ij}]$, which determines a so-called chain put into the Markov's process. For any $ij \in E$ there exists a determined function $Q_{ij}(t)$, which determines a joint probability of an event consisting in that, that the dwell time of the process in the i state is smaller than t and that from the i state the process will go into the j state.

This function satisfies the following conditions:

$$Q_{ij}(0) = 0; \quad i, j \in E,$$
$$\sum_{j \in E} Q_{ij}(\infty) = \sum_{j \in E} P_{ij} = 1, \quad i \in E \qquad (5)$$

The set of functions $Q_{ij}(t)$ creates the $Q(t)$ matrix of the process transition distribution $\{Z_t : t \geq 0\}$. An unconditional probability distribution of the dwell time of the process of operation in the i state is given by the $F_i(t)$ function:

$$F_i(t) = \sum_{j \in E} Q_{ij}(t) \qquad (6)$$

An r-dimensional vector of initial probabilities determines the initial status of a process:

$$\boldsymbol{P} = [P_1, P_2, ..., P_r]$$

where:

$$\sum_{i \in E} P_i = 1 \qquad (7)$$

A thus determined random process Z_t: $t \geq 0$ is a semi-Markov's one given by a ternary: (r, p, Q).

For a so determined process of operation we can create algorithms to determine many indicators of the aircraft availability with regard to a given subset of tasks or its general availability.

Since in the example under discussion the H_1 and H_2 operational statuses compose a subset H_{12} of technical readiness states the semi-Markov's model of the process of operation $\{Z_t, t \geq 0\}$ can be reduced to three states: H_{12}, H_3, H_4 of the incidence matrix:

$$\mathbf{I} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad (8)$$

and the matrix of the process transition distribution:

$$\mathbf{Q}_{ij}(t) = \begin{bmatrix} 0 & Q_{12}(t) & Q_{13}(t) \\ Q_{21}(t) & 0 & 0 \\ Q_{23}(t) & 0 & 0 \end{bmatrix} \qquad (9)$$

The $\boldsymbol{Q}_{ij}(t)$ function depends on the distributions: $G_R(t), F_\tau(t), F_{H_3}(t), F_{H_4}(t)$

Having determined the $\boldsymbol{Q}_{ij}(t)$ function and the initial distribution $\boldsymbol{P} = \{P_i, \ i = 1,2,3\}$, we can define the indicators of the aircraft technical readiness, assuming, that the H_{12} state is a subset of readiness states.

The transition from the H_{12} state into the H_3 state occurs when the following conditions are satisfied:

- the correct operation time $\tau_R \in T_R$ is greater than the expired overhaul life $\tau_\tau \in T_\tau, \tau_R > \tau_\tau$
- the operation life till the τ_τ time is not greater than t, $(S_\tau \leq t)$.

A joint event $\{(\tau_\tau \leq t) \wedge (\tau_R > \tau_\tau)\}$ occurs according to a probability function:

$$Q_{12}(t) = \int_0^t [1 - G_R(y)] dF_\tau(y) \qquad (10)$$

A transition from the H_{12} into the H_4 state occurs when events as follows occur:

- the time of proper operation is not greater than $\tau, (\tau_R \leq t)$;

- the overhaul time τ_τ is longer than the time of correct operation τ_R ($\tau_\tau > \tau_R$).

The joint event $\{(\tau_R < t) \wedge (\tau_\tau > \tau_R)\}$ has a probability function:

$$Q_{13}(t) = \int_0^t [1 - F_\tau(y)]dG_R(y) \qquad (11)$$

A transition from the H_3 state into the H_{12} state occurs when $(T_{H_3} < t)$, therefore:

$$Q_{21}(t) = P\{T_{H_3} < t\} \qquad (12)$$

and from the H_4 into the H_{12} when $(T_{H_4} < t)$, then:

$$Q_{23}(t) = P\{T_{H_4} < t\} \qquad (13)$$

Assuming that the initial state of a process is the H_{12} we obtain the determined initial vector of the process $\boldsymbol{p} = [1,0,0]$. The probability of an aircraft dwell time in the H_{12} technical readiness state over the t time is $1 - F_{12}(t)$. From the formula (14) the following value was determined:

$$P_{12}(t) = 1 - F_{12}(t) = 1 - [Q_{12}(t) + Q_{13}(t)] =$$
$$= 1 - \left\{ \int_0^t [1 - G_r(y)]dF_\tau(y) + \int_0^t [1 - F_\tau(y)]dG_r(y) \right\} \qquad (14)$$

The asymptotic probability of $P_{12}(t)$ of an aircraft dwelling in the technical readiness for $t = \tau$ (flight time) is a product of:

$$P_{12}(t) = P_{12}^{(\infty)} \frac{\int_\tau^\infty [1 - F_{12}(t)]dt}{\int_0^\infty [1 - F_{12}(t)]\,dt} \qquad (15)$$

where $P_{12}^{(\infty)}$ is a limit probability of the aircraft dwell in the H_{12} state, when the process time is very long, $t \to \infty$. The $P_{12}^{(\infty)}$ value represents the aircraft functional readiness (momentary readiness). The $P_{12}^{(\infty)}$ value is determined from the equation in [6].

If the semi-Markov's process $\{Z_t, t \geq 0\}$ of a finite set E is irreducible and non-periodical and random variables T_{Hj} have finite positive expected values $E[T_{Hj}]$, then limiting transition probabilities exist [2]:

$$p_{ij}^{(\infty)} = \lim_{t \to \infty} p_{ij}(t) \quad i, j \in E \qquad (16)$$

$$p_{ij}^{(\infty)} = p_j^{(\infty)} = \frac{\pi_j E(T_{Hj})}{\sum_{r \in E} \pi_r E(T_{Hr})} \qquad (17)$$

where π_j - denotes a limiting transition probabilities of a Markov's chain inserted into the process. This is the frequency of dwelling of an aircraft in the j-th operational state (a subset of states).

The probability π_j, $j = 1,2,3,...$ can be calculated by a method given in [1]:

$$\pi_j = \frac{D_j}{\sum_{i \in E} D_i} \qquad (18)$$

where: D_i - the principal minor of the matrix P obtained by deleting the i-th column and the j-th line denotes the matrix of transition probabilities of the Markov's chain put into the process.

Individual values of the p_{ij} in the model under discussion are:

$$p_{12} = \lim_{t \to \infty} Q_{12}(t) = \int_0^\infty [1 - G_{12}(y)]dF_\tau(y) \qquad (19)$$

$$p_{13} = Q_{13}(t) = \int_0^\infty [1 - F_\tau(y)]dG_r(y) \qquad (20)$$

$$p_{21} = \lim_{t\to\infty} Q_{21}(t) = 1 \qquad (21)$$

$$p_{31} = \lim_{t\to\infty} Q_{31}(t) = 1 \qquad (22)$$

The matrix P has the following form:

$$\mathbf{P} = \begin{bmatrix} 0 & p_{12} & p_{13} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad (23)$$

The matrix D is expressed as:

$$\mathbf{D} = \begin{bmatrix} 1 & -p_{12} & -p_{13} \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \qquad (24)$$

with :

$$\pi_1 = \frac{1}{3 - p_{12} - p_{13}} \qquad (25)$$

$$\pi_2 = \frac{p_{12}}{3 - p_{12} - p_{13}} \qquad (26)$$

$$\pi_3 = \frac{p_{13}}{3 - p_{12} - p_{13}} \qquad (27)$$

$$p_{12}^{(\infty)} = \frac{E[T_{H_{12}}]}{E[T_{H_{12}}] + p_{12}E[T_{H_3}] + p_{13}E[T_{H_4}]} \qquad (28)$$

The expected values of random variables $T_{H_{12}}, T_{H_3}, T_{H_4}$, which are average dwell times of an aircraft in individual operational states can be determined from the following formulas:

$$E[T_{H_{12}}] = \int_0^\infty t[1 - G_R(t)]dF_\tau(t) + \int_0^\infty t[1 - F_\tau(t)]dG_R(t) \qquad (29)$$

$$E[T_{H_3}] = \int_0^\infty t dF_{H_3}(t) \qquad (30)$$

$$E[T_{H_4}] = \int_0^\infty t dF_{H_4}(t) \qquad (31)$$

The expected value of the $T_{H_{12}}$ random variable is also an indicator of the aircraft technical readiness, which characterises an average dwell time in a required subset of states.

A probability of an aircraft dwell time in the technical readiness state determined from the formula (14) is true if a condition that the initial state of the operation process is a state from the set $\{H_1, H_2\}$ is satisfied.

If the operation process initial state differs from the state of readiness, one shall determine the cumulative distribution function of the T_{iA} random variable of the first transition from the initial state $i \in E$, which is not a readiness state, to the A subset – the readiness states set $A \in E$. The cumulative distribution function T_{iA} can be determined from the equation given in [3].

If the A is a subset of states achievable from every state $i \in E \wedge i \notin A$ of the semi-Markov's process the set E of finite states and the $Q_{ij}(t)$ functions are continuous, then there are cumulative distribution functions $F_{iA}(t)$ of the random variable T_{iA}, which satisfy the system of equations [5]:

$$F_{iA}(t) = \sum_{j\in A} Q_{ij}(t) + \sum_{\substack{k\notin \\ k\in E}} F_{kA}(t - x)dQ_{ik}(x) \qquad (32)$$

For the discussed model of aircraft operation:

$$F_{3A}(t) = F_{H_3}(t) \qquad (33)$$

$$F_{4A}(t) = F_{H_4}(t) \qquad (34)$$

The probability $P_{3,1,2}^{(t)}(\tau)$ of an aircraft dwell in a state of technical readiness H_{12} for the time τ after the operation duration time t, when the initial time was the state H_3, can be calculated from the following formula:

$$P_{3,1,2}^{(t)}(\tau) = \int_0^t [1 - F_{12}(t + \tau - x)] dF_{H3}(x) \qquad (35)$$

where the value $1 - F_{12}(t + \tau - x)$ is determined from the relation (14). The $P_{3,1,2}^{(t)}(\tau)$ probability characterises the aircraft technical readiness for action within the time τ after the time t from the beginning of the process starting at the state H_3.

For a process starting at the moment 0 in the state H_4 we have correspondingly:

$$P_{4,1,2}^{(t)}(\tau) = \int_0^t [1 - F_{12}(t + \tau - x)] dF_{H_4}(x) \qquad (36)$$

A more exact analysis of the aircraft technical readiness can be carried out using a more complex model [4].

2.2 Maintenance Time

In the course of preventive maintenance the condition of aircraft assemblies is carefully examined. Probability that during the maintenance a failure/damage to any assembly is found is $Q = 1 - R$. Probability that during the maintenance no failure/damage to any assembly is found is R. This means that the maintenance time is described with a function with independent variables [6]:

$$T = R\,T_{os} + (1 - R)\,T_n \qquad (37)$$

where:

T_{os} – time for maintenance of a serviceable (fit-for-use) aircraft,

T_n – time for correction of a failure/repair of an assembly.

It is assumed that the maintenance of a serviceable (fit-for-use) aircraft consists in the sequentially performed checks and servicing of particular assemblies [1].

$$T_{os} = \sum_{i=1}^{N} \tau_i \qquad (38)$$

where:

τ_i – a random variable of time of maintenance of the i-th assembly of the aircraft.

It is also assumed that in the course of maintenance of the aircraft assembly a failure may occur at the rate λ_{kj}. Reliability of the aircraft in the course of maintenance is [1]:

$$R_o = \exp\left[-\sum_{j=1}^{N} \lambda_{kj}\, \tau_j \right] \qquad (39)$$

Hence, if any symptoms of unserviceability (of becoming unfit-for-use) appear, the aircraft may be re-maintained. Such being the case, the maintenance time is to be found from the following formula [5]:

$$T_{OR} = \frac{T_{os}}{R_o} \approx T_{os}(2 - R_o) \qquad (40)$$

2.3 Repair Time and Costs

The process of repair of any engineered object is very complicated. It consists of some fundamental actions such as [6]:

$T_1 = T_{on}$ - waiting for the repair (mobilization of specific specialists, preparation of equipment);

$T_2 = T_{oc}$ - waiting for spares;
$T_3 = T_{de}$ - part/component disassembly;
$T_4 = T_{me}$ - part/component assembly;
$T_5 = T_{kz}$ - serviceability check.

The wait for the spares occurs with probability P_{12}, whereas we have the spares at our command with probability $1 - P_{12} = P_{13}$. Time of repair of the i-th assembly of the aircraft is given with the following formula [3]:

$$T_{ni} = T_{1i} + P_{12,i}\, T_{2i} + T_{3i} + T_{4i} + T_{5i} \qquad (41)$$

With the following notation introduced:
r_i - probability that the i-th assembly of the aircraf is serviceable (fit-for-use),
q_i - probability that the i-th assembly of the aircraf is unserviceable (unfit-for-use),
probabilities R and Q are found.

Probability that all the aircraft assemblies are serviceable (fit-for-use) is [3]:

$$R = \prod_{j=1}^{N} r_i \qquad (42)$$

Probability that the i-th aircraft assembly is unserviceable (unfit-for-use) is [3]:

$$Q_i = \prod_{\substack{j=1 \\ j \neq i}}^{N} r_j\, q_j \qquad (43)$$

Probability that the i-th and the a-th aircraft assemblies are unserviceable (unfit-for-use) is:

$$Q_{ia} = \prod_{\substack{j=1 \\ j \neq i \\ j \neq a}}^{N} r_j\, q_i\, q_a \qquad (44)$$

Time of repair is given with the formula [3]:

$$T_n = \sum_{i=1}^{N} Q_i\, T_{ni} + \sum_{i<j}^{N} Q_{ij}\left(T_{ni} + T_{nj}\right) + \ldots\ldots \qquad (45)$$

The cost of repair in n maintenance cycles of the interval under consideration is then [5]:

$$C_n = \sum_{i=1}^{N} S_i Q_i\, T_{ni} + \sum_{i<j}^{N} S_j Q_{ij}\left(T_{ni} + T_{nj}\right) + \ldots.. \qquad (46)$$

where: S_j – expenditures for the restoration of the j-th aircraft assembly.

The occurrence of more than one failure/damage at a time is in practice only slightly probable.

Time of accomplishing the preventive treatment is given with the following formula:

$$T_R = R\, T_{oR} + (1 - R)\, T_n \qquad (47)$$

The total cost C_l of executed flights will be:

$$C_l = \sum_{i=1}^{N} L_i \qquad (48)$$

where: L_i – expenditures for particular i-th flights, $i = 1, 2, 3, \ldots N$.

The operational and maintenance cost estimate C_o is given with the formula:

$$C_o = n\left(C_k R\, T_{oR} + (1 - R)\, C_n\right) + C_s + C_l \qquad (49)$$

where:
C_k - expenditures for the preventive maintenance of the aircraft;
C_s - other expenditures throughout the total period of aircraft operation and maintenance, related to the ground handling.

Revenue C_d from the accomplished flights is:

$$C_d = \sum_{i=1}^{u} k_i d_i \qquad (50)$$

where:
k_i - the number of passenger seats occupied in the i-th flight (I = 1, 2, 3, …u),
d_i - revenues gained from particular flights.

The degree of aircraft utilization can be measured with the flying time, in the case of a passenger aircraft – also with the degree of the filling-up of available passenger seats. The total time of aircraft utilization within the sub-set of states H_{12} up to the time instant t is a random variable S_t of distribution defined with the following relationship [1]:

$$P\{S_t < x\} = \sum_{n=1}^{\infty} F_{H_1}^{(n)}(x)\left[F_{H_2}^{(n-1)}(t-x) - F_{H_2}^{n}(t-x)\right] \qquad (51)$$

where n denotes the multiple convolution of distributions of the random variable.

The process of time of aircraft operational use increasing against the calendar time shows the asymptotically normal distribution for $t \to \infty$. The expected value of the distribution is:

$$E[S_t] = \frac{E[T_{H_1}]}{E[T_{H_1}] + E[T_{H_2}]} t = m(t) \qquad (52)$$

whereas the variance is:

$$D^2[S_t] = \frac{D^2(T_{H_1})(E[T_{H_2}])^2 + D^2(T_{H_2})(E[T_{H_1}])^2}{(E[T_{H_1}] + E[T_{H_2}])^3} t \qquad (53)$$

The utilization factor of an aircraft in the state of readiness is:

$$\rho = \frac{E[T_{H_1}]}{E[T_{H_1}] + E[T_{H_2}]} \qquad (54)$$

where:
$E[T_{H_1}], E[T_{H_2}]$ – expected values of flight and duty times.

The factor of the filling-up of available passenger seats γ is:

$$\gamma = \frac{\sum_{i=1}^{u} k_i}{u \cdot N} \qquad (54)$$

where:
N - the number of available passenger seats in the aircraft.

General coefficients of aircraft service effect are as follows:
- economic

$$E_c = \frac{C_o}{C_d} \qquad (55)$$

- operational

$$E_w = \rho \cdot \gamma \qquad (56)$$

3 Conclusions

The in the service data bank stored fractions of time in the existence of particular objects and assemblies and parts thereof, ones between some selected events or classes of events, build up suitable sets of random variables of times of the probabilistic model of the process of aircraft service. Distributions of these random variables can be described with analytical models, parameters of which characterize particular qualitative features of the process of aircraft service and properties of objects.

The measure of aircraft readiness describes the object itself and the logistics system, within which the object is embedded. The probability value that expresses the readiness translates into the percentage share of time the aircraft remains at the disposal of the transportation system manager. This value depends on the

maintenance program pre-set by the aircraft manufacturer. The program includes:

- the schedule of maintenance activities and diagnostic checks throughout the operational use of the aircraft,
- maintenance procedures and procedures of other service-oriented ventures.

The above-determined rates of the aircraft readiness within the logistics system express the potential to effectively perform its functions. The aircraft fill-up rate (the degree of having available seats filled with passengers) describes the degree to which those capabilities can be utilized. The general rate of effectiveness that describes the aircraft availability and market-defined demand for a passenger aircraft can be expressed with the product of the above-mentioned rates of readiness and fill-up.

References

[1] Barziłovic E., Mezencew V., Savenkov M., *Nadieżnost aviacionnych sistem*. Transport, Moskwa 1982.

[2] Grabski F., *Teoria sami-markowskich procesów eksploatacji obiektów technicznych*. WSMW Gdynia 1985.

[3] Woropay M., Żurek J.: *Model gotowości do działania pojazdów interwencyjnych w systemie transportowym*. Bydgoszcz-Radom 2003.

[4] Żurek J.: Ziółkowski J.: Model gotowości do działania pojazdów interwencyjnych**.** *X Kongres Eksploatacji Urządzeń Technicznych*, Stare Jabłonki 2005.

[5] Jaźwiński J., Żurek J.: *Wybrane Problemy Sterowania Zapasami.* Instytut Technologii Eksploatacji, Radom 2007.

[6] Żurek J.: *Modelowanie Nadążnych Systemów Bezpieczeństwa.* Instytut Technologii Eksploatacji, Radom 2010.

Visionary Concepts for Future European Seaplane Operations

Giangiacomo Gobbi, Ladislav Smrcek, Roderick Galbraith, Alan Canamar
Aerospace Engineering, Glasgow University, Glasgow G12 8QQ, UK

Jana Halirova
Faculty of Transportation Sciences, Czech Technical University, Prague, Czech Republic

Keywords: *seaplane operations, sustainability, concept design*

Abstract

The objective of this document is to establish a common understanding of current strengths and weaknesses of existing seaplane/amphibian transport system as well as future opportunities for a new seaplane/amphibian transport system. In this report a market analysis is discussed to show the actual point in which seaplanes stand in today's European market. Starting from this point of view, strengths and weaknesses of existing seaplane market/operation/design are taken in account in order to understand which aspects need more attention, and, on the other hand, what are the areas where existing seaplanes are already superior to other means of transportation. The European Vision 2020 for aeronautics, and the concept of sustainable aviation stated within, is here considered as guideline in the strength/weakness assessment. On the base of this analysis some suggestions are then made to obtain a better, more reliable and least expensive seaplane operation for the market today, in a near future and long term period.

1 Introduction

Seaplanes have been in the market for over 80 years. However, advance modifications or new designs have not been created since the 1950's [1]. Why is it that seaplane design had not been developed over the years? The answer is that seaplanes do not have a wide market as land planes do. The main use of seaplanes is in the private sector, and most of these designs are modified small landplanes, such as the Cessna 185, which is simply adapted with floats, to become a seaplane. The creation of new concepts is expensive and industry is not interested because the market is not very reliable. However, with the increase of tourism around the world, a new opportunity for seaplanes had arisen. Therefore, some natural tourist places are inaccessible to arrive by other means of transportation such as cars, buses, trains and even landplanes. So the solution to this problem is seaplanes. Not only tourism benefits from seaplanes. In North America, especially in Canada (e.g. Figure 1), the large number of bodies of water and the remoteness of many important locations has produced a healthy seaplanes culture [2]. Another important use for seaplanes is to combat fires. Seaplanes are adapted as water bombers to carry up to 12,000 liters of fluid, such as the Beriev Be-200, and combat forest fires [3]. United States, Canada, Greece, Portugal, and Russia are some countries that rely on water bombers for fire-fighting. In Europe, however, the seaplane market is not as well developed as in North America. Most seaplanes are owned privately, and some are used as water bombers. Seaplane

Airlines are scarce, and they have to compete against other types of transportations, such as ferries and trains.



Figure 1 Seaport in Vancouver Harbour, Canada

2 Seaplane Operations

To get an overview about the current application of seaplanes and amphibians an online survey has been created and made accessible to operators worldwide during the Fusetra (Future Seaplane Traffic) project [4,5].

The following topics have been identified as subject of interest:

- General Information about Seaplane Operators
- Operational Issues
- Pilots, Regulations and Certification
- Infrastructure and Aircraft
- General issues and comments on the future development of the seaplane transport system.

3 General Information about Seaplane Operators

3.1 Participants' origin

The distribution of survey participants, given in Figure 2, shows that around one half them was European, more than one third North American and the rest from India and Australia. The European participants are equally distributed over the continent with slightly more participation in France and the United Kingdom, Figure 3. Surprisingly operators from countries with long coast lines, for example from the Iberian Peninsula, the Spanish Isles or Scandinavia (except Norway) did not participate. There are no participants from Russia, which is home to a very active seaplane industry and has a lot of inland waterways.

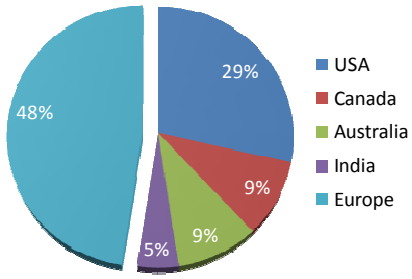


Figure 2: Origin of survey participants

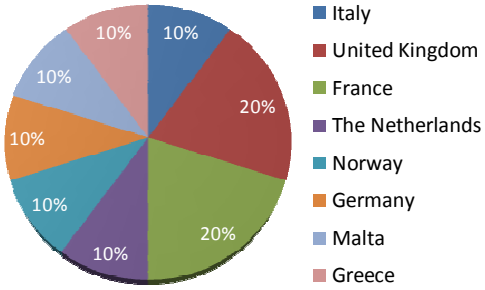


Figure 3: Origin of European survey participants

3.2 Participants' field of service

The participants' companies or clubs were found from 1930 to now. Almost two thirds of them were found in the current decade. This result strengthens the assumption that a long-lasting seaplane business is very difficult to establish and maintained, but without proven information of former operators it only stays an assumption.

Concerning the services offered by the participants, a general statement has to be made, that most of them offer more than one. The activities are divided into those given in the following list:

- Commercial airline, Commercial airline aspirant (or former commercial airline)
- Flight training, flying club
- Manufacturer
- Charter, Tourist
- Consultant

Popular combinations of services are flying schools that offer charter flights, generally the combination of charter and scenic flights and former commercial airline or aspirants that offer their services as consultants. Commercial airline aspirants are either working on obtaining their Air Operator Certificate (AOC) or already abandoned this plan.

4 Seaplane operations

4.1 Aircraft in operation

Figure 4 shows the types and amount of aircraft in operation 2010. The large portion of the deHavilland models is due to the fleet of the world's largest seaplane operator that almost only consists of DHC-2 and DHC-3. Its fleet marks almost half of all the aircraft considered for this study. The Cessna models 172, 206 and 208 are also widely used. The 172, that is also very popular with a conventional landing gear is the widest spread one of them.

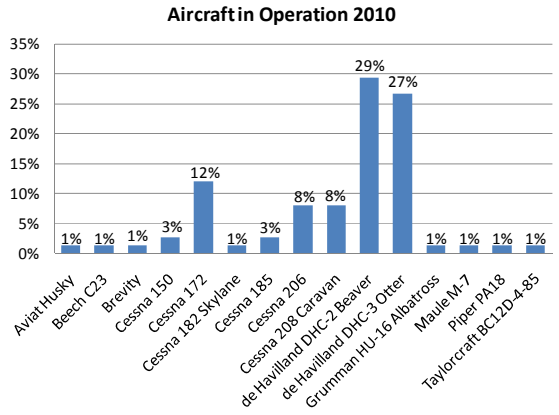


Figure 4: Aircraft in Operation 2010

When regarding the passengers that the above aircraft types are certified for, the distribution is given in Figure 5. The portion of aircraft with more than 15 passengers results from one Grumman HU-16 operated by a US based company. While models with one to three passengers are obviously mainly used for flight training, scenic and charter flights, the aircraft used for passenger transportation offer seven to fifteen seats.

The results on the undercarriage used are again mainly influenced by the fleet of the world's biggest seaplane airline that uses only straight floats. On the rest of the aircraft, more amphibian systems are mounted.

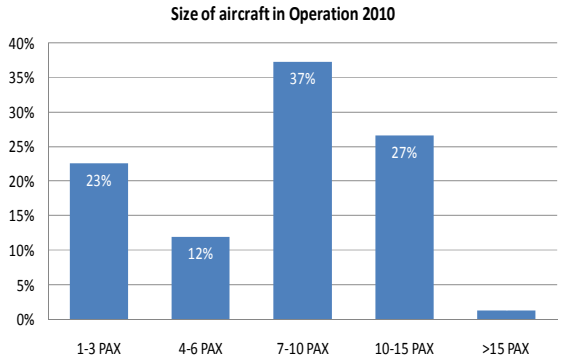


Figure 5: Size of aircraft in operation 2010

5 Certification

5.1 Pilots

The general situation, when summarizing all results about the comments on the availability of pilots, is not alarming. Almost three quarter of the participants do not characterize the situation as critical. Dividing up the continents shows that in North America the availability of pilots is unproblematic for over 85%, while for two-thirds of the European participants it is critical and challenging for the remaining one-third. In Asia and Australia the situation is generally characterized as challenging.

It was further asked where the pilots employed with an operator received their original flying license. Without exception it was issued in the country the operator is based in.

5.2 Operators

All of the participants considered in this section own the aircraft operator certificate (AOC), except for the US based flying schools and the commercial airline aspirants. In all cases it was issued by the national aviation authorities

(NAA) of the country the operator is based in. When looking at the participants' description of the certification process, it is clearly to be seen that only in a minor number of cases it was considered unproblematic. All these cases are North American companies. The Canadian Department of Transport is explicitly mentioned for an uncomplicated working relationship. In Europe the process is mostly seen as critical. In one case, a participant describes his impression that his NAA seemed to be complicating the process willingly. Several statements from participants worldwide say that there is a lack of understanding and sensibility for seaplane operations in the NAA.

When asked if they were assisted with the licensing process by their NAA, North American participants generally answered that they were not, but the process was feasible and known. A Canadian operator was assisted when looking for a new aerodrome. European participants complain about the unclear regulations and a missing point of contact within EASA. Expanding the question to the expectations they have for a central certification process governed by EASA or a central institution, and which points should be included, various points were mentioned. A specific European concern is to modify EU-OPS, so that for international business, it is not necessary to study the varying national laws. Furthermore a seaplane licence rating and standardisation in issuing landing sites was prompted. One European operator wished for a distinction between commercial and private operations with respect to the level of experience.

More than half of the operators are affected by special regional regulations concerning the use of waterways. Besides the mentioned national parks, they face generally restricted areas, excessive diffusion of water plants or are restricted to coastal regions, as reported from Norway or the United Kingdom. When asked if the compliance with both, maritime and aviation regulations leads to a conflict, one-third thought that they do. It was stated that maritime regulations do not consider the lack of manoeuvrability and ability to come to a sudden stop when compared to a boat.

6 Infrastructure and aircraft

6.1 Configuration of seaside landing site

The following table gives an overview about installations currently used by seaplane operators. The use of moorings, pontoons and respective foot bridges is commonly part of a seaport. Amphibian operators need not rely on seaside infrastructure and can used land bound landing strips, of course. The use of own maintenance hangars, offices and fuel stations is also not related to business size (or aircraft operated). Additionally, the use of emergency equipment seems not to be regulated differently from country to country.

Table 1 Landing site installations

Installation item	Operators using installation [%]
Moorings / Pontoons	50%
Foot bridges	41%
Navigation lights	5%
Maintenance site/ Hangar	18%
Office	36%
Fuel Station	32%
Emergency / Fire Services	14%

Other items mentioned apart from the above list were safety boats, fuel barges and navigation marks.

6.2 Future aircraft requirements

Main requirements towards future aircraft were aim of the survey too. The requested payload lies between 250kg and 1500kg, while speed is ranging from 140-180 knots. Additional comments on features which should be considered in aircraft development in future:

- Capable of operating in open / rough water (good sea state capability)
- High wing
- Amphibian
- IFR capable (Instrument Flight Rules)
- Engine designed to cope with very short cycles 5/hr

- Hot salt water tolerant corrosion
- Good visibility for passengers
- Suitable for use in confined areas
- Suitable for working with boats
- Low operating costs

6.3 Main Problems in Seaplane Aviation

Table 2 shows the main problems indicated by seaplane operators. The table distinguishes between worldwide and European operators in order to identify problems specific to European operations. It can be seen that availability of pilots and suitable aircraft is a major problem in Europe. Other major problems are aviation authority regulations and their implementation. On a worldwide basis environmental issues of seaplane operations are posing difficulties to operators during the licensing process of seaplane bases.

Table 2 Main problems of seaplane operators

Main Problems of Operations	Worldwide	Europe
Availability of licensed pilots	18%	50%
Availability of suitable aircraft	27%	50%
Safety issues	14%	20%
Passenger reservation about seaplanes	0%	0%
Opposition of environmental authorities	41%	20%
Aviation authorities regulations	32%	40%
Naval authorities regulations	18%	20%

7 Strengths, Weaknesses, Opportunities, and Threats analysis of existing seaplane operations

The aim of this SWOT (**S**trengths, **W**eaknesses, **O**pportunities, and **T**hreats) analysis is to recognize the key internal and external factors that are important to seaplane operations. The swot analysis may be then split into two main categories as follow:

- Internal factors: strengths and weaknesses internal to this particular type of transportation.
- External factors: opportunities and threats presented by the external environment.

Strengths and weaknesses of seaplane operations are here analyzed under the light of the "European Aeronautics: a vision for 2020" document [6], where the concept of sustainability is introduced and made the kernel of the aviation future. EU vision 2020 in not a deadline, but a sensible reflection on what should lie ahead for Europe in the near future in order to win global leadership in aeronautics. In vision 2020 aeronautics must satisfy constantly rising demands for lower costs, better service quality, the highest safety and environmental standards and an air transport system that is seamlessly integrated with other transport network.

Skies have to be always safer and the most advance automated systems have to be integrated to eliminate accidents. Aircraft need to be cleaner and quieter and the environment sustainable with the contribution of the aeronautic sector. The definition of sustainability states that "sustainability is the concept to endure". It depends on the wellbeing of the natural world as whole and the responsible use of natural resources. One EU (European Union) main objective, in this regard, is to halve, by 2020, carbon dioxide (CO2) emission, perceived noise pollution, and reduced nitrogen oxide (Nox) emission by 80% from 2000 levels. In conclusion it can be said that if a generation ago the imperatives were: higher, further and faster, then, according to the vision 2020 guidelines, these have become: more affordable, safer, cleaner and quieter.

7.1 Strengths

One of the major deterrents facing the seaplane market today is the opposition by environmental authorities on the perceived impact of seaplane. The main argument is based on the noise impact of seaplane landing, taxiing and taking off, which is known to exceed the ambient noise level. Additionally, there is a belief that noise, landing and take-off all impact on wildlife. A current example of this is the on-going dispute between Loch Lomond Seaplanes

and Trossachs National Park. Moreover, as mentioned before, also worldwide the greatest obstacle facing seaplanes is considered to be the opposition of environmental authorities. In Europe this was also agreed by 20% of operators [5].

- air/water pollution:

Only few studies have been completed to assess the seaplane environmental impact anywhere in the world and in many cases these are independent studies carried out by private seaplane operators [7,8]. The most inclusive and unbiased is probably an investigation conducted by US Army corps of Engineers (USACE) [9]. The outcomes were:

- Air quality: no impact
- Water quality: no impact
- Soil quality: no impact
- Wildlife: no impact
- Fisheries: no impact
- Hydrology: no impact

It is true that carbon emission generated from seaplane exceed the emission produced by boats. However, consideration should be given to the fact that the number of boat movements within any given area greatly outweighs seaplane movements in this area. Additionally, it should be considered that the next propulsion generation (which is already tested) will have much lower noise and carbon emission levels. Attention should also be drawn to the fact that seaplanes do not discharge sewage or oily bilge water and are not treated with toxic anti-fouling paints unlike boats. Seaplane exhaust are emitted into the air, much above the water giving low water impact, and currently used seaplane fuel does not contain the flammable and volatile compound MBTE (Methyl Tertiary-Butyl Ether), which is found in boats. Moreover, seaplane propellers are located away from the water, giving no disturbance on sediments or marine life, and they are near negligible polluters in regard of foul water and waste from chemical toilettes. Evidently, a further study validated that floatplanes generate no more than a three inch wake without any shoreline erosion effects [10].

Seaplanes have relatively low impact on noise pollution too. The majority of noise is generated during take off when high engine power is required to make the seaplane airborne. The following table lists typical noise levels for various operations at typical distances from the sound source and, once again, highlights the minimal impact seaplanes produce [11].

Table 3 Noise levels for various operations

Noise	dBA	Example
Military jet	120+	
Jet ski	110	e.g. watersports on lake
Chainsaw	100-104	e.g. tree felling/ forestry/ logging
Grass Cutting	88-100	Golf courses
Tractors	95	e.g. general operations
All terrain vehicles	85	
Speedboat	65-95	e.g. watersports on lake
Seaplane	75	on take-off only @ 300m (20 sec)
Inside car – 30 mph	68-73	
Normal conversation	65	

Attention should be also paid to the fact that the figure quoted is representative of the seaplane taking off, a short period of daytime-only occurrence which, compared to taxiing and landing, requires the highest throttle power.

In conclusion it may be said that seaplane do not have negative effect on hydrodynamics, hydrology, water quality, air quality, wildlife fisheries and birds or noise pollution when compared to existing background activities on lakes and seaports.

7.1.1 Society's needs

Air travel does not develop in a vacuum: its size, shape and success will be determined by society as a whole. Nowadays there are specific aspects of air transport that can be better or only satisfied by seaplane/amphibian operations. The most noticeable strengths in this regard are:

- Very versatile type of transportation.
- Point to point connections.
- Connections to very difficult to reach places.

- Safe and efficient surveillance in otherwise inaccessible destinations.
- Monitoring of wildlife and management of national parks.
- Very good safety records with few incidents during takeoff, landing operations or related to collisions with boats.
- Sightseeing tours/tourism.
- Ability to conduct rescue operations over large bodies of water, water bombers.
- Avoid the ever congested airfield, holding patterns and control sequences.
- No need for runway infrastructures, "unprepared" landing strip, smaller landing fees than landplanes.
- Access from 40% (flying boats) to 70% (amphibian plane) more of the earth's surface area than a conventional land plane.

7.2 Weaknesses

Seaplanes today are "endangered species" and although they posses undoubted potential, the lack of ability to unlock this potential is due to numerous problems. These are of a various nature and involve different aspects of seaplane/amphibian's environment. Certainly, the design aspect is a major impediment on seaplane advancement and is linked to many other areas. In fact, as with the introduction of new efficient commercial aircraft designs, the use of the seaplane declined, no new advanced designs have been made, and most extant seaplanes existing these days are approaching the end of their operating life. This situation has resulted in a scarcity of modern and cost-efficient seaplanes. The lack of innovative designs and use of today's technology then force seaplanes to VFR (Visual Flight Rules) and make them not suitable in adverse weather conditions or rough waters. In addition, some environmental issues could, in the near future, change what is currently a strength factor into a weakness. As stated before, vision 2020 aims to reduce polluting emissions by 50% for CO2 (Carbon dioxide) and by 80% regarding Nox (Nitrogen oxide). Alternative fuels and new generation engines, together with better aerodynamic performances, must be considered in order to keep these values as low as possible and match the suggested targets by the year 2020.

Finally, but equally important, the limited amount of seaplane bases and missing standard infrastructure equipment is surely a weak point that limits the seaplane market. It means that refueling and regular maintenance are factors which need serious consideration.

7.3 Opportunities

There is huge room for improvements in seaplane operations and many opportunities that can be exploited in such market. While demand is difficult to forecast without a detailed market research and an overview of current trends, something that is not available to fledgling industries, it can be presumed that demand should arise if the industry can offer a different service from large commercial airlines, either in terms of savings, convenience or novelty. Following is a list of the main features that may be considered as reliable new opportunities for seaplane:

- Easy usability among places with lots of islands and area/s with (many) resource/s of water.
- Faster service compared to ferries when connecting mainland-islands or island-island (e.g. Greece, UK, Ireland, etc) and the possibility to fly directly from major inland cities catering also specific groups of commuters in their daily journeys [12].
- Unconventional experience from transport (especially for tourists).
- Transport with quick dispatching.
- To shorten travel times avoiding the use of a combination of other means of transportation (e.g. Malta-south coast of Sicily) or considerable time savings that can be made where travel by any land based means is significantly time consuming.
- Avionics systems (lighten the burdens on the pilot, help making correct decisions and

reduce human error, night flight). In fact, seaplanes are limited to daytime VFR. Then the way to eliminate this disadvantage is by adding advance cockpit technology, or the used of advance gear such as GPS (Global Positioning System), radar, laser altimeters, gyros, advance sensors, among other gear.

- Larger seaplanes with better range, more seats and less affected by weather/water conditions.
- Efficient, safe, comfortable infrastructures [13] (seaports, docking facilities, accessibility…).
- Air freight services: cargos travel by air because it is more competitive.
- Modifications of existing planes with innovative new design. Based on the market research and the technological review, the creation of a new seaplane design will require time, manufacturing costs, regulation and certification, and social acceptance.
- Investments in new technology, materials and new seaplanes/amphibians advance design. When new advance design is involved, it should be consulted with operators, due to future equipment plans, and maritime authority regulations should be considered in advance of the design process. However, it may be expected that new solutions that lower drag when airborne, maintenance times and costs, and enhance competitiveness in cost/seat/miles ratio will be always looked forward by operators.
- Add value to the air transport market by opening up more locations to air travel and in doing so make it more convenient, while reducing the congestion on airfields and offering significant time savings to passengers.

7.4 Threats

For seaplanes to really take off there are a number of barriers that must first be overcome. This paragraph highlights the major threats that seaplane operation is facing today and the fundamental issues that need to be addressed [14]:

- Possibly difficult accessibility of airport (to replace automobile and railway means of transport is very hard in this case because of difficult approach of airports).
- Public perception of light aircraft safety may impact on the acceptability of seaplane transportation. However, it should be noted that in the UK there has not been a single reported accident according to their air accidents investigation branch (AAIB) [15], though this is in part due to the fact that there have been historically very few seaplane operated in the UK. [16].
- Acceptance from population and environmental activists.
- Fly time limitations. Alleviation on this regulation is needed so as to better meet the requirements of seaplane operations thus making them more financially sustainable without any subsequent of flight safety standards.
- Lack of a minimum level of training and acceptability of Dock Operating Crew so as to be multifunctional with regard to, assisting in the arrival and departure of aircraft on pontoons or piers, passenger handling, as well as manning the requirements of Rescue and Fire Fighting activities.
- Certification process for new seaplanes.
- General regulations: government regulation and control includes both aviation authority regulations and naval authority regulations. Nowadays there is not a set of unified regulations throughout Europe and these can also be sometimes in conflict.
- Corrosion resistance.
- Seaplanes are still too much depended on the weather conditions.

8 A new aircraft Conceptual Design and Ideas

In order to conduct an efficient seaplane design for the near future aiming to reduce costs and time on design, manufacturing and

regulations, existing certified aircrafts presented will be configured into a seaplane configuration by adding a floating device. These aircrafts have common similarities that are essential for the aircraft to be converted into a seaplane. They must have high wing configuration and have the capability of Short Takeoff and Landing (STOL). Some advanced ideas analyzed were the use of retractable floats, catamaran hull, navigation aids and hydrofoils, among others. Many proposed ideas were analyzed, but due to the complexity and high costs of some of these, it was decided to use retractable floats to reduce drag and the use of catamaran technology that will increase hydrodynamic performance of the seaplane.

The use of retractable floats gave a decrease in drag of around 5% compared to the floats in an extended position shown in Table 4.

Table 4 Drag Calculations for original design and seaplane configuration

Flat Plate Drag Area Breakdown [m^2]	*Aircraft*	*Seaplane [Extended]*	*Seaplane [Retracted]*
Fuselage	0.205	0.205	0.205
Wing	0.319	0.319	0.319
Horizontal Tail	0.078	0.078	0.078
Vertical Tail	0.055	0.055	0.055
Subtotal	1.279	1.279	1.279
Boat Hull	0	0.107	0.078
Floats	0	0.071	0.049
Strutting	0	0.037	0.012
Total	1.279	1.494	1.419
C_d	0.037	0.043	0.041
C_d Increment	0	0.0062	0.0041
Drag [N]	5724.63	6682.83	6346.82
Drag Increase		14.34%	9.80%

The seaplane in the retracted position improve the flight performance in which less thrust is generated, a greater turn radius for steady level turn was observed, higher absolute and service ceiling values were obtained, and better speed performance is given in each of the flight segments (takeoff, climb, descent, and landing). In general, retractable floats gave a great improvement in thrust and level turn, especially at high speeds. A final CAD model representation was elaborate to calculate the dimensions, observe the mechanism of the retractable floats, and show the location of the boat hull. In the near future, with the aid of this CAD model, it will be conducted CFD, structural and hydrodynamic analysis of the seaplane design. An example of how this seaplane will look adapting the catamaran technology and the use of retractable floats is shown in Figure 6.



Figure 6 CAD Model of Seaplane with retractable floats

9 Conclusion

Throughout the paper, the reader has been presented with a wide variety of different aspects of the existing seaplane operation. This was achieved through the investigation of current regulations, novel seaplane design, strengths/weaknesses and opportunities/threats analysis for a complete overview of all the aspects that such a market is facing today, will have to deal with or take advantage of in the near future to become a feasible alternative to other means of transportations. The European market has slumped and now faces a real challenge of over coming to obtuse and sometimes opaque aviation regulatory requirements that are completely insensitive to particular seaplane operations. Certainly, this is the main challenge to be taken in account when considering a bright future for seaplane operation in Europe. European regulation is not in its current state universally applicable and is, therefore, a stumbling block against further

seaplane development. Of particular note are the regulations surrounding seaplane base/landing sites which can vary wildly on a local, national and international level. A second issue is its acceptability. There are currently many reasons that affect the acceptability of seaplane operations both from a commercial standpoint but also from the stand point of general public. It may be suggested that the two most important ones are concerns about safety and environmental impact. It is, however, shown in the paper how studies on sites demonstrate the low impact that they have on the environment. Furthermore, seaplanes exhibit very good safety records, with very few accidents in general. Other main points that look to be more urgent than others in today's EU market, are the lack of qualified pilots, investment in new technologies, the need of larger seaplanes with more seats, better range and less dependent on weather/water conditions and the availability of safe, comfortable seaports or docking facilities.

However, seaplanes are uniquely placed as point to point connections and should certainly be marketed as such for new routes. They are very versatile and can cover very different type of missions that are peculiar of only this mode of transportation. They have access to more of the planet surface than landplanes, using unprepared landing strips, that makes them suitable for direct link between city centres (all major cities in Europe are situated near large enough bodies of water) and sightseeing, safe and efficient surveillance/monitoring of particular areas such as wildlife/national parks, ability to conduct rescue operations over large water bodies and fast and efficient way in fire fighting when used as water bombers. In conclusion it may be said that seaplane operations are, without doubt, a unique case in the whole transportation system, there is big room for improvements and a lot of new opportunities for existing seaplane operators and new emerging players. Only investing in the potentiality of this market can lead the way to a rapid and continuous expansion in the next decade. At the same time this will also make apparent to the competent organisms the necessity of new, common European standards, shaped on, and in line with the new market requirements and seaplane capabilities.

References

[1] Huda S., "Amphibian Aircraft Concept Design Study," Dept of Aerospace Engineering, University of Glasgow, UK, 2009.

[2] MacGregor Garcia G. K., "Future Seaplane Traffic" Dept of Aerospace Engineering, University of Glasgow, UK, 2009.

[3] "Beriev BE-200" beriev.com, 2011, http://www.beriev.com/eng/Be-200_e/Be-200_e.html [Cited Jan 26, 2011]

[4] http://www.fusetra.eu

[5] http://www.fusetra.eu/documents/malta/Survey_ Results_Malta.pdf

[6] Report of the group of personalities, "European aeronautics: a vision for 2020. Meeting society' needs abd winning global leadership", Luxembourg: Office for Official Publications of the European Communities, 2001.

[7] http://www.seaplanes.org/advocacy/environment.pdf

[8] http://www.seaplanes.org/advocacy/booklet.pdf

[9] http://www.seaplanes.org.au/PDF/Seaplanes-The_Facts.pdf

[10] http://www.seaplanes.org.au/PDF/Seaplanes_and_ the_Environment.pdf

[11] http://www.harbourair.ie/environmental-impact.pdf

[12] Office for National Statistics, "An investigation into the location and commuting patterns of prt-time and full-time workers in the United Kingdom, using information from the 2001 Census1 Alistair Dent & Stephen Bond", Executive Summary, UK, 2008.

[13] Department of Transport, "Public experience of and attitudes towards air travel", UK, 29th July 2010.

[14] Lightening B., "Road Map for Fullfilling Market and Operator Needs in Seaplane Operation Within Europe", Fusetra, 2011.

[15] http://www.aaib.gov.uk/home/index.cfm

[16] Transportation Safety Board of Canada, "A SAFETY STUDY OF SURVIVABILITY IN SEAPLANE ACCIDENTS", Report Number SA9401, Canada, 1994.

AIRPORTS OF THE FUTURE

Marc Noyelle
Académie de l'Air et de l'Espace/Air and Space Academy, France

Keywords: *capacity, environment, security, service, baggage, information*

Abstract

Airports of the future will have to meet 4 major challenges: capacity, environment, security and quality of service for passengers.

The number of PKT (passenger x km) will be multiplied by 3.3 between 2010 and 2050, and the number of flights will double with main hubs growing the fastest.

Most of the world's 1600 existing commercial airports will need considerable investments (infrastructures, equipment and ground access) and approximately 200 new platforms will be built. With a few exceptions, they will meet the demand.

The airports of 2050 will resemble our current airports, with long runways, given the fact that vertical and/or short take off and landing aircraft will remain scarce since they require too much energy per passenger and are therefore too expensive.

The services provided (including baggage and personal information) should be considerably better than today.

Introduction

The Air and Space Academy (AAE) set up a Foresight Commission made up of Academy members and other aviation experts to look into the challenges facing civil air transport by 2050 and to put forward some recommendations when this important task is completed in early 2012.

Marc Noyelle conducted the studies concerning operating structures (airlines and airports) and future air transport costs. The present paper is mainly the result of these reflections concerning airports as of summer 2011.

For at least the next 40 years, Airports of the future will resemble our current airports, with long runways: vertical and/or short take off and landing (V/STOL) aircraft will remain scarce since they require too much energy per passenger and are therefore too expensive.

Airports will have to face 4 major challenges: capacity, environment, security and quality of service for passengers. In this paper, we will see:

1- Future demand (2050) and airport capacities
2- Environment and access to airport-cities
3- Security

4- Quality of service for the passengers (and their luggage)
5- Conclusions: Airports in 2050

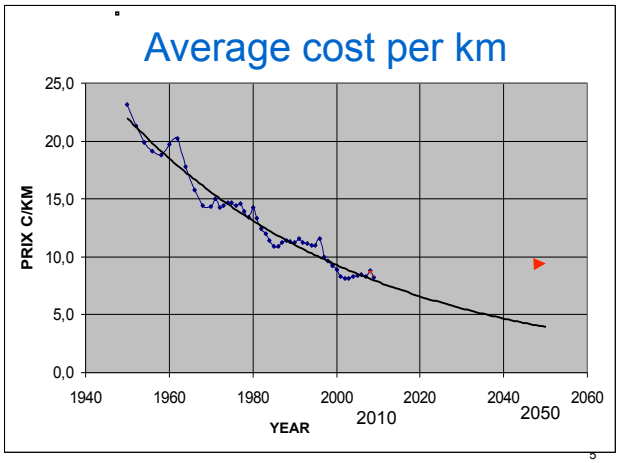
1 DEMAND AND AIRPORT CAPACITY

1.1 Evaluation of the traffic and the average fare in 2050

A clear relationship was highlighted between traffic and GDP: in a given country, it was shown that expenditure on air transport is a constant part, 0.85%, of the GDP.

The cost of the average ticket will increase by only 10 to 20%, while the cost of air fuel will triple: the share of spending on air fuel will grow from 25% in 2010 to almost 50% in 2050, while other components of the cost of a ticket (aircraft and maintenance, staff, ATM, ground handling and catering, airport, commercial and administrative expenses) will be reduced by 20%.

Altogether, we think that the average cost per kilometre of a plane ticket will remain under 0.10 US $.



Source : Stanford University + AAE

According to our assessment of growth in GDP in the 6 main regions of the world, the number of PKT (passenger x km) will be multiplied by 3.3 between 2010 and 2050, and the number of flights will double (x1.4 in Europe/USA, and x3.1 elsewhere in the world) with main hubs growing the fastest.

1.2 Airport Capacity

The capacity of an airport depends on *4 factors: 2 "air-side"* (capacities of the runways and average aircraft passenger load) *and 2 "land-side"* (capacity of terminals and capacity of the environment). The last factor contains 2 components (access and acceptable noise for the surrounding area).

The number of movements (landing or take off) depends on the number of runways (of proper dimensions) and on the possibility of exploiting them with adequate taxiways and sufficient parking place for the aircraft, and of course also on the ATM (Air Traffic Management) which will be more efficient than today.

It depends also on local geography (and climate), aircraft opening hours, type of demand (hub or not, rush hours, etc.), defrosting equipment, etc.

Paris-CDG (Roissy) for instance, with its 4 parallel, well oriented runways today has a programming capacity (for establishing slots) fixed by air traffic control at almost 120 movements per hour, which in clear weather makes it possible to fit in more actual movements (sometimes over 160 flights in a single hour).

The highest "runway" capacity in the world is to be found in Dallas FW airport with 260 movements per hour (7 runways)!

With 500,000 movements per year, CDG is far from being technically saturated, although the pressure of demand at certain peak times is very high. London/Heathrow, for example, counts 450,000 movements per year on only 2 runways, with less favourable weather conditions and significantly larger average aircraft size than at Paris/CDG.

The larger the aircraft capacity, the greater the air-side capacity of the airport, measured in numbers of passengers.

The *average passenger load* globally is of 90 passengers per aircraft in 2010. This figure is heavily influenced by the very low passenger load in the USA due to the high number of low capacity domestic flights.

The average aircraft passenger load varies from one region to another. It is 80 in Chicago/O'Hare, an airport in which the

runways are not saturated, 118 passengers per aircraft at CDG but over 150 at Heathrow and 220 on average in Tokyo/Haneda (world record).

The world's largest airport, Atlanta, sees over a million movements per year on 5 runways, with an average carrying capacity of 90 pax/plane. This hub has not yet reached 100 million passengers per year, but could easily pass this threshold without building another runway.

The *number of passengers an airport can deal with depends on the number of terminals and their surface area*. They must be arranged in an optimal way with enough doors, adequate gangways and constant efforts to iron out bottlenecks which can shift according to security and border control regulations, international or domestic flight destinations, the size of the aircraft, etc.

This presupposes necessary investment in bigger aircraft and a constant search to meet passengers' needs (with efficient luggage sorting systems and enough properly sized delivery belts, police booths and security posts, etc.) and sufficient personnel and public services (border and customs control) available at a given time.

Optimal organisation of all ground-based (passenger) and air-based (aircraft maintenance, cleaning, refuelling, etc.) services clearly also contributes to airport capacity.

Lastly, capacity depends on the political will to provide *proper ground access to the airport and* to set in place a noise management plan so as to stay within an *environmentally acceptable* limit.

This "environmental acceptance" can be very constraining and catastrophic if it translates into inflexible regulations. Thus for instance, the fixing of a ceiling of 250,000 movements per year in 1994, with a ban on night flights, immediately put a stop to Orly's development and the corresponding jobs. It would have been *wiser to set a limit on the measured impact* (a noise ceiling for instance), thus allowing a certain development, given the progress in noise levels of aircraft.

1.3 How many commercial airports in 2050?

The number of flights (regular + charter + freight) should increase on average by nearly 40% in Europe/USA by 2050 and elsewhere in the world by 3.1. According to a BCG study "Airports, Dawn of a new era", hubs will grow noticeably more than other airports. *The average number of passengers per plane would increase, from 90 in 2010 to 125 in 2050*, for regular flights (which make up the large majority).

Roughly 1,600 airports are currently open to commercial traffic in the world, 600 of which are in North America and 400 in Europe.

Commercial airports in the WORLD	Total (Incl. France)	>1Mpax / year	>10Mpax / year	>33Mpax / year
Europe	400 85	175 16	47 3	6 1 London,CDG, FRA,MAD
North America	600	120	48	14 ATL,Chicago,LA NY
Latin America	200	65	8	0
Africa + Middle East	200	45	7	1 Dubaï
Asia + Pacific	200	90	40	9 PEK,HND,HKG, JAK 9
Total in 2010	1600	495	150	30

In 2008, up to 500 airports dealt with more than 1 million passengers, 175 of which were in Europe, 120 in North America, 90 in Asia/Pacific, 65 in Latin America and 45 in Africa/Middle East. Nearly a third of world traffic is concentrated around the 30 largest airports.

This list changes regularly: Beijing thus went from 8th largest in 2008 to 2nd in 2010 with nearly 20 million extra passengers in two years. Indeed a second airport is planned to the South of Beijing around 2017, with a capacity of 130 Mpax.

China had 4 airports in the list of the 30 largest airports in 2010 as against only 2 in 2008. Sydney and Jakarta also featured for the first time in 2009 whilst two American and two European airports disappeared.

Dubai is now in 15th place with 42Mpax; 10 years ago it had 4 times fewer passengers. The Dubai hub has a particularly high number of long-haul/long-haul connections, due to its favourable geographic situation. It is limited however in terms of its surface area (1400 hectares, only 2 runways, hemmed in by the city), hence the 2nd airport currently being built in Djebel Ali, 40 km to the South.

Paris CDG is in 6th place with close to 60Mpax. It is a classic hub with a large available surface area (3300ha), greater than Heathrow and Fraport combined, which will not spare it from experiencing congestion problems by 2050.

By analysing each of the 150 airports with over 10Mpax in the world, it emerges that most could deal with 2, 3 or even 4 times more flights by 2050, or else rely on nearby airports, existing or new.

However no satisfactory solution exists for some major capitals: London, whose 5 airports (Heathrow, Gatwick, Stansted, Luton and City) are all too small. Los Angeles whose 4 airports, like New York (unless the former Floyd Bennett Field base next to JFK were given over to commercial traffic), are too small. Tokyo with a congested Haneda, although Narita could conceivably be extended. Istanbul and Mumbai have two airports close to saturation point and are looking for a third site (difficult in the case of Mumbai), as is Shanghai which could be saturated in 2025 despite the potential of Pudong. Hong-Kong is planning to build a third runway, but what then? Jakarta is looking for a site for a second airport around 2020, as is Sydney, for which the chosen site is no longer available…

Few airports in the world will be able to deal with over 1 million movements a year, or 110 to 160 million passengers in 2050.

In practice, the number of airports will rise in certain parts of the world such as China which is currently building dozens of new airports to complement an exceptional construction programme underway of over 20 high-speed rail routes!

But in Europe and America, the number of airports should not rise appreciably because of the difficulty of finding appropriate sites and residents' frequent opposition (NIMBY "Not In My Back Yard" syndrome).

In Europe and the USA, the number of airports is, with some notable exceptions, sufficient to meet the demand expected by 2050, but this requires major investments into and around many existing airports (extensions and access), as well as improved ground transport to enable network operations around certain saturated airports.

Elsewhere in the world, we estimate that around 200 new airport platforms will be built as well as adapting existing airports in order to meet demand in 2050.

2 ENVIRONMENT AND ACCESS TO AIRPORT CITIES

2.1 Ground access

An airport is often judged in terms of its ground access.

The problem is that this aspect does not come under the responsibility of the airport manager, but of (sometimes numerous) public authorities. It is however essential to provide for excellent access:

- Rail access to the airport city and surrounding region: rapid metro system like Heathrow Express, and high-speed rail connection. There could also be high-speed freight rail connections between certain major airports and ports, at least in Europe, Japan and China.
- Other public transport: good bus and taxi (both individual and collective) network.
- Efficient car hire system (50% of the car rental market in France comes from airports).
- Suitable road access and practical car parking facilities, with good links to terminals (automated people-mover if necessary).
- Helicopter services in certain specific cases.

2.2 Controlling pollution around the airport.

Aircraft have already made huge progress in terms of noise pollution and will continue to do so.

The best compromise is and will continue to be sought so as to meet their insertion into the environment on a long-term basis:

- *Noise*: promote less noisy air procedures (eg: *CDA, Continuous Descent Approach*, tested in Los Angeles and Orly in 2007, and from 2011 in Atlanta, San Diego, etc.) and limit noisy night movements. *Research and experimentation must go into elaborating such procedures*.
- *Urban planning*: check that reasonable extension of the airport is provided for by preserving non-constructible areas and also offering sound-proofing grants to existing residents or even purchasing their properties in certain cases.
- *CO2, NOx and other gas emissions*:
  - All airport buildings and vehicles must meet the stiffest requirements and be operated in a cost-effective sustainable way;
  - Aircraft management on the ground must be improved in order to reduce the lapse of time between authorisation to begin taxiing for take-off and effective take-off. More accurate forecasting of take-off times will help optimise use of en-route air space and destination airport capacities (CDM).
  - Energy provision of the aircraft (electricity, air conditioning) should come from the ground as soon as possible after the aircraft has landed.
- Relations between residents and other airport stakeholders: encourage a relationship of mutual, enduring trust with the surrounding.

2.3 Airport cities

Major airports already are, and future airports will even more so be, a *major entry gate to the world.*

Like all major transport intersections, airports generate commercial *activities that go way beyond the simple transfer of passengers and goods.*

Airports do not only welcome the innumerable airlines, passengers and accompanying persons, taxis and other means of transport, but also service businesses for passengers and aircraft, freight and postal services, and countless offices, hotels, import-export firms, congress centres, businesses, etc. attracted by this “frontier” location.

These groupings of businesses are turning into *highly dynamic airport cities. An airport is not only a place of transit, but a destination in its own right which gives it an important local profile*. Thus the site of Roissy-Charles de Gaulle, with a surface area of some 8,000 acres belonging to Aéroports de Paris, already generates some 100,000 on-site jobs in 2010 (in over 1,000 companies), and many others in the surrounding area, where 40 years ago it was almost deserted. The number of jobs increases by several thousand per year, except during crises.

Future airports will no longer be considered simply as a nuisance for residents with the only benefit going to “strangers”, but as generating wealth and employment locally. This will hugely facilitate their acceptance.

3 SECURITY

Security measures, aimed at protecting against hostile actions, evolve very rapidly and will undoubtedly make great progress by 2050.

Demands on the part of security authorities have risen steeply since the terrorist attacks of 11 September 2001. Nowadays it is the most constraining function and (along with baggage) the butt of the greatest number of complaints.

National and European regulations are superimposed. Certain countries can enforce complementary formalities as a reaction to events, which can further complicate control procedures. Security not only has a cost (in France the security tax has been raised to 11.5€ per departing passenger), but is also and above all an obstacle to improved aircraft productivity,

as well as representing a loss of time and even dignity for passengers.

Systems set up in airports for security - passenger flow separation, efficient security checkpoints, etc. - *have necessitated in the terminals 15% extra space per passenger in the 10 years from 2001 to 2011, a steep rise which will undoubtedly continue.*

Security is a national activity. Its cost, however, is entirely shouldered by passengers in Europe, which is not the case for rail transport, nor is it the case in the USA.

Security *represents nearly 20% of airport turnover and more than a third of airport operational expenses.* This is in danger of increasing with increased demands on the part of security authorities

It is important both to *put more efforts into making malicious acts more difficult in airports and into elaborating controls that are both highly reliable and acceptable to passengers.*

New technologies are becoming available: large scanners, for instance, to monitor passengers with their baggage as they walk along a corridor equipped with sophisticated sensors, without (ultimately) the person being monitored having to stop; in May 2011, the European Parliament authorised these body scanners which give the person being controlled the possibility, in exchange, of refusing a manual body search.

Many experiments are underway globally giving passengers the choice between these rapid, efficient systems, and the slower, often badly tolerated frisking.

It is vital that industry continues research into improving security control systems where possible.

4 QUALITY OF SERVICE FOR PASSENGERS

4.1 Passenger information

Accurate information in due time is of *primary importance to passengers*, who are always under a certain amount of stress: it is never easy to find the right panel and/or an available information desk to obtain reliable, comprehensible (in one's own language) information.

This is *particularly crucial during periods of disruption*, which are frequent in such a complex, fragmented transport chain due to the number of players and the many meteorological, mechanical or social variables.

In order to provide reliable information (aside from the rapid sharing of information between the different personnel present in the airport which is provided by CDM, Collaborative Decision Making), the classic display panels and signs, regularly updated, will more and more often be combined with small computer terminals in the hands of personnel in contact with passengers, and passengers' mobile phones with GPS or other positioning means. Text messages for instance informing connecting passengers as they land of the path and distance to the boarding gate have already met with success.

4.2 Baggage

Baggage handling (for departures) includes check-in, security controls, sorting and dispatching of tens of thousands of cases per day in the major hubs. It is the responsibility of each airline, which usually makes use of facilities placed at its disposal by the airport.

Because of increased security demands, these facilities are more and more sophisticated. The automated baggage sorting system in the East part of CDG2, for instance, deals with 50,000 bags per day.

For the same security reasons, the idea of taking charge of baggage before the airport and as far as the final destination has been put aside, despite its attractiveness.

Baggage handling largely determines the efficiency of an airport, with an exorbitant overall cost (investment and operation). The thorniest problem is that of connections, given that roughly half of passengers and ¾ of hold baggage are in transit.

The question of baggage has almost always been underestimated. Many existing airports are cruelly lacking in the space needed in order to improve the situation whilst meeting new security standards.

The aim is to increase the system's flow and efficiency with a maximum baggage delivery time of 20 minutes for passengers arriving at their final destination, and 40 minutes transfer time from one hold to another, for connecting passengers, after the aircraft's arrival at its block. Dependability must also be increased by cutting the number of badly routed bags by a half or two thirds, in order to reach an error rate of under 1% (this figure can be over 3% on occasion). IATA estimates that routing errors cost its members more than 3.3 billion$ each year and endorses the BIP Baggage Improvement Programme.

Techniques employed successfully in one airport should ideally be transposed to all others (interoperability). *Luggage losses could be drastically reduced for instance by using* a single international system in all airports to identify each piece of luggage (as is the case for maritime transport containers); this would involve placing an *RFID (Radio Frequency Identification Device*) during manufacture in all suitcases and travel bags sold in the world (this does of course raise other problems).

Unfortunately we are still far from this kind of interoperability and there is plenty of room for improvement.

Another point: one day *mechanical aides will probably have to be installed to handle baggage* so as to make this job less physically demanding, as already tested at Nice airport and in Sweden. *This will also take up room in the airport buildings…*

One could also radically change the process by encouraging passengers by means of financial incentives to carry their own baggage as close to the aircraft as possible, maybe even into the aircraft. This cheap solution would nonetheless require extra room in the cabin and also in the airports if passengers required help carrying their suitcases. This radical simplification went on one century ago for the railways; it should be examined, although it seems that it would never be able to be completely the rule for air transport…

4.3 Other services

The development of e-services for check-in is inevitable: over half of all check-ins, ¾ in the case of domestic flights, are "self-service" nowadays; further services are optional, with the passenger choosing to pay for them or not.

As well as traditional services such as car parking (with several different options available according to length of stay, the possibility of booking a parking place, etc.), well-stocked shops, restaurants for all tastes, lounge areas, customised reception of persons or groups, etc., many other "premium" services are or will soon be available:

- "fast track": a special pass to enable the holder to skip security control queues, with the option of personal assistance,
- reception of and assistance to PRM persons of reduced mobility,
- relaxation area, even a hotel bedroom in a reserved zone in the case of a very long connection (an option that already exists in Dubai and shortly in CDG),
- cyber-centres, Wi-Fi connection and other Internet services,
- gaming and sports areas, libraries, children's areas, …
- valet parking, porter services, etc.

Ensuring excellent service for every passenger will remain a real challenge in practice. It will always require organisation and determination and careful attention at all times on the part of those working in the airport.

NB: we cannot talk here of airport management nor Collaborative Decision Making whose aim is to make operational decisions swifter and more efficient, nor safety aspects (runway incursion, etc.), nor border control and the contribution of biometrics in making identification of persons simpler and more reliable, etc. For such aspects the reader is referred to the Air and Space Academy's Dossier 33 on "Airports and their Challenges" published in late 2010.

5 CONCLUSIONS: AIRPORTS IN 2050

- The number of passengers will triple and that of flights will double.
- *Airports of the future will resemble our current airports with long runways* given the fact that V/STOL (vertical/short take off and landing) aircraft will remain scarce since they require too much energy per passenger.
- The size of the aircraft will remain in the "80 box: 80 meters long, 80 meters wide and 80 feet high".
- *Most of the 1600 commercial airports in the world will need to invest in order to meet demand and about 200 new platforms will be built essentially outside Europe and the US.*
- Huge difficulties are facing only a few major cities like London, New York, Los Angeles, Tokyo, Shanghai, Hong-Kong, Jakarta, Istanbul…
- *The main hubs will develop more rapidly than the point-to-point airports*.
- Airports will have to concentrate on environmental aspects (noise, urban planning) and build up relationships with the community based notably on trust by developing airport cities.
- Need to improve CDM and ATM as well as:
  - *Ground access*: intermodality
  - *Security*: efficiency combined with respect for passengers' dignity
  - *Baggage*: efficiency and interoperability
  - *Good and Personalised information* for each passenger
- *Build large, adaptable airports so as to offer the best services at acceptable prices*.

We are building today the airports of tomorrow.

Paris (France) September 2011,

Marc Noyelle

A few references

[1] Dossier AAE n°33 « Airports and their Challenges » Nov. 2010
[2] EADS « Friend Lean Airport of the future » June 2011
[3] IATA « Simplifying the Business » 2011
[4] European Commission « Flightpath 2050 » 2011

CEAS 2011

Experimental Setup in Aerodynamics

Experimental and Numerical Investigations of Vortex Generator Jets for Active Flow Control at High Reynolds numbers

M. Casper, S. Mahmood, P. Scholz and R. Radespiel
Technische Universität Braunschweig, Institut für Strömungsmechanik, Braunschweig, Germany

Keywords: *Active Flow Control, Vortex Generator Jets, High-Lift Airfoil, High Reynolds number, Scaling effects*

Abstract

An active flow control (AFC) experiment with a 2D two element high-lift airfoil equipped with vortex generator jets at Reynolds numbers of up to Re=12.2·10^6 was performed in the cryogenic test facility DNW-KKK in Cologne, Germany. The measurements were conducted to study Reynolds number scaling effects. The results show that AFC with vortex generator jets at high Reynolds numbers is possible but for the given setup the benefits decreases the higher the Reynolds number. This effect is investigated via numerical simulations by the density, temperature, and velocity ratio between the free stream and the jet.

1 Introduction

An airfoil in high-lift configuration could be a complex system consisting of a main-wing with a slat, one or more flaps and the corresponding kinematics to position these parts relative to the main-element. Although all these parts are retracted in cruise flight the performance of an aircraft is influenced. A slatless wing will result in a seamless leading edge region and offers the ability to use laminar airfoils with less drag. There is no need for the corresponding kinematics of the slat which will reduce the weight of an aircraft. Both advantages will reduce the fuel consumption in cruise flight and lowers the costs as well as the impact to the environment.

But the aerodynamic function of a slat must be replaced otherwise such an airfoil can stall at the leading edge. To suppress stall active flow control (AFC) is a promising technique. One method is the vortex generator jet (VGJ) [1] which enhances the momentum transport from the boundary layer edge close to the surface and suppresses stall [2]. Compared to the passive devices, the vane vortex generators (VG) [3] as well as the hidden vane vortex generators [4], the orifices of the VGJ's are only small disturbances in the surface of the airfoil but such a system needs compressed air. So, these devices must work efficient to generate strong vortices with a minimal mass flow. The VGJ's can be used statically as well as dynamically which is assumed to be more efficient [5, 6, 7]. So, there exist a large number of geometric and operational parameters to optimize. This includes the jet size, the jet angles, the distances between several jets arranged to an array as well as the mass flow, the frequency, and the duty cycle.

The actual study reports about VGJ's close to the leading edge on a two element 2D high-lift airfoil, the DLR-F15 airfoil model, consisting of a main-wing and a single flap with the focus on Reynolds. For this high-lift airfoil the sources of the scaling effects are [8]: the decrease of the boundary layer thickness with the Reynolds number, the disappearance of the laminar separation bubble with the Reynolds number, the movement of the transition with different Reynolds number, and the changes in the interaction between the wake of the main-wing and the flow over the flap. With respect to wind tunnel experiments it may be noticed that the increase of the Reynolds number has a favourable impact also on the side wall effects. The boundary layer thickness decreases with the Reynolds number. The wind tunnel experiments were performed in the cryogenic wind tunnel facility DNW-KKK in Cologne, Germany. The Reynolds number varied from Re=2.1·10^6 to Re=12.2·10^6 and the Mach number from M=0.15 to M=0.2. Numerical simulations with a single VGJ and a flat plate without pressure gradient will help to understand the measured effects.

2 Experimental setup

The experiments were conducted with the DLR-F15 model, which is part of several studies without and with AFC [9, 10, 11, 12, 13]. It consists of four elements, the leading edge box, the main center-element, the shroud, and the flap as seen in Fig. 1.

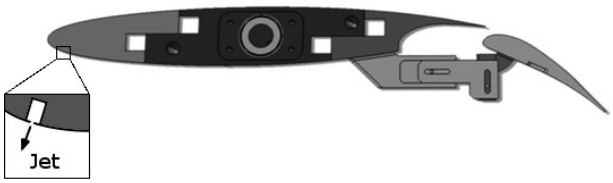


Fig. 1 The DLR-F15 airfoil model with the four segments (from left to right) leading edge box, main center-element, shroud, and flap; sketch is not to scale

The flap is connected to the shroud by four flap tracks, which allow variable positions of the flap. The main-wing (leading edge box, main center-element, shroud) is made of Weldural (AlCu6.5Mn0.3), which is a high strength aluminium with low internal stresses. The flap is made of glass-fibre reinforced plastic (GFRP) with almost identical thermal expansion characteristics. Due to the fact that the wind tunnel is cooled down to 100K identical thermal expansion characteristics are important, otherwise stresses will deform the flap which influences the measurement results. The wing span of the airfoil is 2.4m and with retraced flap it has a reference chord length of 600mm. The model has approximately 85 static pressure tabs in the middle section to measure the pressure distribution of the high-lift airfoil. Sidewall effects are observed by 17 static pressure tabs close to the sidewall. The influence of the sidewalls is reduced by VG's close to the test section on the suction side of the airfoil. A five-hole probe wake rake traverse allows to verify the extent of the two-dimensionality of the flow.

The AFC system, the vortex generator jets, is positioned at 1% chord length on the pressure side and allows different actuator jet configurations. A previous work [14] shows that in case of a leading edge stall a actuator position on the pressure side should be preferred. The AFC system consists of the round actuator orifices, settling chambers, solenoid valves and fed lines as seen in Fig. 2.

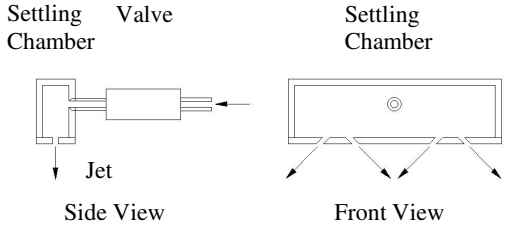


Fig. 2 Principle side and front view of the actuation system consisting of the solenoid valve, the settling chamber, and the actuator orifice with the jet

The airfoil model is separated in the middle and has one fed line per side. One fed line is connected to 18 solenoid valves. One valve feeds one settling chamber and four actuator orifices with compressed, non cooled nitrogen. To realize the cryogenic measurements the wind tunnel model must be prepared. The solenoid valves have an allowed temperature range from -5°C to +40°C. With respect to the cryogenic conditions the valves must be heated and controlled. Therefore, heater foils (Minco), pt1000 temperature sensors, and thermal insulations are used. The controllers, industrial

thermostats (Jumo eTron T), are outside the cooled area and are not sensitive to electromagnetic distortions on the measurement line. As well as the wind tunnel model the fed lines must be prepared, too. The two fed lines are installed differently. One fed line is in a warm area and the other one is in a cooled area, the model process chamber, under the test section. This will cool down the gas in this fed line and will result in a non uniform blowing of the VGJ´s. To prevent this, the fed line in the cooled area is heated and controlled, too. Additional pt1000 temperature sensors inside the wind tunnel model are used to check this. An identical flow rate is realized by a mass flow rate sensor (Testo 6442) and a pressure regulating valve (Festo MPPE) for each supply line. The flow rate is used to calculate the moment coefficient as defined by [12]

$$c_\mu = \frac{\dot{m} \cdot u_{jet}}{\frac{\rho_\infty}{2} \cdot U_\infty^2 \cdot S} \qquad (1)$$

The calculated values c_μ of the experiments are in Tab. 1.

Tab. 1 Momentum coefficients of AFC test rows

M [-]	Q [l/min]	c_μ [-]	f [Hz]	dc [%]
0.15	1460	0.0014	100	50
0.15	2400	0.0053	100	50
0.15	3200	0.0112	100	50
0.15	2400	0.0053	-	100
0.2	1200	0.0005	100	50
0.2	2400	0.003	100	50
0.2	2400	0.003	-	100

Previous experiments [12] in the DNW-KKK with this airfoil model were performed to test different flap settings and the actual AFC system for the ability to suppress stall. The most promising configuration with AFC, defined by a high $c_{L,max}$ and $\Delta c_{L,max}$, is a flap gab of 2,7%, a flap overlap of 0.5% and a flap angle 45°. The AFC system in the previous and this experiment is a round orifice with a diameter of d=9mm, which points in spanwise direction β_{jet}=90° and has angle of α_{jet}=30° between the jet and the surface. The actuators are arranged to produce counter rotating vortex pairs. The distance between two orifices with common flow down is 9d and 24d for the common flow up configuration.

The wind tunnel tests were performed in the cryogenic test facility DNW-KKK in cologne, Germany. It is a closed circuit, continuous, low-speed wind tunnel with an ambient pressurized test section of 2.4m x 2.4m. The wind tunnel is cooled by injecting liquid nitrogen and temperatures down to 100K are possible. This allows variations of the Reynolds and Mach number independently. The available Reynolds and Mach number envelope with the DLR-F15 airfoil model is presented in Figure 3 and the tests with AFC are marked with red circles. The Reynolds number varied from Re=2.1·10^6 to Re=12.2·10^6 and the Mach number from M=0.15 to M=0.2.

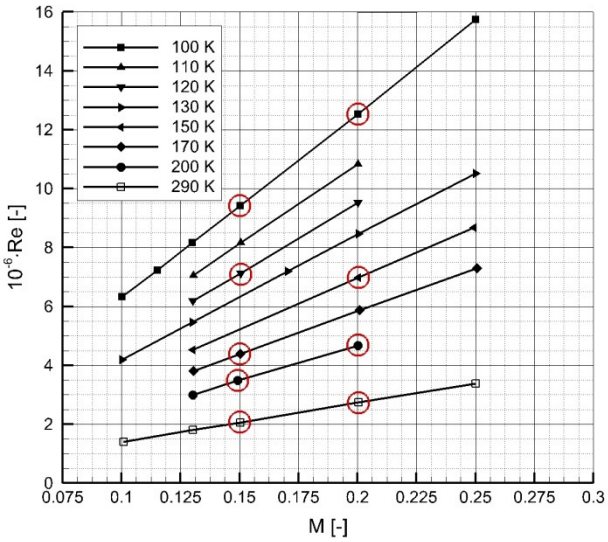


Fig. 3 Reynolds- and Mach number envelope of the DLR-F15 airfoil in the test section of the DNW-KKK

3 Numerical setup

In order to study the effect of VGJ in cryogenic conditions. Numerical studies have been carried out with AFC over the flat plate at lower Mach number using circular VGJ as used in [15]. Since the numerical AFC results show good agreement with the experiments [15,16] conducted in standard operating conditions. Therefore this test case was selected to study the influence of hot jet interacting with the cold free-stream in cryogenic conditions. The numerical flow solver used is the DLR TAU-code [17] developed by the German Aerospace Centre. The TAU-code is flexible in solving

different types of grids like hybrid meshes having prism layers and tetrahedral mixed with hexahedral cells. In order to accelerate convergence low-Mach number preconditioning with multigrid is employed. The inviscid terms in the governing equations were descritized using the second-order Central scheme with scalar dissipation.

In order to simulate the effect of VGJ, grid generation has been carried out using the Chimera approach [18, 19]. This makes it easier to handle the large scaling differences between the flat plate and the VGJ [15]. The total mesh size is $3.8 \cdot 10^6$ million grid points having fully structured meshes as shown in Fig. 4 generated with GRIDGEN V15.13 [23] package. Fig. 4 shows the Chimera grid setup with local grid including the VGJ. The selected grid resulted in grid independent solution as explained in [15]. For the present study, Menter's Shear Stress Transport turbulence model [20] is used as it performs well in the separated regions [15, 12]. Cryogenic nitrogen behaves as real-gas that is not thermally perfect. Such that the ratio of specific heats and viscosity are functions of pressure and temperature [21]. For this study it has been assumed that the cryogenic nitrogen gas behaves as an ideal-gas assumed to have negligible real-gas effects at very low Mach number. This means that the nitrogen gas was assumed to be thermally and calorically perfect. The test matrix studied is shown in Tab. 2. Where T_{jet}, ρ_{jet} and m_{jet} defines the jet temperature, density and mass flow rate. *M* and *Re* defines the dimensionless flow parameters as Mach and Reynolds number. The *DR* defines the density ratio between the jet and free-stream density. Where *PR*, *TR* and *VR* defines the pressure, temperature and velocity ratio between the respective jet and free-stream conditions. The VGJ is located at $x/\delta = 53$ behind the flat plate and based on the different free-stream flow conditions the resulting Reynolds number at this location is shown in Tab. 2. Where δ is the boundary layer thickness determined in experiments at jet location. The experiments identified the efficient range of flow control parameters i.e. pitch angle, $\alpha_{jet} = 30° - 45°$ and skew angle, $\beta_{jet} = 90° - 105°$ [21, 22]. The angle between the surface and the blowing direction is called the pitch angle. The skew angle is defined as the angle between the blowing axis and the free-stream direction. The flow control parameters used for the present simulations were among the best configuration of $\alpha_{jet} = 45°$ and $\beta_{jet} = 90°$.

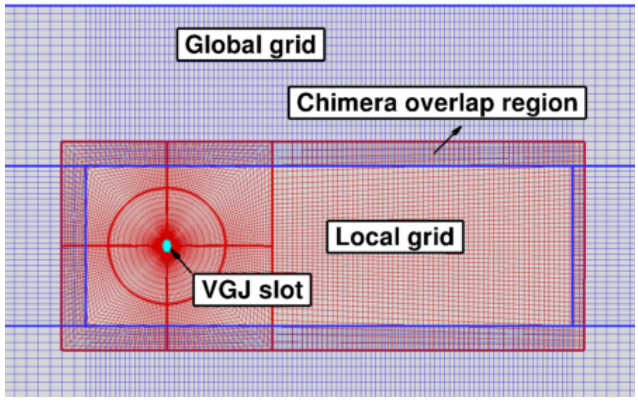


Fig. 4 Chimera grid setup of circular VGJ

Tab.2 Numerical test matrix

VGJ flow conditions			Free-stream conditions				VGJ ratios			
description	ρ_{jet} [kg/m^3]	m_{jet} [kg/s]	ρ_∞ [kg/m^3]	T_∞ [K°]	*M* [-]	*Re* [-]	$\frac{\rho_{jet}}{\rho_\infty}$	$\frac{p_{jet}}{p_\infty}$	$\frac{T_{jet}}{T_\infty}$	$\frac{U_{jet}}{U_\infty}$
1-warmjet	1.22	$2.58 \cdot 10^{-3}$	1.22	273	0.072	$7.67 \cdot 10^6$	1.0	1.032	1.0	2.5
2-warmjet	1.22	$2.40 \cdot 10^{-3}$	3.45	100	0.072	$33.0 \cdot 10^6$	0.35	1.020	2.73	4.36
3-warmjet	1.22	$2.80 \cdot 10^{-3}$	3.45	100	0.072	$33.0 \cdot 10^6$	0.35	1.065	2.73	7.38
4-coldjet	3.45	$4.0 \cdot 10^{-3}$	3.45	100	0.072	$33.0 \cdot 10^6$	1.0	1.020	1.0	2.63

4 Experimental results

Lift and drag of the airfoil are estimated by the pressure distribution and are only corrected by the dynamic pressure in the test section. All presented result are measured without fixed transition otherwise, the test cases are specially marked (non filled symbols, TF).

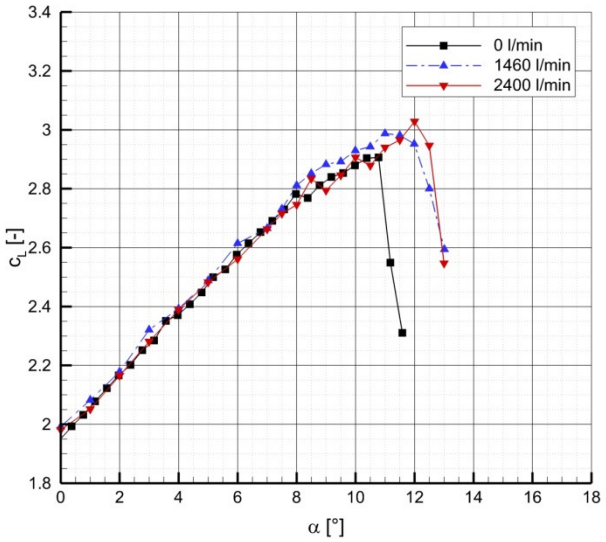


Fig. 5 Lift curves without and with AFC (Q=[1460, 2400]l/min, f=100Hz, dc=50%) at Re=4.2·106 and M=0.15

Fig. 5 shows lift curves without (black) and with (coloured) AFC. The Reynolds number is Re=4.2·10^6 and the Mach number is M=0.15. The airfoil, or rather the reference without AFC, stalls abrupt at an angle of attack (AoA) of α_{max}=10.8°. The measured maximum lift coefficient is $c_{L,max}$=2.91. With AFC two different flow rates Q=[1460, 2400] are tested. The actuation frequency is f=100Hz and the duty cycle is dc=50%. The linear part of the lift curves is unchanged whereas the maximum lift coefficient $c_{L,max}$ and the corresponding AoA α_{max} changes. Depending on the flow rate both values increases. With a flow rate of Q=2400l/min the benefits are $\Delta c_{L,max}$= 0.12 and $\Delta\alpha_{max}$=1.2°. To analyze the stall behaviour Fig. 6a displays corresponding pressure distributions (c_p) of the airfoil without AFC at three different AoA α=[6, 10.8, 11.6]°. At an AoA of α=6° the main element has an attached flow whereas the flow of the flap separates stream down the suction peak. At the maximum lift coefficient α_{max}=10.8° the suction peaks of the main element and the flap are higher. The stalled airfoil α=11.6° still have a significant suction peak on the main element and nearly constant pressure behind the suction peak to the trailing edge. So, the pressure distribution in the middle section indicates a trailing edge stall. The stall behaviour changes with AFC as seen in Fig. 6b. The figure includes three pressure distributions with AFC at the AoA α=[6, 12, 13]°. The suction peak of the main element increases from α=6° to α=12° at the maximum lift. With increasing AoA, the stalled airfoil, there is a small area with constant pressure behind the suction peak. Then, between x/c=0.15 and the trailing edge, the

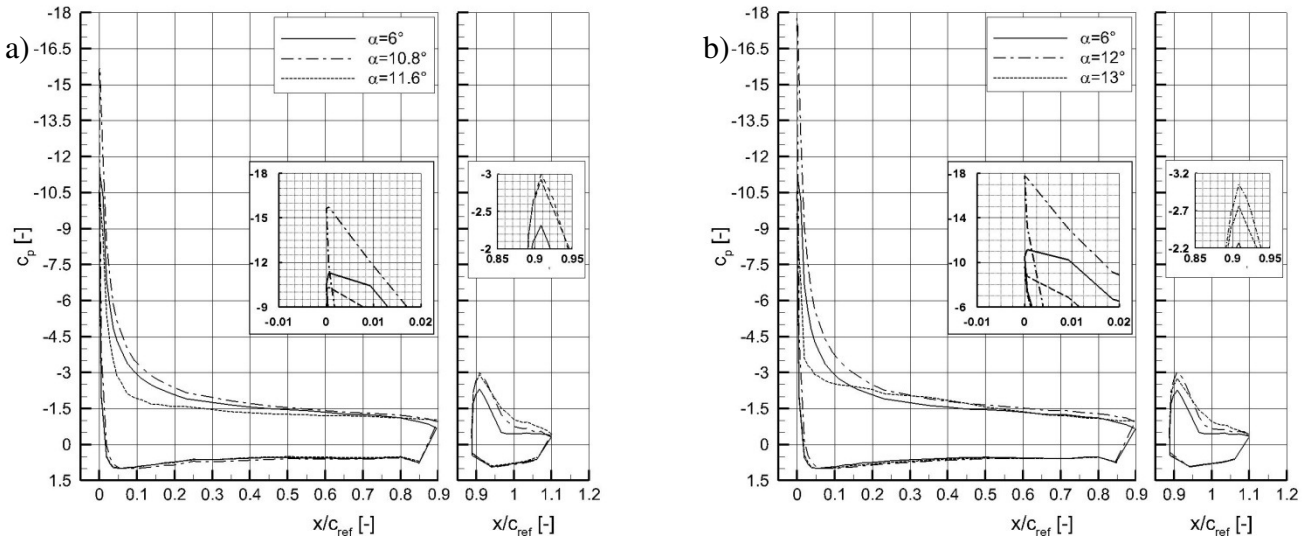


Fig.6 c_p without (a) and with (b) AFC (Q=2400l/min, f=100Hz, dc=50%) at Re= 4.2·10^6 and M=0.15

pressure increases and indicates a turbulent leading edge stall.

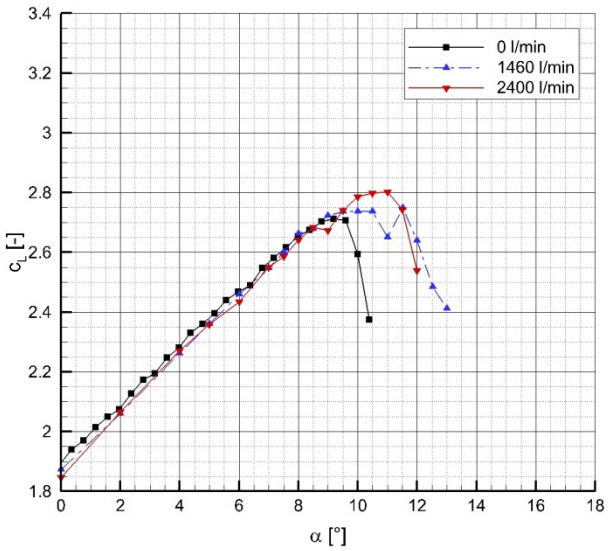
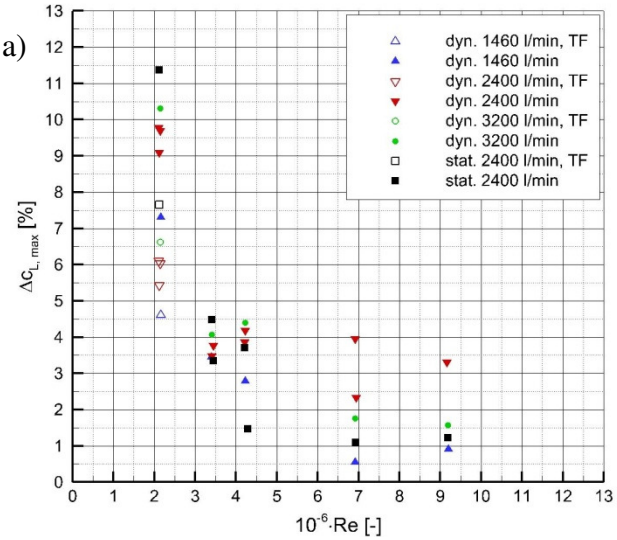


Fig. 7 Lift curves without and with AFC (Q=[1460, 2400]l/min, f=100Hz, dc=50%) at Re=9.2·106 and M=0.15

Lift curves with (black) and without (coloured) AFC at a higher Reynolds number of Re=9.2·10^6 and a Mach number of M=0.15 are presented in Fig. 7. The AFC parameters are identical to these in Fig. 5 and due to an identical Mach number the momentum coefficient c_μ is identical, too. Without AFC the airfoil stalls at a lower AoA and maximum lift coefficient compared to the reference at the Reynolds number Re=4.2·10^6 in Fig. 5. With AFC the maximum lift coefficient as well as the corresponding AoA increases depending on the flow rate. The differences to the reference are $\Delta c_{L,max}$= 0.09 and $\Delta\alpha_{max}$=1.8° with a flow rate of Q=2400l/min.

Figure 5 and 7 show that the reference, without blowing, changes with Reynolds number and thus the local reference at the corresponding Reynolds number has to be use to estimate the benefits ($\Delta c_{L,max}$, $\Delta\alpha_{max}$). In case of the maximum lift coefficient the percentage value will be used. The effect of the AFC system with different Reynolds number at a Mach number of M=0.15 is presented in Fig. 8a. It should be noticed that at the lowest Reynolds number Re=2.1·10·^{6} two references are used to calculate the benefits. The filled symbols represent the cases without fixed transition. Then the airfoil stalls due to the burst of a laminar separation bubble which results in higher benefits. The non filled symbols are the cases with fixed transition. The figure includes three dynamic AFC test rows. The actuation frequency is f=100Hz, the duty cycle is dc=50% and the flow rate varies Q=[1460, 2400, 3200]l/min. Additionally one test row with static AFC and a flow rate of Q=2400l/min is shown, too. All test rows increases the maximum lift coefficient compared to the local reference but the benefits become smaller with increasing Reynolds

a)



b)

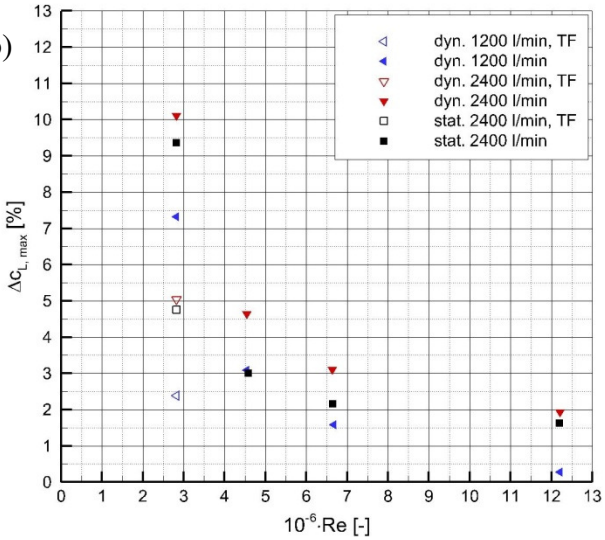


Fig.8 Lift increase with dynamic AFC (Q=[1460, 2400, 3200] l/min, f=100Hz, dc=50%) and static AFC (Q=2400l/min) at M=0.15 (a) and with dynamic AFC (Q=[1200, 2400] l/min, f=100Hz, dc=50%) and static AFC (Q=2400l/min) at M=0.2 (b)

number. Comparing the different flow rates of the dynamic test rows at the Reynolds number of Re=2.1·10^6 the benefits increases the higher the flow rate. The behavior changes with increasing Reynolds number. At a Reynolds number of about Re=4·10^6 are nearly identical. With higher Reynolds number the benefits with the flow rate of Q=3200l/min have a value between the test rows of Q=1460l/min and Q=2400l/min. This behavior indicates that there exits an optimum depending on the Reynolds number. Comparing the static and dynamic AFC test rows with identical flow rate the figure shows that in most cases, or rather at higher Reynolds number, dynamic AFC is superior to static AFC. At lower Reynolds number this differs. This trend becomes more clear presenting the results of test rows at a higher Mach number of M=0.2. At all Reynolds numbers dynamic blowing is superior to static blowing. The other trends described in Fig. 8, the influence of the Reynolds number and the flow rate, are visible in Fig. 8b, too. However, due to the limited test rows with different flow rates a maximum or optimum is not visible.

During the measurements most of the AFC parameters were varied. Summarizing the best AFC test cases independent of the AFC parameters result in Fig. 9a. The AFC parameters are in Tab. 3. The Mach number is M=0.15. The figure clearly show that the benefits $\Delta c_{L,max}$ decreases the higher the Reynolds number. At higher Reynolds number no AFC parameter combination results in comparable high benefits as measured at the lowest Reynolds number of Re=2.1·10^6. The benefits of $\Delta\alpha_{max}$ in Fig. 9b corresponds to this. So, these results and the results of Fig. 8 points to a non optimized AFC system for high Reynolds number or an influence of additional parameters. As already described, a constant flow rate results in a constant c_μ at different Reynolds number and constant Mach number. But other parameters changes. The actuators are blowing with gaseous nitrogen with a nearly ambient temperature whereas the wind tunnel is cooled down to 100K at the highest Reynolds number. This changes the density of the free stream, too. Additionally, the free stream velocity decreases with increasing Reynolds number and constant Mach number. Summarizing this, c_μ is constant whereas the density ratio ρ_{jet}/ρ_∞, the temperature ratio T_{jet}/T_∞ and the velocity ratio U_{jet}/U_∞ vary. A more detailed study of the measurement results is presented in [13].

Tab. 3 AFC parameters for the best test cases

Re·10^{-6} [-]	Q [l/min]	c_μ [-]	f [Hz]	dc [%]
2.1	2400	0.0053	65	50
3.4	2400	0.0053	150	50
4.2	2800	0.0079	100	50
6.9	2400	0.0053	100	50
9.2	2400	0.0053	100	50

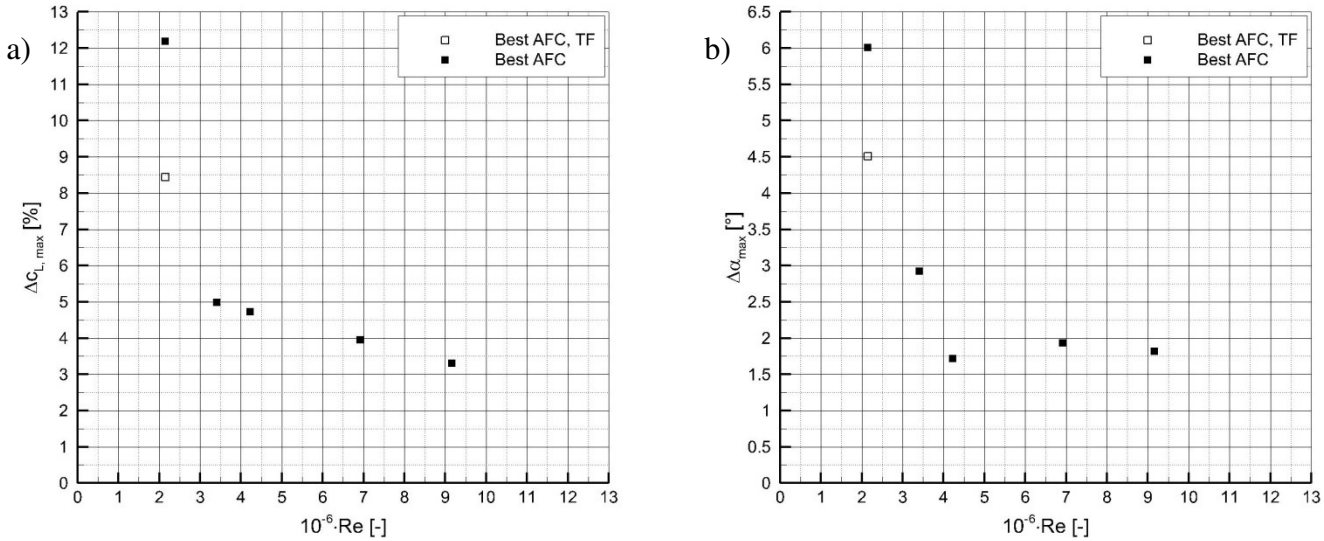


Fig.9 Lift increase (a) and increase of the corresponding AoA (b) of the best AFC test cases at M=0.15

4.1 Numerical results

Fig. 10 shows the spanwise averaged skin friction coefficient (c_f) for the different test cases mentioned in Tab. 2. The jet location is marked with an arrow at x/δ=53. The effectiveness of the VGJ was judged w.r.t the reference case labeled as 1-warmjet computed under normal free-stream conditions at same Mach number. The black solid line shows the baseline (no AFC) results. It can be observed from the results that the combination of the warm jet with *VR*=4.36 interacting with the cold free-stream conditions decreases the efficiency of the VGJ. This means that the warm jet at higher Reynolds number under cold

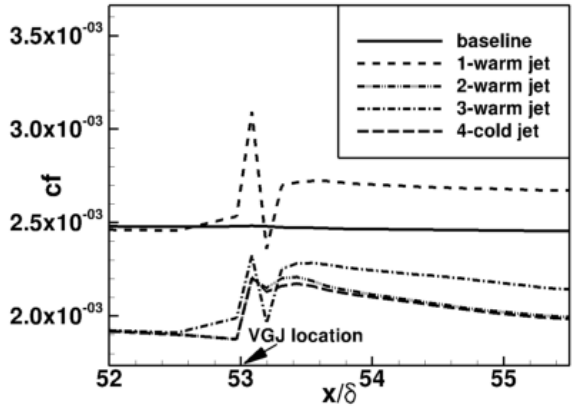


Fig. 10 Skin friction distr. downstream of the VGJ

conditions decreases the effectiveness of the VGJ as observed in the AFC experiments over an airfoil. This deficiency could be due to

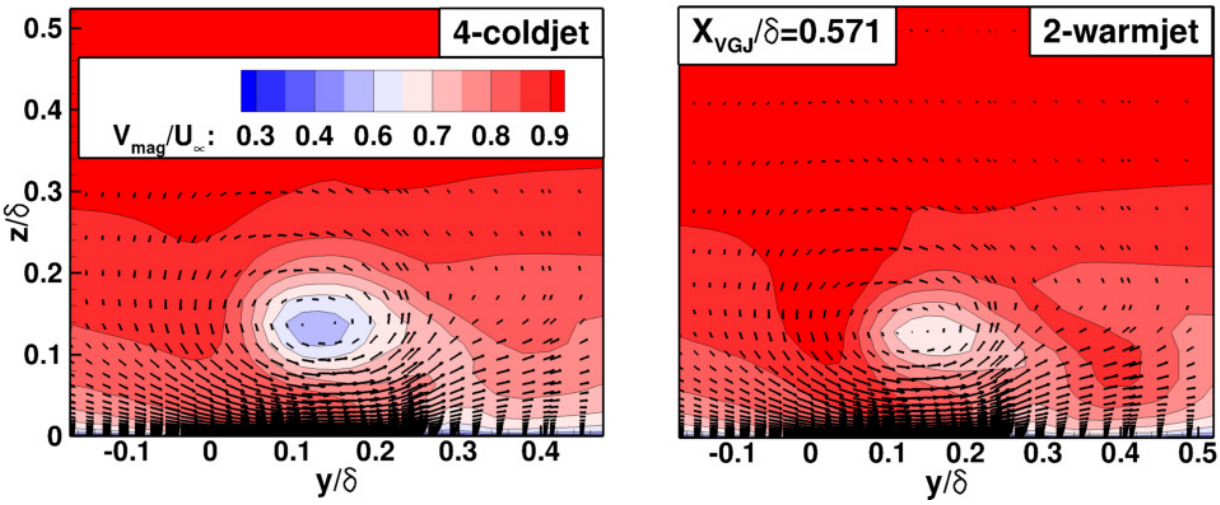


Fig. 11 Normalized velocity magnitude contours in the cross-flow plane

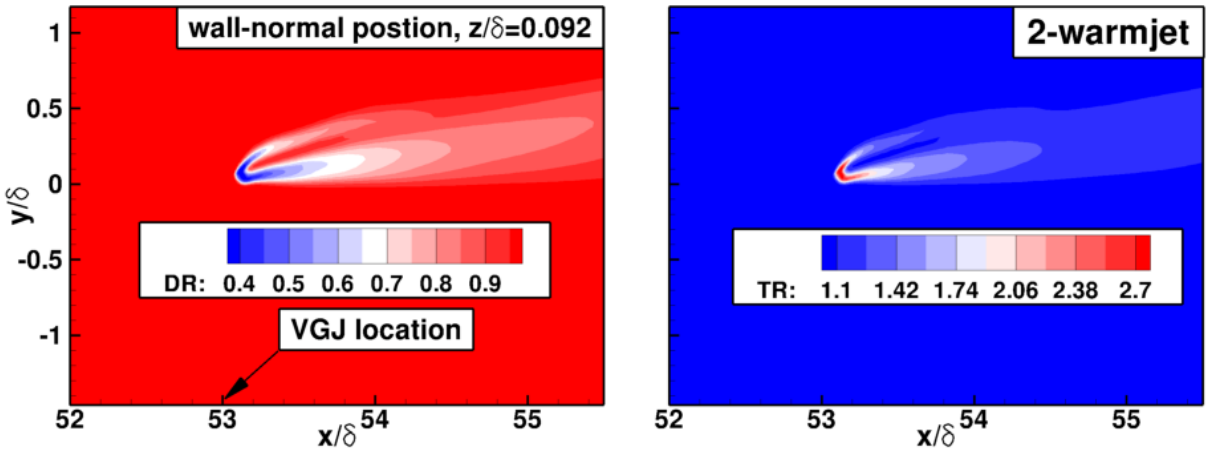


Fig. 12 Influence of density and temperature at wall-normal position

the strong inertial effects at higher Reynolds number resulting in negligible viscous effects.

The similar trend was observed with 4-cold jet with VR=2.63 under cold free-stream conditions. But the effect of higher PR and VR=7.38 with 3-warmjet does influence the results downstream of the VGJ by predicting higher c_f value. It is worth mentioning that the computations with 2-warmjet and 4-coldjet cases were carried out at approximately same c_μ=0.0013, but resulted in different velocity ratios of $VR_{2\text{-warmjet}}$=4.36 and $VR_{4\text{-coldjet}}$=2.63, similar PR's.

Fig. 11 shows the normalized velocity magnitude contours downstream of the VGJ at X_{VGJ}/δ=0.571, where y/δ and z/δ defines the spanwise and wall-normal position. The results showed the presence of the vortex embedded within the boundary layer for both the cold and warm jet under cryogenic conditions. It can be observed that both the cold and warm jet predicted similar vortex core position. The effect of temperature and density was shown in fig. 12 at wall-normal position of z/δ=0.092 for the 2-warmjet. A strong influence of the temperature and density variation was noticed within immediate vicinity of the VGJ. But this effect diminishes very rapidly due to the strong upstream body force in the form of heavier density nitrogen gas operating at higher Reynolds number.

5 Conclusion

This paper describes experimental and numerical results about vortex generator jets for AFC at high Reynolds number. The experiments were performed with cryogenic conditions and a warm actuator jet. The experiment allows the following conclusions:

$\Delta c_{L,max}$ [%] reduces the higher the Reynolds number which was observed at two different Mach numbers

The optimum flow rate, or rather momentum coefficient, with respect for high $\Delta c_{L,max}$ changes with increasing Reynolds number

In most cases dynamic AFC is superior to static AFC.

It has been demonstrated numerically that the efficiency of the AFC conducted over the flat plate decreases at higher Reynolds number. The results showed that warm jet interacts with the cryogenic cross flow but the hot stream of fluid gets diffused quickly resulting in degradation of the VGJ performance. Future numerical investigation should consider the AFC at higher free-stream Mach numbers under cryogenic conditions and also study the real-gas effects. Then it should be clarified if VGJ's for AFC could be optimized for high Reynolds numbers for ambient and cryogenic conditions to enhance the performance.

Acknowledgments

The work was funded by the German government within the framework of the 4th Aerospace Research Program (LuFo IV) in the project M-Fly (coordinated by the German Aerospace Center DLR) for the experimental investigations. The experiments in the DNW-KKK performed in close collaboration with the DLR Braunschweig, and have been extensively supported specifically by Jochen Wild, Vlad Ciobaca, Martin Sitzmann, Ralf Rudnik and Marvin Gülzow. Furthermore the authors would like to thank Heribert Bieler and Burkhard Gölling (both Airbus Operations GmbH) for their ongoing support.

References

[1] Godard, G., and Stanislas, M., "Control of a Decelerating Boundary Layer. Part 3: Optimization of Round Jets Vortex Generators," Aerospace Science and Technology, Volume 10, Issue 6, 2006, pp. 455-464.

[2] Johnston, J.P., and Nishi, M., "Vortex Generator Jets-Means for Flow Separation Control," AIAA Journal, Vol. 28, No. 6, 1990, pp. 989-994.

[3] Godard, G., and Stanislas, M., "Control of a decelerating boundary layer. Part 1: Optimization of passive vortex generators," Aerospace Science and Technology, Volume 10, Issue 6, 2006, pp. 181-191.

[4] Bauer, A. B., "Hidden Vortex Generators," US-Patent 4039161, 1977

[5] McManus, K., Legner, H.H., and Davis, S.J., "Pulsed Vortex Generator Jets for Active Control of Flow Separation," 25th AIAA Fluid Dynamics Conference, Colorado Springs, USA, 20-30 June 1994, AIAA Paper 94-2218.

[6] Johari, H., and Rixon, G.S., "Effects of Pulsing on a Vortex Generator Jet," AIAA Journal, Vol. 41, No. 12, 2003, pp. 2309-2315.

[7] Ortmanns, J., Bitter, M., and Kähler, C.J., "Dynamic vortex structures for flow-control applications," Experiments in Fluids, Vol. 44, No. 3, 2008, pp. 397-408.

[8] AGARD Fluid Dynamics Panel, "Scale Effects on Aircraft and Weapon Aerodynamics," AGARD-AG-323, July 1994

[9] Wild, J., Pott-Pollenske, M., and Nagel, B., "An integrated design approach for low noise exposing high-lift devices," 3rd AIAA Flow Control Conference, San Francisco , USA, 5-8 June., 2006, AIAA Paper 2006-2843.

[10] Wild, J., Wichmann, G., Haucke, F., Peltzer, I., and Scholz, P., "Large Scale flow control experiments within the German Flow Control Network," 47th AIAA Aerospace Sciences Meeting, Orlando, USA, 5-8 Jan., 2009, AIAA Paper 2009-0530.

[11] Scholz, P., Kähler, C.J., Radespiel, R., Wild, J., and Wichmann, G., "Active Control of Leading-Edge Separation within the German Flow Control Network," 47th AIAA Aerospace Sciences Meeting, Orlando, USA, 5-8 Jan., 2009, AIAA Paper 2009-0529.

[12] Scholz, P., Mahmood, S.S., Casper, M., Radespiel, R., Sitzmann, M., and Ciobaca, V., "Experimental and Numerical Investigations on the Control of Airfoil Stall using Vortex Generator Jets", 5th Flow Control Conference, Chicago, Illinois, June 28-1, 2010, AIAA Paper 2010-4250.

[13] Casper, M., Scholz, P., Radespiel, R., Wild, J., Ciobaca, V., „Separation Control on a High-Lift Airfoil using Vortex Generator Jets at High Reynolds numbers," AIAA-2011-3442 41st AIAA Fluid Dynamics Conference and Exhibit, Honolulu, Hawaii, June 27-30, 2011

[14] Scholz, P., Casper, M., Ortmanns, J., Kähler, C.J., and Radespiel, R., "Leading-Edge Separation Control by Means of Pulsed Vortex Generator Jets," AIAA Journal, Vol. 46, No. 4, 2008, pp. 837-846.

[15] Mahmood, S., Radespiel, R., "RANS Simulation of Jet Actuation in a boundary layer flow using Chimera Grids", Deutscher Luft- und Raumfahrtkongress, Aachen, Germany, September 8-10, 2009.

[16] Ortmanns J., Aktive Grenzschichtbeeinflussung mittels pneumatischer Wirbelgeneratoren bei großen Reynoldszahlen, Institut für Strömungsmechanik, TU Braunschweig, Dissertation ZLR-Forchungsbericht 2009-03.

[17] Schwamborn D., Gerhold T. and Heinrich R., "The DLR TAU-Code: Recent Applications in Research and Industry", *In: P. Wesseling, E. Oate, J. Priaux (Eds), ECCOMAS CFD*, *Egmond aan Zee*, The Netherlands, 2006.

[18] Madrane A., Raichle A. and Stuermer A., "Parallel implementation of a dynamic overset unstructured grid approach", ECCOMAS 2004, Jyväskylä, July 24-28, 2004.

[19] Schwarz T., "An Interpolation Method Maintaining the Wall Distance for Structured and Unstructured Overset Grids", In: Proceedings of the CEAS 2009 conference. CEAS 2009 European Air and Space Conference, October 26-29, 2009, Manchester, UK.

[20] Menter F.R., "Zonal two-equation k-w turbulence model for aerodynamic flows", *AIAA paper*, AIAA 93-2906, 1993.

[21] Adcock, J. B., "Real-Gas Effects Associated With One-Dimensional Transonic Flow of Cryogenic Nitrogen", NASA TN D-8274, 1976.

[22] Ortmanns J. and Kähler C.J., A Single Round Vortex Generator Jet at High Reynolds Number, FLUCOME 2007, Tallahassee, Florida, USA.

[23] Commercial CFD software by POINTWISE, Inc. http://www.pointwise.com/gridgen

An experimental case study of an Air Outlet: Drag components and Efficiency parameters

T. Gericke, P. Scholz and R. Radespiel
Technische Universität Braunschweig , Germany

Keywords: *Air Outlet, Drag, Efficiency, Particle Image Velocimetry*

Abstract

A model of an Air Outlet was built to accomplish an experimental case study of its drag and efficiency. The Air Outlet basically consists of a bend channel and an adjustable flap. Experiments were conducted in a low speed wind tunnel at different flap angles and mass flows. In addition the modular conception of the model allowed testing of two different outlet-wall connections. For the computation of drag the control-volume approach is used. Furthermore the drag components of the different vortices are computed using the momentum equation and an efficiency parameter E_{ff} is established. The results clearly show that a flap angle of 45° produces the greatest loss of momentum and thus the maximum drag. Furthermore, the comparison of the total drag for different flap angles make clear that a flap angle of 45° results in the highest total drag.

1 Introduction

It is generally expected that the next aircraft generations will provide better fuel economy, emission- and noise reduction with an increased safety and reliability. Therefore it is necessary to use configurations and architectures that promise new optimization potentials. Nowadays bleed air is used in aircrafts to get compressed and fresh air for the cabin air conditioning, which decreases the engines efficiency and durability mainly due to the bleed air extraction inside the engines. Here, a bleedless engine design should lead to an increased efficiency and durability when the engines only generate propulsion. Furthermore the expected reduction of tubings which are needed in a bleedless engine design throughout the aircraft in comparison to a bleed engine design probably increases space and decreases weight. For a bleedless engine design new and additional air intakes and outlets are fundamental components in order to achieve the goal of a bleedless engine design.

The purpose of this work is to investigate an Air Outlet (AO) mainly with standard particle image velocimetry (2C2D PIV) and stereoscopic particle image velocimetry (3C2D PIV) [1] in order to better understand the flow phenomena in and around an AO. Here the main focus is the drag of the AO and its different components as well as a definition of a reasonable efficiency parameter [8]. The AO shown in Fig. 3 consist of a bended channel

and an adjustable flap. The outlet is not covered with an aerodynamically formed covering in which the flap is submerged. In addition two different outlet-wall connections are investigated (Fig. 1), in order to study the different detachment behavior of those geometries. The total drag and its components are calculated by deriving the forces from the wake and flow far field information, acquired with PIV, itself using the control-volume approach. These approache remove the necessity of additional and /or intrusive instrumentation behind the model or of the model itself [2-6].

2 Theoretical Background

The classical control-volume approach is a common used technique to relate the thrust and drag of an object to the wake and far flow field. Required information of the flow field are velocity, density, pressure, shear stress and gravity. In most of the cases gravity and shear stress can be neglected. The momentum equation in integral form (Eq. 1) leads directly to the thrust or drag, by integrating the flow field information along a line (2D case) or over an area (3D case).

$$\int \frac{\partial(\rho u_i)}{\partial t} dV = -\int u_i \rho u_j dA_j + \int p dA_i - \int \tau_{ij} dA_j + \int \rho g_i dV \qquad (1)$$

In most of the cases the outer contour of the control volume is so orientated that the free stream velocity is perpendicular to the in- and outlet faces and parallel to the side walls. This restriction simplifies most of the terms in the equation. PIV does not provide pressure information, therefore is it necessary to use a formulation in which the pressure can be eliminated or has to be measured separately. Furthermore, the size of the outer contour has to be chosen very carefully in order to not miss any flow related information. However, the distance to the boundarys of the control-volume is absolutely arbitrary.

3 Experimental Setup

The flap of the AO model is made out of glass fiber reinforced plastic, has a wingspan of b 180 mm and a chord length of c 222 mm. Furthermore, mainly plexiglas was used for other components in order to attain optical access for PIV measurements inside the model (Fig. 3). Additionally, the modular conception of the model allowed the testing of two outlet wall connections (round and bevel).

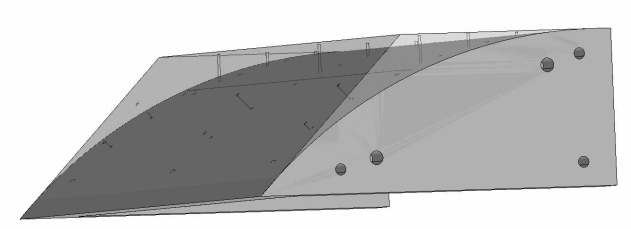


Fig. 1 Both outlet-wall connections assembled into one plot. Round (grey) and bevel (transparent)

Oil flow pictures were generated using fluorescent paint and ultraviolet light. About 78 pressure ports were embedded in total into the flap with 39 pressure ports on every side. Here, the pressure ports were arranged so that 1/3 of the ports build one line on the left side, in the middle and on the right side of the flap surface. Moreover, 20 additional pressure ports were installed before and behind the AO model all aligned with the middle section of the model and a total of 50 ports are integrated in the bended channel. All pressure ports were connected to a PSI 8400 SDI multichannel pressure transducer with an accuracy of 0.12%.
Experiments were performed in a Göttingen type wind tunnel with a test section of 0.80 m x 0.80 m (Tu = 0.2% at 53 m/s and 0-25 kHz) located at the Technical University of Braunschweig. Freestream velocity was constantly set to Mach = 0,34 giving a Reynolds number of Re = 310796. The volume flow through the AO was provided by a radial-blow fan with an Δp_{max} of 100 mbar and measured with the thermal mass flow measuring system t-mass 65I at a distance of 6 m before the AO model. Directly before the AO model entrance a 1,1 m long plexiglas tube was installed to ensure appropriate inflow conditions. Required forces were determined by using the control-volume

approach [7]. Particle image velocimetry (PIV) measurements were made with a setup for standard PIV (Fig. 2) and stereoscopic PIV consisting of one or two LaVision Imager pro X 11 megapixel cameras with a zeiss lens of 105 mm focal length. A Litron double-cavity pulsed Nd:Yag laser with 180 mJ per pulse at a wavelength of 532 nm was used to illuminate the flowfield.



Fig. 2 2C2D PIV setup

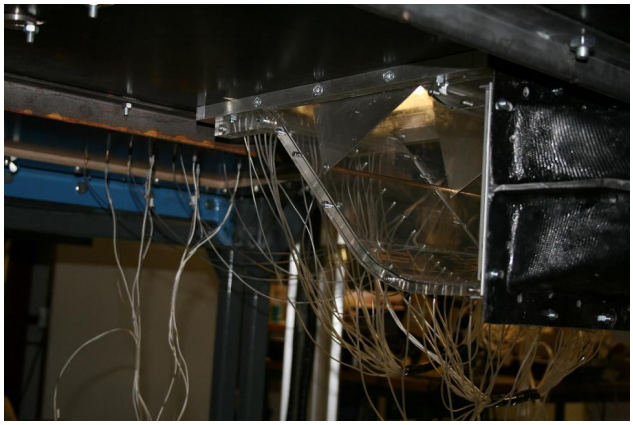


Fig. 3 AO model integrated into the wind tunnel

3.1 PIV measurements and analysis

As mentioned above the PIV measurements were made with a setup for standard PIV and stereoscopic PIV mainly consisting of one or two LaVision Imager pro X 11 megapixel cameras and a Litron double-cavity pulsed Nd:Yag laser. In-house build Scheimplufg-adapter were used additionally for the stereoscopic setup. The time between images is chosen as a function of the in-plane speed and was set to 10 µs for 115 m/s for standard PIV and to 12 µs for stereoscopic PIV. The seeding material consists of sunflower oil droplets with a particle size of about 2µm. In total five different measurement planes were investigated (Fig. 4). Measuring planes 1-3, positioned in the middle section of the model, cover the flow inside the AO, before the flap and the aggregation of the free stream and AO flow. All planes were measured using the standard PIV setup. Plane 4 and 5 were measured using the stereo PIV setup in order to cover all three velocity components of the highly 3D wake flow. Plane 4 is located 30 and plane 5 120 mm behind the model. Both planes cover an area of about 0.08 m².

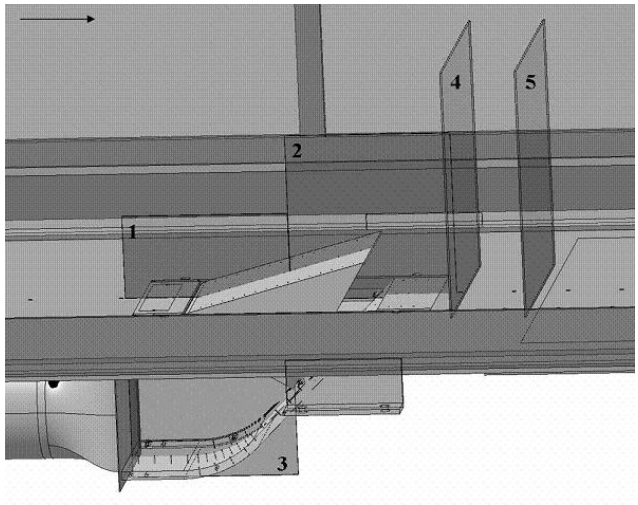


Fig. 4 AO model integrated into the wind tunnel showing the PIV measuring planes

A variety of different flap angles and mass flows was measured and a total of five hundred image pairs for each measurement were recorded. To minimize reflection interference on the images a fluorescent paint called FP R6G (Flow Visualisation Components) is used to coat the bottom of the measurement planes. This paint has the property of absorbing the laser light with a maximum absorption of 540 nm to about 99%. The remaining part is reflected in a spectral range shifted to orange with a maximum emission of 625 nm. Optical filters are used to filter wavelengths other than 532 ±5 nm in order to reduce reflections.

Data analysis was taken out with Davis 7.2 (LaVision, Göttingen). First of all image pairs were processed using count(xy)-(avg/min /max)(x,y) of time series.

Here two passes through all images are necessary which are 1) calculation of the minimum (or average or maximum) pixel intensity value at each pixel position and 2) subtraction of this background image from all images in the set. Thereafter a multi pass with a window size of 64 x 64 and 32 x 32 was conducted. Overlap was set to 50% using 1 and 3 passes. Image postprocessing was realized using correlation peak validation.

3.2 Momentum analysis

The momentum equation was solved by integrating the velocity over the area of interest using Tecplot 360 2009 Release 1 (Tecplot, Inc, Bellevue). 3C2D PIV data showed a small divergence for v_z in x-direction concerning all measurements. This was corrected by 1) calculating the slope of the v_z graph at different y positions and 2) using the calculated slope to homogenize the v_z component by calculating v_z corrected v_{zc}. Another correction was carried out in order to remove the velocity increase due to the displacement, caused by the flap at different flap angles. Average velocity was calculated for every measurement in an undisturbed flow region and the difference between the velocity at the inlet and outlet was subtracted from the outlet velocity. This approach eliminates the pressure term from Eq. 1. Gravity were also neglected in this case. Taking into account those corrections the momentum was calculated using the Riemannsches Integral method available in Tecplot 360. Momentum at the inlet and outlet of the AO is integrated along a line. Assuming that the velocity distribution is constant over the cross section, it is then possible to compute the whole momentum from 2C2D PIV measurements with only the flow information from the middle section. The different measurements planes enable the computation of total AO drag (cw AO_{total}), the momentum loss inside the AO (cw $AO_{internal}$) and the momentum loss due to the vortices shed from the flap (cw $AO_{vortices}$). Total drag (Eq. 2) is calculated by subtracting the momentum of plane 5, measured with AO, from plane 5 measured without AO.

$$\dot{p} = \rho \int u_z u_z dA \qquad (2)$$

Momentum accumulated in the vortices is the sum of the momentum in x- and y-direction. Equation 3 shows the formula for the momentum in x-direction, which is the same for the y-direction.

$$\dot{p} = \rho \int u_x u_z dA \qquad (3)$$

Momentum loss inside the AO is calculated using the same principles.
Concerning the efficiency of the system a efficiency parameter called E_{ff} is introduced in Equation 4 [8]. Here the volume flow V, the pressure difference between free air and the entrance of the AO Δp, drag D and free stream velocity u_∞ is considered.

$$E_{ff} = \frac{\dot{V}(-\Delta p)}{D u_\infty} \qquad (4)$$

The numerator in the expression represents the power output which can also be called the useful work, whereas the denominator represents the power input. The power input, in this case, is the power needed to overcome the drag of the model. Moreover, the effective drag coefficient cw_{eff} is calculated using Equation 5. A total of four coefficients were calculated namely cμ AO, cw AO_{total}, cw $AO_{internal}$ and cw $AO_{vortices}$ using A_{ref}, which is equal to the flap surface. The momentum coefficient cμ AO is calculated by taking into account the momentum leaving the AO.

$$cw_{eff} = cw + c\mu \qquad (5)$$

4 Results

4.1 Flowfields

Results of the 2C2D and 3C2D PIV measurements for AO bevel outlet-wall connection, 45° and 1,992 kg/s are assembled into one plot in Fig. 5. They cover the whole flowfield around the AO, in the center section of the AO model before and after the flap and inside the AO.

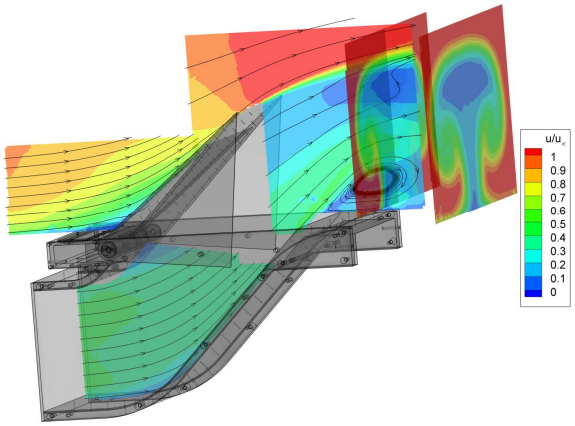


Fig. 5 2C2D and 3C2D PIV measurements assembled into one plot for AO bevel outlet-wall connection, 45° and 1,92 kg/s

Here a complex 3D flow structure behind the AO with a vortex shed from the sides of the flap and a recirculation zone downstream of the outlet-wall connections is clearly visible.

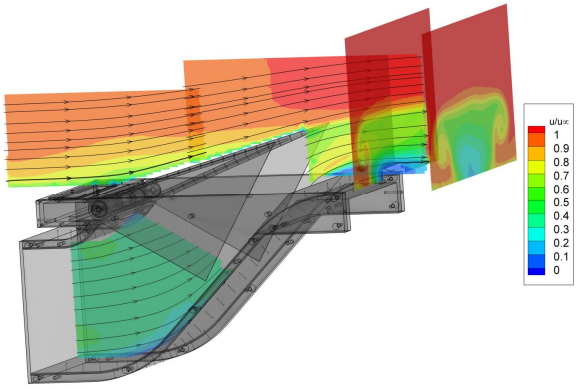


Fig. 6 2C2D and 3C2D PIV measurements assembled into one plot for AO bevel outlet-wall connection, 21° and 1,76 kg/s

Furthermore, indications of three additional flow detachment zones, two inside the AO in the upper and lower redirection region and one at the beginning of the flap (left hand side), are observed. Figure 6 shows the same measurement planes for bevel outlet-wall connection, 21° and 1,76 kg/s. Compared with 45° and 1,92 kg/s the main differences are 1) the vortex shed from the flap is only half the size 2) the recirculation zone downstream of the outlet-wall connection is much smaller, 3) the flow field leaving the AO is almost uniformly distributed and 4) the three additional flow detachment zones are reduced. The extended flap causes a loss of momentum whereas the outlet flow of the AO reduces the momentum loss due to an extra energy supply into the flow. Furthermore, depending on the flow around the flap, different amounts of energy are fed into the vortices, which also generate drag. Detailed 3C2D recordings of plane 5, downstream of the AO, for the above mentioned cases are displayed in Fig. 7 and 8.

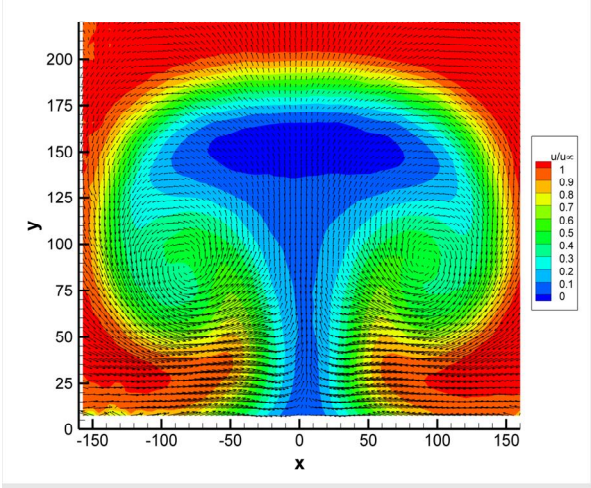


Fig. 7 3C2D PIV measurement (plane 5) for AO bevel outlet-wall connection, 45° and 1,92 kg/s

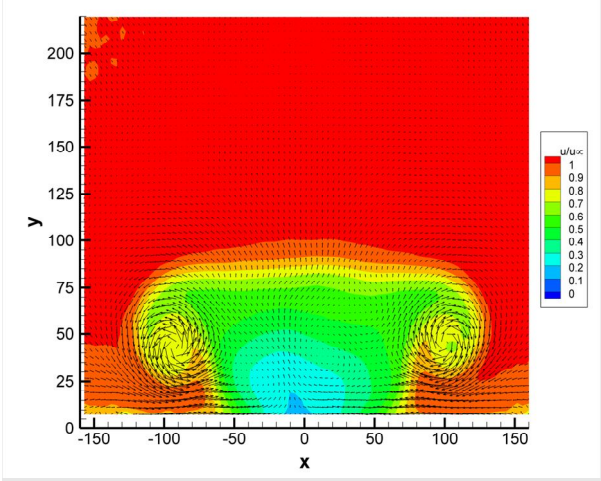


Fig. 8 3C2D PIV measurement (plane 5) for AO bevel outlet-wall connection, 21° and 1,76 kg/s

An oil flow picture (Fig. 9) shows the streamlines for the before mentioned 45° and 1,92 kg/s case. The separation line upstream of the flap is clearly visible. Two vortices develop downstream of the AO opening and the streamline curvature in this area shows a convergent behavior at the reattachment zone. Moreover, the streamlines on the flap are divergent and therefore form the vortices shed from the flap.

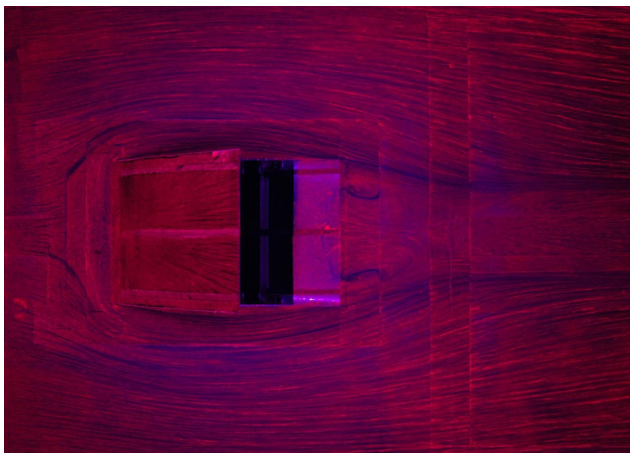


Fig. 9 Oil flow picture of AO bevel outlet-wall connection, 45° and 1,92 kg/s

4.2 Drag and Efficiency

Momentum and drag coefficients for both configurations, diagrammed in Fig. 10 and 11, show that the coefficients at a flap angle of 45° are much higher than for all other cases.

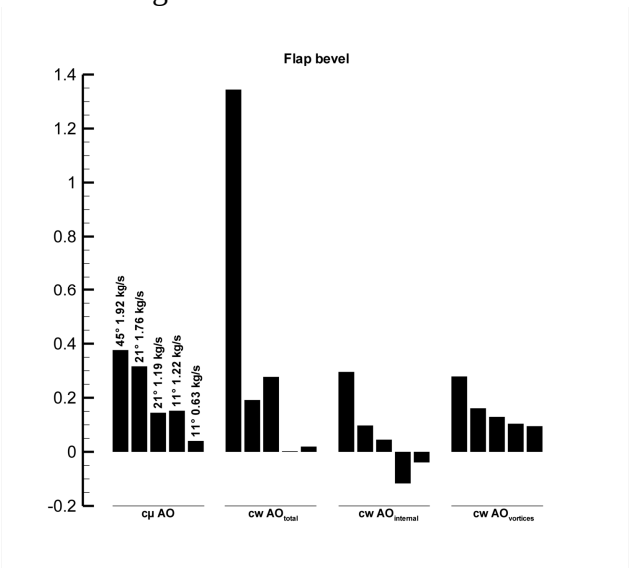


Fig. 10 Momentum and drag coefficients of AO bevel outlet-wall connection

Moreover, cw AO_{total} for 45° predominate all other cases by at least 75%. Vortex drag coefficient cw AO_{vortex} rises slightly until a flap angle of 21°, the increase to 45° is twice the increase for smaller flap angles (about 40%). The inner drag coefficient cw $AO_{internal}$ is negative (momentum gain) due to the acceleration of the internal flow for a flap angle of 6° and 11° and positive (momentum loss) for 21° and 45°.

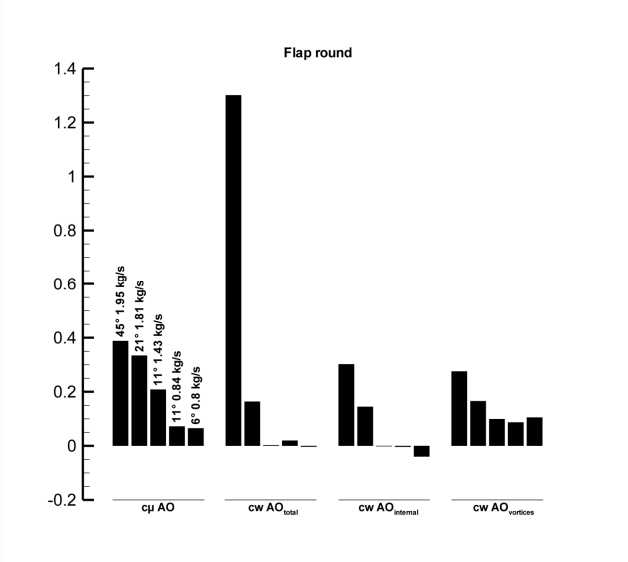


Fig. 11 Momentum and drag coefficients AO round outlet-wall connection

The introduced effective drag coefficient cw_{eff} (Fig. 12) for a given flap angle and different mass flows stays at the same level with increasing mass flows.

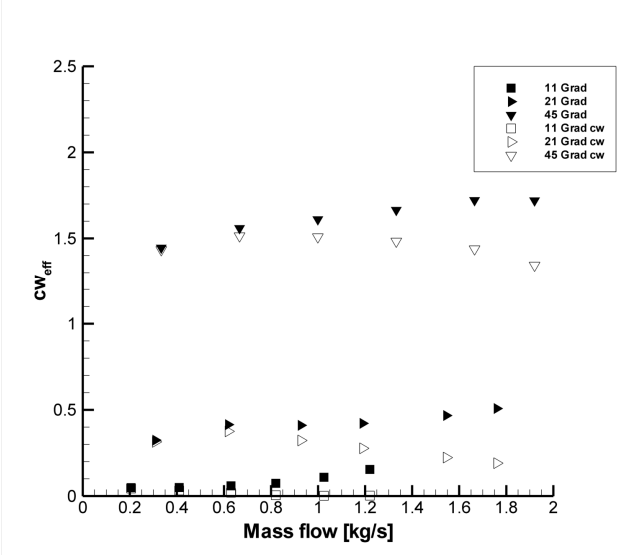


Fig. 12 cw_{eff} over mass flow for AO bevel outlet-wall connection 11°, 21° and 45°

For a flap angle of 45° cw_{eff} is three times higher then at 21° and even more when compared with 11°. Furthermore the effect of cμ AO can clearly be seen by the diverging plots for cw_{eff} and cw. The reader should note that the energy necessary to produce cμ AO is not taken into account. E_{ff}, the efficiency parameter, almost shows no effect of the power output compared to the power input at a flap angle of 45° (Fig. 13). This means that the power input, drag times free stream velocity, is predominant and with this mass flows the contribution of the power output to drag reduction is negligible.

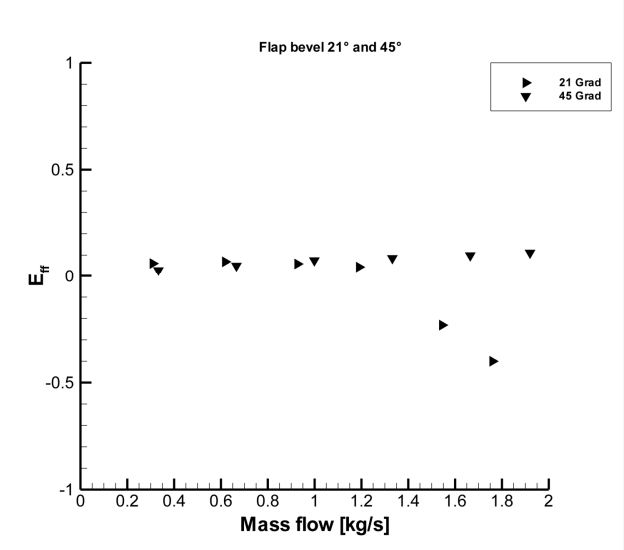


Fig. 13 Efficiency parameter over mass flow for AO bevel outlet-wall connection 21° and 45°

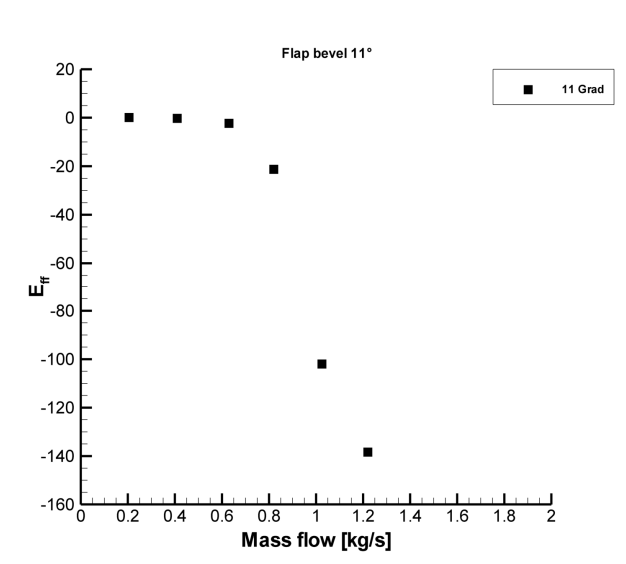


Fig. 14 Efficiency parameter over mass flow for AO bevel outlet-wall connection 11°

At 21° (Fig. 13) a mass flow of more than 1,2 kg/s enhances the efficiency of the systems slightly, whereas at 11° (Fig. 14) a mass flow over 0,6 kg/s increase the efficiency to 140% at 1,2 kg/s. Note that the negative E_{ff} results from the negative pressure difference between the free air and the entrance of the AO Δp. This means that a negative E_{ff} is not equivalent with a poor efficiency. In fact, the results show that a negative E_{ff} means that the power output is greater then the power input. This on the other hand doesn't mean that thrust is generated.

5 Conclusion

Momentum and drag coefficients show a clear dependency of the flow phenomena on the flap angle and the mass flow, whereas the outlet-wall connections barely make a difference. Except for drag coefficient cw AO_{inner}. In this case the smaller outlet area of the outlet-wall connection bevel leads to a higher outlet velocity and therefore to a higher momentum which reduces cwe_{ff}. It should be noted that the before mentioned reduction of cw_{eff} is small. In terms of heat problems, with the structure, in the vicinity of the outlet-wall connections the bevel reduces those due the bigger recirculation zone and hence a much smaller flow velocity over a greater area directly at the wall. cw_{eff} mostly depends on the flap angle. Different mass flows only have a little influence in this case. Which make clear that the total drag level, of the system, strongly depend on the flap angle. Results clearly show that a flap angle of 45° is not useful due to the high drag. The introduced efficiency parameter E_{ff} points out that at a flap angle of 45° the power output is negligible compared to the power input. Whereas the efficiency at a flap angle of 21° and 11° increases in consequence of the higher power output and lower power input. Nevertheless, this approach neglect the energy consumed, for example, by a jet pump to generate the power output and that the power output reduces the drag. In cwe_{ff} the power output (cμ) is added to cw resulting in a higher drag with increasing mass flow and flap angle. Future work should include the creation of an efficiency parameter which includes the energy

used to create the power output and the loss of momentum due to turbulent kinetic energy and dissipation.

References

[1] Raffel, M., Willert, C., Wereley, S. and Kompenhans, J., Particle Image Velocimetry, a Practical Guide, 2nd ed., Springer-Verlag, Berlin, Heidelberg, New York, 2007.

[2] Drela, M., Power Balance in Aerodynamic Flows, AIAA Journal, AIAA, New York, Vol. 47, No. 7, July 2007.

[3] van Oudheusden, B. W., Scarano, F., Roosenboom, E. W. M., Casimiri, E. W. F. and Souverein L. J., Evaluation of integral forces and pressure fields from planar velocimetry data for incompressible and compressible flows, Springer-Verlag, Berlin, Heidelberg, New York, Experiments in Fluids 2007 43:153-162.

[4] van Oudheusden, Casimiri, E. W. F. and Scarano, F., Aerodynamic Load Measurement of a Low Speed Airfoil using Particle Image Velocimetry, 45th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 2007.

[5] van Oudheusden, Scarano, F. and Casimiri, E. W. F., Non-intrusive load characterization of an airfoil using PIV, Springer-Verlag, Berlin, Heidelberg, New York, Experiments in Fluids 2006 40:988-992.

[6] Unal, M. F., Lin, J. C. and Rockwell, D., Force prediction by PIV Imaging: A momentum-based approach, Journal of Fluids and Structures, 1997 11, 965-971.

[7] Batchelor, G. K., An Introduction to Fluid Dynamics, Cambridge: Cambridge University Press, 1967.

[8] Rokus, F. and Troller, TH., Tests on Ventilation Openings for Aircraft, Journal of the Aeronautical Sciences, April 1936.

Wing-Tip Shape and the Axial Flow Structure in Wing-Tip Vortices

R. M. Buffo and R. Hoernschemeyer
Institute of Aeronautics and Astronautics, RWTH Aachen University, Germany

Keywords: *Vortex, Axial, Wing-Tip Shape*

Abstract

In order to identify the impact of different wing-tip shaping on vortex structure and axial velocity near the centre of the vortex core experiments using the 3C-PIV technique have been conducted. Three different shapes were used. They differ in the rounding of the tip side as well as the rounding of the planform layout between the leading edge and the tip edge. The experiments focus on the development of the vortex from 1 to 6 chords downstream of the wing. The core centre axial velocity is related to the intensity of core circulation. The distribution of axial vorticity along the span and the position of the core developing along the chord hold responsible for the radial distribution of circulation. This way it is possible to assign vortex characteristics and downstream behaviour to geometrical features of the tips.

Nomenclature

b, s, c	= wing span, half span, chord
r, r_c	= vortex radius, core radius
t	= airfoil thickness
x, y, z	= Cartesian coordinates
u, v, w	= local vel. in Cart. Coordinates
v_t	= tangential velocity
Re	= Reynolds number
S, A	= area of total wing, one wing-tip
U	= freestream velocity
α	= angle of attack
ω	= vorticity
Γ	= circulation
Λ	= aspect ratio of the wing
PIV	= Particle Image Velocimetry
CCD	= Charge-coupled device

1 Introduction

1.1 Axial Velocity Motion

A basic understanding concerning the occurrence of axial velocity motions at the centre of the core was first presented by Batchelor [1]. By means of the conservation equations for an axially symmetric line vortex he concluded that the axial motion is dependent on equilibrium between circulation and certain losses of head. In fact, a velocity excess or defect can be understood as a result of pressure gradients along the vortex axis and said pressure being related to centrifugal forces of the turning motion of the vortex.

Soon also the dynamic behaviour of the axial velocity was investigated. Singh et al. [2] and

Phillips et al. [3] present a high amount of axial fluctuations on a ring around the vortex core. Devenport [4] describes the mean and turbulent structure in the vortex core and in the surrounding spiralling shear layer. His measurements were conducted with hot-wire technique using a *NACA0012* half-wing model. By means of a spectral decomposition he finds axial velocity fluctuations inside the core, which are animated by the turbulence of the shear layer which wraps around the core, said fluctuations being named "inactive motion buffeting". Turbulence on her part is caused by the stretching and skewing which are caused by the vorticity in the shear layer and the rotation of the vortex itself. Outside the core he finds small-scale turbulent structures, which position themselves in four defined regions around the core. Axial velocity profiles through the centre of the core show – for increasing angle of attack – an axial velocity wake diverting from the centre of the core towards outer regions. This either indicates a ring of diminished velocity or a more asymmetric distribution of locally reduced axial velocity.

Spalart [5] soon delivers a summary of the contemporary investigations of the axial velocity and highlights the influence of viscosity inside the vortex, which leads to a diminishment of axial velocity from jets to wakes.

Amongst others, Huppertz [6] continues with the results of Devenport concerning the measured fluctuations in the core region and holds them responsible for wake-like axial velocities in the core. Having used the 3C-PIV technique, his own measurements show two separate regions of increased and two of decreased axial velocity related to the freestream value. The structure and placement of these regions are not clarified. Assumptions are made towards a velocity increasing effect through the momentum of outside flow being enclosed in between the shear layers and a decreasing effect through velocity fluctuations in the shear layer – thereby citing Devenport.

Looking at the roll-up process as such it can be concluded so far, that the matter of vortex roll-up is strongly involved with mechanisms between the core, respectively the core circulation, and the surrounding shear layer which acts via turbulence and friction onto the core and the axial velocity inside. Furthermore, some investigations indicate a ring-like distribution of axial velocity respectively separate regions outside the core.

1.2 Tip-Shape Influence

Moving forward, the impact of the wing-tip shape is now considered. Apparently, most measurements were either carried out with symmetrical *NACA*-airfoils or the *Clark-Y*-airfoil.

Using a *NACA0015* with a half-circle rounded tip side and a sharp tip side McAlister et al. [7] found a *1,45 U* excess (jet) in the case of the sharp shape and a *1,52 U* excess in the case of the rounded shape at a close $x/c = 1$ position behind the wing at $Re_c = 1{,}5 \times 10^6$. Both excesses diminish to a *1,1 U* excess at $x/c = 4$ downstream. Furthermore their measurements reveal higher amounts of circulation at the very ending of the wing in the case of the sharp tip.

Vitting [8] presents LDA-measurements having used a *Clark-Y* half-wing model. He compares a sharp and a rounded edge at *10* chords downstream and finds an excess of *1,25 U* for the sharp edge and a deficit of *0,8 U* for the rounded edge. For the rounded edge he finds a axial development from an excess of *1,4 U* at *0,6* chords to a deficit of *0,7 U* at *16* chords. By comparing a rectangular wing to a wing with double-trapezoid layout he finds a deficit of *0,7 U* for the latter which is constant for all measured distances from *0* to *2* chords. Vitting assumes that high axial excesses can only appear and prevail, if a highly concentrated vortex is formed and said vortex develops without entraining losses of axial momentum through the boundary- or shear layer respectively. The rounded tip and the double-trapezoid layout are said to lead to an involvement of shear stress reducing axial motion in the core.

Anderson et al. [9] explain jet-like fluid motions with a component of downwash

pointing in direction of the axis due to the spiralling angle of the flow. Hence the question arises if there is no backward facing component as well. Furthermore, they merge different investigations on the influence of sharp and rounded wing-tips with differing Reynolds numbers. They identify the axial velocity in streamwise direction to be linearly connected to a non-dimensional circulation parameter representing wing loading. Furthermore, they find the vortex radius to be the dominant scaling factor for said parameter. In doing so they are able to harmonise results from sharp and rounded tip sides by scaling them with the vortex radius. Comparing a sharp and a round tip side they find a *1,45 U* excess for the rounded tip and a *1,15 U* excess for the sharp tip, measuring at $\alpha = 10°$, $Re_c = 7{,}5 \times 10^5$ and $x/c = 1$. The authors suggest that the sharp tip produces a larger core due to unsmooth separation and therefore, assuming equal core circulation, the circulation density, as suggested to be the reason for high axial velocity, is comparably low.

Summing up said results it needs to be asked why the sharp edge produces higher jets than the rounded type and vice versa. Though there probably is a slight sensitivity towards Reynolds number what the rounded tip is concerned, it is not probable that a further diminishment of Reynolds number could invert both results leading to a higher jet for the sharp tip and a lower jet for the rounded tip respectively. Therefore, it is reasonable to ask, if these investigations and comparable ones (e.g. Vitting) differ because of the choice of airfoil (*NACA0012/15* and *Clark-Y*) or if also the Reynolds number influences the downstream gradient of axial velocity reduction.

Lee and Pereira [10] provide further insight into the overall axial velocity structure in planes perpendicular to the flow. They conduct measurements with a *NACA0015* half-wing model with a half-circle rounded tip side and a sharp tip side. The rounded tip vortex contains a higher core circulation and thus higher axial excess. For the sharp shape the development of the axial component along the downstream distance from *0* to *4* chords behind the wing starts at *1,35 U* and ends with *1,1 U* and for the rounded shape starts at a relatively low *1,3 U*, soon rises above the sharp value and ends at *1,2 U* ($\alpha = 10°$, $Re_c = 3{,}07 \times 10^5$).

At smaller angles of attack core circulation is also smaller and so eventually axial velocity drops below the freestream value. It is concluded that in the case of small angles of attack the shear layer diminishes the core axial velocity whereas at higher angles the shear layer prevents disturbances to intrude into the core. Tip condition is not seen to influence core condition.

1.3 Intermediate Summary

It can be concluded so far that there is a high amount of investigation concerning the influence of wing-tip shapes.

So far, most studies agree on core circulation steering axial motion at the beginning of the roll-up process. Furthermore, shear stresses may have a diminishing influence on core axial flow in the progress of the wake spiralling around the core.

Harmonisations have been presented considering that the influence of Reynolds number, wing-tip shape and downstream interaction of core and shear layer can all be scaled through the core radius. Performing this model very near downstream of the wing showed good agreements. Still, different wing-tip shapes such as rounded layouts have not been considered so far.

Since results from symmetric *NACA*-airfoils and the asymmetric *Clark-Y*-airfoil differ, it is also not clear whether different airfoil-shapes produce even opposite influences concerning the axial motion. As few data over a wider range of downstream positions exist, it may be presumed, that different downstream courses of the axial velocity may cross one another leading to different results at different stations downstream. Considering the possibility of differing gradients of the axial development in the case of the rounded tip due to Reynolds number this aspect does not at all seem trivial. So far, not enough data of the downstream development is available.

The exact velocity structure in the surrounding flow field is not fully understood yet. Conventional point-wise measurement techniques do not open the full scope to view the mechanisms. Furthermore, there are certain aspects, which indicate the occurrence of asymmetrical or at least ring-like structures instead of monolithic Gaussian distributions.

Lastly, further classifications need to be developed concerning different round shapes apart from the half-round tip side.

2 Experimental Arrangement and Procedure

The investigations presented here were conducted in an enclosed circulating water-tunnel with a measuring section of *1.000 x 540 x 540 mm³*. The tunnel's duct-contraction is *1,8*, the level of turbulence is *2-3 %* in the duct and the freestream velocity has a continuous range from *0* to *4 m/s*. All measurements were performed at freestream velocities between *2,4* and *2,7 m/s* to account for changing water temperature with respect to a constant Reynolds number. Figure 1 shows the water-tunnel.

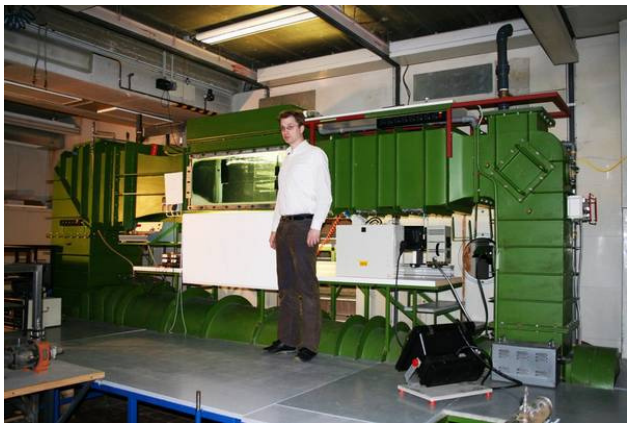


Fig. 1 The circulating water-tunnel at ILR, RWTH Aachen University

The wing was mounted vertically in the ceiling of the test-section. The position of the measuring plane and so the calibration of the 3C-PIV setup was identical for all measurements. The wing and therefore the distance of the wing's trailing edge to the measuring plane can be adjusted through removable segments which complete the measuring section's ceiling. The coordinate system and the general setup can be read from fig. 2. For ease of view the water containers avoiding optical refraction were excluded.

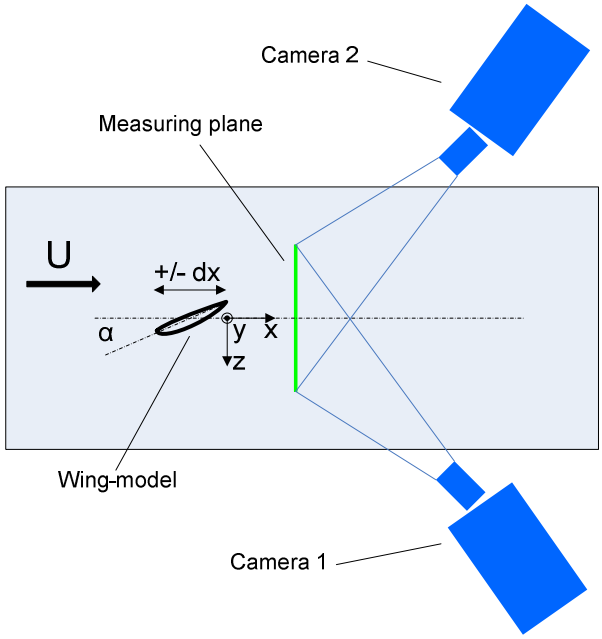


Fig. 2 General experimental setup viewed from above the measuring section

To account for differing kinematic viscosity due to rising water temperature while measuring, freestream velocity was adjusted to guarantee a constant Reynolds number.

The wing has a *Clark-Y*-airfoil and was equipped with three different wing-tips. A special naming is introduced in order to describe the geometrical features of the tips, i.e. *RS-050-050*, *RL-30-80* and *RL-20-20*. *"RS"* respectively *"RL"* stand for a **r**ounded **s**ide respectively a **r**ounded **l**ayout. In case of the *RS*-tip the first and second row of numbers indicate the rounding of the edges between the pressure-side- (respectively the suction-side-) and the tip-side-surface in percent of local airfoil thickness. This way, *RS-050-050* is the known half-circle tip side. In case of the *RL*-tips the first and second row of numbers indicate the length of the major and minor axis of an ellipsis divided by two, shaping the layout between the leading edge and the side edge. Figure 3 illustrates the aforementioned definitions.

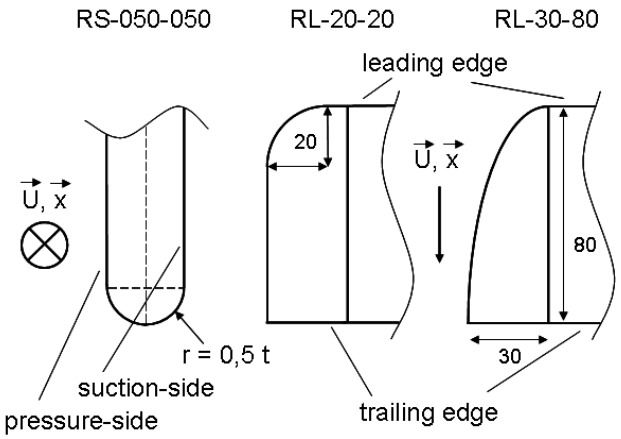


Fig. 3 Definitions for the naming of the wing-tips

The wing-tips are mounted and positioned to the wing body by magnets and pins which provide a clean surface and easy modification of the tips.

The investigations were done with the following parameters: Wing-tips: *RS-050-050*, *RL-30-80* and *RL-20-20*. The aspect ratio was fixed to $\Lambda = 5$. To account for the changing area of the tips the span was adjusted according to eq. (1), where s_{tip} has the reference length of *30 mm*. The Reynolds number was $Re_c = 200.000$, measurements were taken at the downstream positions from the trailing edge at $x/c = \{0,0\ ;\ 1,0\ ;\ 2,0\ ;\ 3,0\ ;\ 4,0\ ;\ 5,0\ ;\ 6,0\}$ with $\alpha = 10°$.

$$\Lambda = \frac{b^2}{S} = \frac{(\,2(\,s_{body} + s_{tip}\,))^2}{2((\,s_{body} s_{tip}\,) + A_{tip}\,)} \qquad (1)$$

For the stereographic PIV measurements a pulsed laser powered by *120 mJ* with an effective wavelength of *532 nm* as well as 2 CCD-cameras PCO.2000 with a resolution of *2048 x 2048 px²* mounted on adapters that fulfilled the Scheimpflug criterion were used.

Every series was recorded with *500* images at *1 Hz*. Using objectives with a focal length of *180 mm* the size of the viewing area was *60 x 40 mm²* which was divided into correlation windows of *64 x 64 px²* with an overlap of *75 %*. This way the overall resolution is approx. *6,8 vectors/mm²* and the measured minimum core with approx. $r = 3$ *mm* could be resolved by *190* vectors.

After correlating the velocity vectors in Davis 8.0.7 by LaVision they were imported into MATLAB for further computation. To account for fluctuating motions of the vortex during recording, the vortex centre of every image was calculated by the centre of mass of vorticity. Then every image was aligned to one common vortex centre before the mean vector components were calculated. This method provides velocity profiles which are not stretched and suffer from lowered peaks.

The complete investigation was done with one calibration thanks to the moveable wing in streamwise direction. The water temperature was taken before every recording series and thus the freestream velocity was conditioned. Due to high resolution the vortex fills out the complete viewing area; in order to obtain the freestream velocity, extra measurements of the freestream were performed.

To measure the axial component in the core special care needed to be taken to minimise systematic errors which may lead to a visibility of particles in off-planes in front and behind the light-sheet. Seeding density because of high resolution must be high but a secondary emission of scatter-light to particles in off-planes must be avoided. If the tunnel is not absolutely clean of air, bubbles of air will stream to the low pressure region due to the intensity of the core, line up on the vortex axis and reflect laser light to the off-planes. Therefore special care was taken to provide a focussed and thin light-sheet (*2 mm*). A small aperture of $N = 8$ secured a minimum depth of focus with a maximum light-sheet spread-angle and minimum laser power to avoid an overexposure of the particles.

3 Results

In section two the method for obtaining the vortices centre was explained. Figure 4 shows the downstream development for the centre axial velocity component u_{cd}/U for each measured configuration. The differences between the selected wing-tips can be observed.

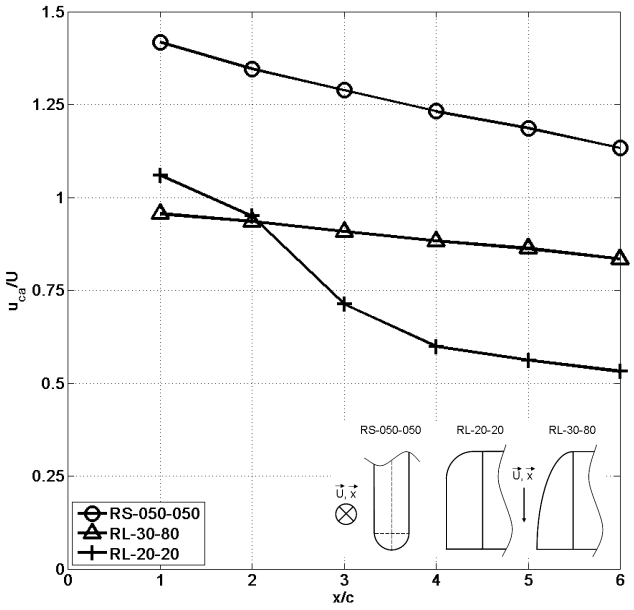


Fig. 4 Downstream development of the centre axial velocity component u_{ca}/U

All courses begin either jet- or wake-like but all diminish in axial-speed with differing slopes. The rectangular wing configuration *RS-050-050* produces high axial velocities, which start and remain jet-like. With a lower slope, u_{ca}/U of the *RL-30-80* develops being completely wake-like. The *RL-20-20* experiences a sudden decrease during the first chord wise distances.

Considering the courses in fig. 4, the development of a jet – if at all – initially occurs while the flow rolls up around the wing-tip as there is no reason for the flow to accelerate far in front of the wing and the flow behind the wing experiences shown velocity drop. Since the core is formed at the wing-tip, therefore the highest initial axial flow is produced by the vortices core structure at the wing.

Ultimately, axial motion along the vortex axis can only arise through some pressure gradient along said axis – apart from head losses which are likely to reduce axial motion even more. A negative gradient and thus falling pressure will produce an acceleration compared to freestream velocity and likewise a positive gradient and thus a rising pressure will lead to a diminishment of axial centre flow. Considering the courses in fig. 4, the flow in the core experiences differing courses of counter pressure.

Through a vortices' turning motion the pressure at the axis diminishes in order to balance the centrifugal force. By setting up equilibrium of force experienced by an infinitesimal piece of fluid the pressure drop from far outside the vortex to its centre can be expressed by the radial distribution of circulation, eq. 2.

$$\frac{p_{ca} - p_{\infty}}{\rho} = \frac{1}{4\pi^2} \int_{r=0}^{\infty} \frac{\left(\frac{\Gamma(r)}{Ub}\right)^2}{\left(\frac{r}{b}\right)^3} dr \qquad (2)$$

Here, $\Gamma/(Ub)$ and r/b are normalised values of circulation and radius. Obviously, the key factor causing axial motion is the sum of radial fragments of the normalised form of $\Gamma(r)^2/r^3$ representing the intensity of circulation towards the centre of the vortex. Furthermore, by the nature of said factor, the portions of circulation closest to and inside the core as the radius approaches zero are emphasised most. Figure 5 displays the radial distribution of circulation in the core region (for all tips $r/s < 0{,}02$) and the nearest region outside the core at the first plane at $x/c = 1$. The curved slope in the case of the *RS-050-050* between $r/s = 0{,}02$ and $0{,}05$ represents the addition of vorticity by the shear layer during its first spiral around the core.

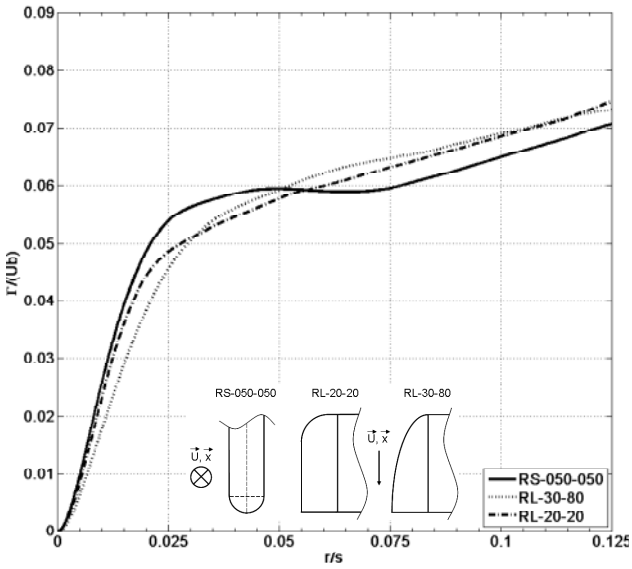


Fig. 5 Radial distribution of the circulation coefficient $\Gamma/(Ub)$ at plane $x/c = 1$

Since the state of circulation approaching r = 0 is crucial, the core and the circulation gathered there need to be considered in detail.

The following figures give insight into the behaviour of circulation intensity in the core.

Figure 6 displays the mentioned intensity of circulation in the core.

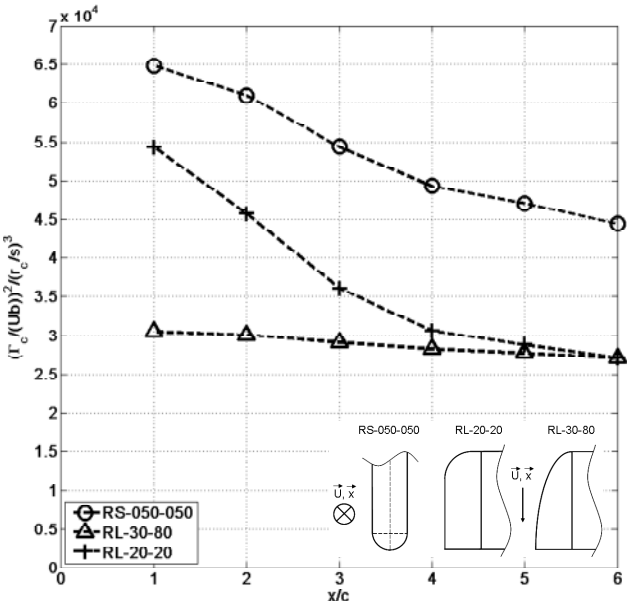


Fig. 6 Downstream development of the intensity of circulation of normalised Γ_c^2/r_c^3 inside the core

Furthermore, core circulation is preserved downstream, as seen in fig. 7.

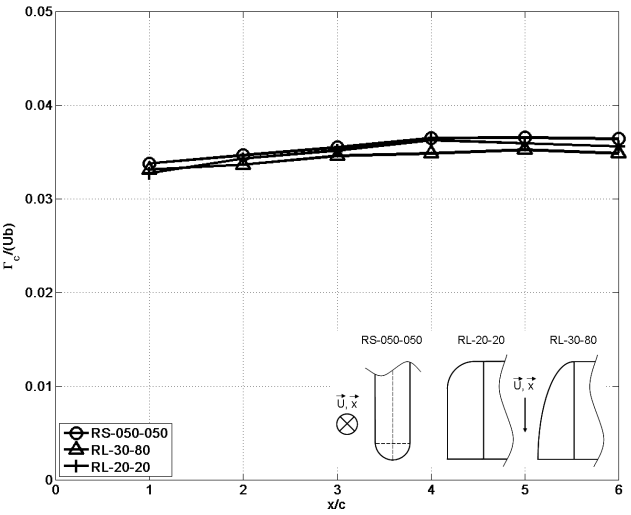


Fig. 7 Downstream development of the core circulation coefficient $\Gamma_c/(Ub)$

In the course of viscosity in the region of the highest tangential gradients, i.e. in the ring-like region at the boundary of the core, peak tangential velocity reduces which leads to a widening of the core, as can be seen in fig. 8 and fig. 9 respectively.

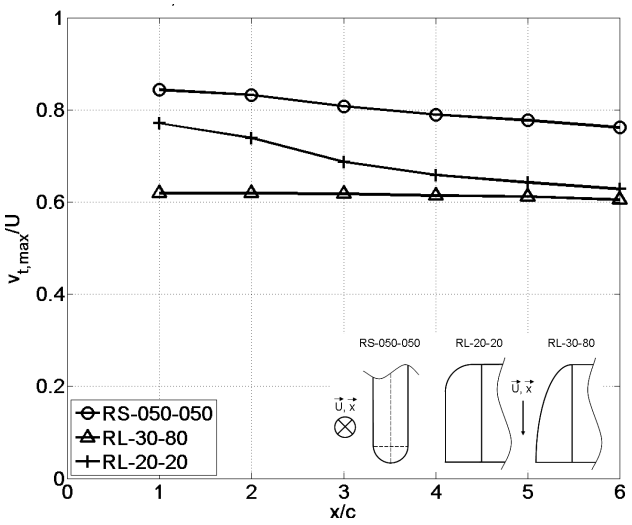


Fig. 8 Downstream development of the peak tangential velocity component $v_{t,max}/U$

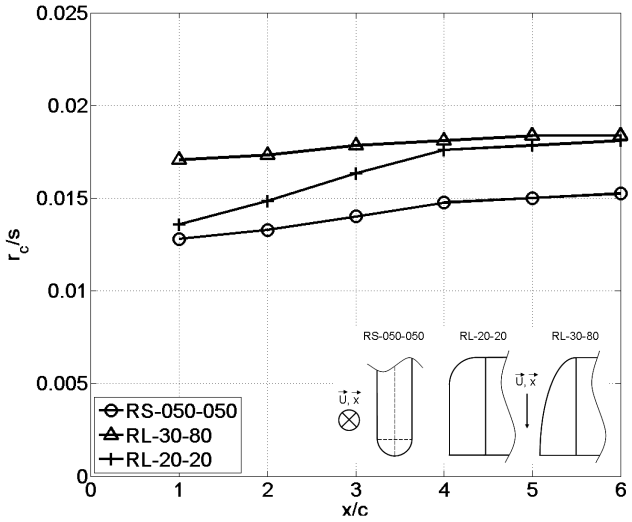


Fig. 9 Downstream development of the core radius r_c/s

Since the level of circulation remains constant and the core widens, the intensity of circulation diminishes shifting circulation to greater radii. Eventually, this leads to a pressure drop in the centre of the vortex.

It needs to be questioned now, from which part of the vortex during roll-up the influences on viscosity originate. Figure 10 shows the downstream development of the tangential velocity profile in the case of the *RL-20-20* tip, whose vortex experiences the greatest fall of axial motion. It can be seen that the development of the outer part of the tangential profile is independent from the widening of the vortex core in the course of viscous diffusion.

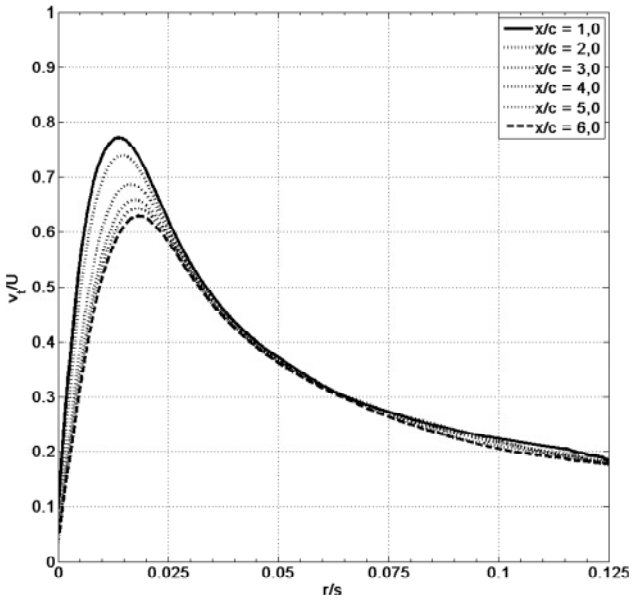


Fig. 10 Downstream development of the tangential velocity distribution $v_t(r)/U$ of *RL-20-20*

In addition, fig. 11 illustrates the radial distribution of circulation in the last plane *x/c = 6*. Comparing the courses to those in fig. 5 (radial circulation distribution at *x/c = 1*) it can be seen, that circulation outside the first spiral of the shear layer (*r/s > 0,05*) continues to be filled up due to further layers of viscosity being added onto the core. But it is not likely, that this addition has impact on the core region. Therefore, the influence due to viscosity is a function of time and not of place. This leads to conclude, that the initial conditions at *x/c = 1* inside the core need to be focussed on.

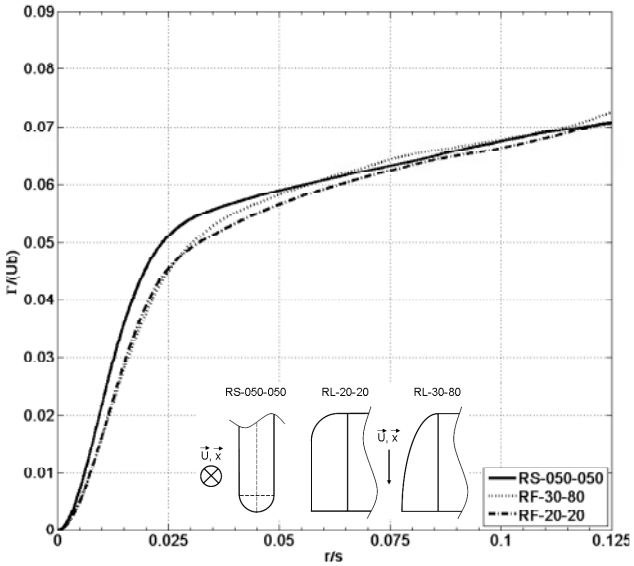


Fig. 11 Radial distribution of the circulation coefficient $\Gamma/(Ub)$ at plane *x/c = 6*

Focussing on the core region, the question what causes different states of friction needs to be answered. Friction occurs while there are high gradients in the tangential distribution which means that the radial development of circulation has a sharper bending outside of the core circulation region. A more continuous and smooth bending is achieved by adding circulation which is provided by vorticity coming from the layer.

Since it is known that vorticity in said layer increases along the span towards the wing-tip, it is a question were the core is positioned to receive this portion of vorticity at the very outside of the core. Figure 12 shows the positions of the cores of the different configurations.

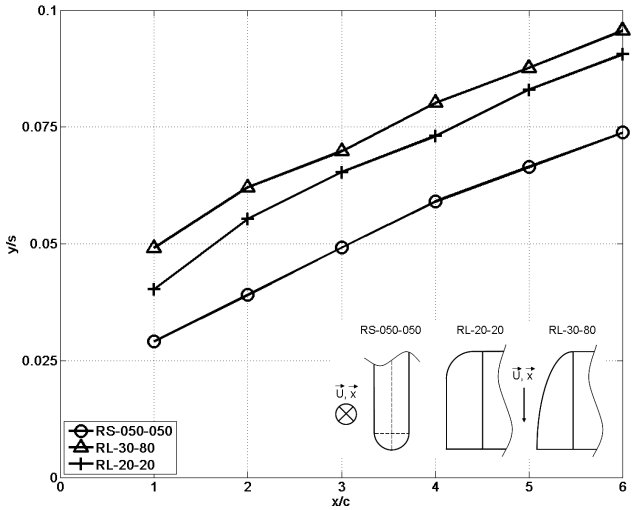


Fig. 12 Downstream development of the position of the vortex axis in the x,y-plane

It can be seen that the *RS-050-050*, shaping a rectangular wing, forms a vortex at the very tip of the wing. On the other hand, *RL-20-20* forms a vortex axis which is shifted towards the wing's root and *RL-30-80* even more.

Considering the vortex axis positions in fig. 12 and the core radii in fig. 9, the initial distribution of circulation in fig. 5 can be explained.

The *RS-050-050* forms at the tip of the wing where the axis is lead along the tip side. The core develops smoothly around the rounding and therefore creates a small and intensive core due to the maximum intensity of vorticity at the very tip of the wing. Core circulation and core

radius form the initial gradient of circulation, which is the largest in the case of the *RS-050-050,* as can be seen in fig. 6. Having formed the core behind the wing the outer portion of the core experiences a great amount of vorticity brought to it by the shear layer, therefore explaining the continuous and smooth increase of circulation outside of the cores, as seen in fig. 5.

The *RL-20-20* and *RL-30-80* form their vortex cores on the suction side of the wing at the wing-tip. This can be seen in fig. 12 by their axis positions, which are both shifted towards the root of the wing. While rolling on the surface of the wing, boundary layer is involved into core production leading to a small circulation gradient inside the core. Since the cores evolve at a more inside position of the wing, there is no great circulation overshoot outside the core brought by the vorticity layer – as experienced by the rectangular wing configuration.

The question is now, what mechanism leads to the vast diminishment of axial velocity in the case of the *RL-20-20* in contrast to the *RL-30-80* since both seem to have a similar roll-up process. The question can again be answered by a higher amount of viscosity at the core edge region for the *RL-20-20* due to higher tangential gradients. A hypothesis can be formulated:

Figure 13 shows the distribution of vorticity in a plane directly at the leading edge (*x/c = 0*) of the wing with the *RL-30-80* tip. It can be seen, that, unlike any other wing-tip, the *RL-30-80* forms two vortices during roll-up over the wing's surface.

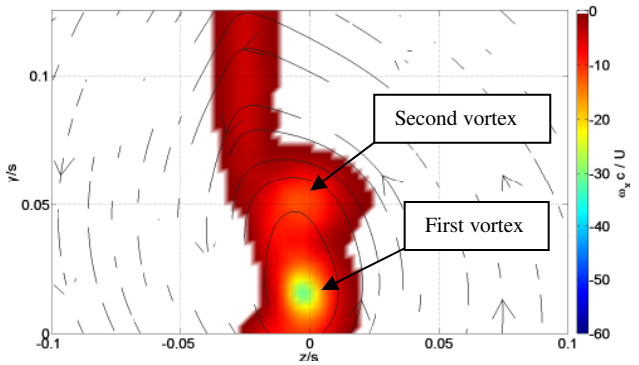


Fig. 13 Contour of normalised vorticity $\omega_x c/U$ for *RL-30-80* at *x/c = 0*. For clarity, vorticity below *1 %* of the maximum value is not displayed

Said vortices merge into one vortex as soon as the first plane *x/c = 1* is reached. The first and major vortex is the tip vortex (*y/s = 0,02*), the second vortex is similar to that separating from the edge of a slender delta wing (*y/s = 0,05*). Since the shape of the *RL-30-80* has a slender portion at the front this seems to be a valid explanation. Also, the vortex core is widened because of both vortices gradually forming one core. Considering the starting point of the second vortex, which must be at the curved leading edge of the tip, and the position of the vortex axis from fig. 12, it can be assumed that the vortex axis is parallel to chord direction. Having developed further towards the trailing edge, the core is contacted by the vorticity layer and thus the curved slope of the radial distribution of circulation can be explained.

In contrast to this curved slope, the *RL-20-20* shows a more edgy transition from core circulation to outer core circulation. This edgy course is caused for the same reason as the tangential velocity profile having a sharp maximum, see fig. 10, namely the non-existence of axial vorticity components added to the flow outside the core.

A second hypothesis can be formulated: The rounding of the *RL-20-20* attaches the core flow to the surface leading the core axis on the suction surface of the wing and then diagonally towards the root of the wing and not parallel to chord as in the case of the *RL-30-80*. This way, the core axis is placed parallel to the streamlines on the suction side which are also directed towards the root of the wing. On the path to the leading edge, boundary layer material is entrained into the core. As soon as the leading edge is reached, the core does not experience axial vorticity components coming from the vorticity layer but more a tangential velocity component from the flow of the pressure side of the wing, which is directed towards the tip of the wing. Therefore, no major vorticity component is realised in the region outside of the core and the radial distribution of circulation develops a sudden edge. The edge i.e. the sharp peak of tangential velocity forms high shear stresses which diminish peak tangential velocity

severely while the vortex is pacing downstream. This high reduction of tangential velocity, as can be seen in fig. 10, leads to a severe drop in negative pressure at the core and so axial core velocity is reduced as can be seen.

4 Conclusions

Experiments have been performed and results presented in order to assign vortex structure and downstream behaviour to geometric parameters of the wing-tip shape.

It was found that the rounding of the wing's layout has a different impact than the rounding of the sides of the tip. A first step towards a classification was shown and a common establishment of said classification is encouraged.

In general, the wrapping of vorticity around the core leads to a curved and smooth transition of the radial distribution of circulation from the core region to the outside of the core and thus leads to a wider peak slope of the tangential velocity profile.

This wider slope leads to a lower amount of shear stresses in the region of the core boundary where the highest tangential gradients can be observed. Therefore a sharp velocity profile leads to a high amount of viscous stress. Said stress leads to a diminishment of peak tangential velocity and because of conservation of core circulation the core is widened. A widening of the core and a diminished tangential velocity diminish the radial gradient of circulation in the very near of the vortex axis, said gradient being the key factor for negative pressure on said axis.

Therefore, on the one hand, a high initial radial circulation gradient and a region of high vorticity outside the core leads to high axial jets which are preserved over a long pace downstream. This is obviously the case for wing-tips shaping a rectangular wing, where the rounding of the tip's side delivers a high core circulation gradient.

Respectively, on the other hand, a low initial radial circulation gradient and a region of low vorticity outside the core lead to a lower axial velocity, which diminishes quickly downstream.

References

[1] Batchelor G. K., "Axial Flow in Trailing Line Vortices", *J. Fluid Mech.*, Vol. 20, Part 4, 1964, pp. 645-658.

[2] Singh P. I., Uberoi M. S., "Experiments on Vortex Stability", *The Physics of Fluids*, Vol. 19, No. 12, 1976, pp. 1858-1863.

[3] Phillips W. R. C., Graham J. A. H., "Reynolds-Stress Measurements in a Turbulent Trailing Vortex, *J. Fluid Mech.*, Vol. 147, 1984, pp. 352-371.

[4] Devenport W. J., Rife M. C., Liapis S. I., Follin G. J., "The Structure and Development of a Wing-Tip Vortex", *J. Fluid Mech.*, Vol. 312, 1996, pp. 67-106.

[5] Spalart P. R. "Airplane Trailing Vortices", *Ann. Rev. Fluid Mech.*, Vol. 30, 1998, pp. 107-138.

[6] Huppertz G. M., "Experimental Investigation of Engine Jet / Vortex Interaction", Ph.D. Dissertation, Institute of Aerodynamics, RWTH Aachen University, Germany, 2010.

[7] McAlister K. W., Takahashi R. K., "NACA 0015 Wing Pressure and Trailing Vortex Measurements", NASA Technical Paper 3151, AVSCOM Technical Report 91-A-003, 1991.

[8] Vitting T., "Struktur von Flügelrandwirbeln und Maßnahmen zur Wirbelabschwächung", Ph.D. Dissertation, Institute of Aeronautics and Astronautics, RWTH Aachen University, 1991.

[9] Anderson E. A., Lawton T. A., "Correlation between Vortex Strength and Axial Velocity in a Trailing Vortex", *J. Aircraft*, Vol. 40, No. 4, 2003, pp. 699-704.

[10] Lee T., Pereira J., "Nature of Wakelike and Jetlike Axial Tip Vortex Flows", *J. Fluid Mech.*, Vol. 47, No. 6, 2010, pp. 1946-1954.

CEAS 2011

CFD Applications 1

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

CFD analysis of wind tunnel flow in open section model rotor tests

M. Biava
AgustaWestland-Politecnico Advanced Rotorcraft Center, Milano, Italy,
e-mail: biava.awparc@agustawestland.com

L. Vigevano
Dipartimento di Ingegneria Aerospaziale - Politecnico di Milano, Milano, Italy,
e-mail: luigi.vigevano@polimi.it

Keywords: CFD, wind tunnel, rotor, flow breakdown

Abstract

The paper presents a qualitative CFD analysis of the flow inside the open test section of the Politecnico di Milano large wind tunnel, with account for rotor effects by means of an actuator disk model. The analysis of the performed steady state simulations allows to determinate approximatively the flow breakdown boundaries in terms of wind tunnel speed and rotor thrust. Indications are also given on how to detect the proximity of flow breakdown conditions by simple pressure measurements during wind tunnel operation.

1 Introduction

Despite the increasing capability of CFD methods in predicting rotor performance, the use of wind tunnel test measurements still plays an important role in the design of new rotor systems. However, it is not obvious how to operate the model rotor and the wind tunnel so as to closely reproduce free air conditions. In fact, rotor tests at low speed and high thrust in a wind tunnel may experience what is known as *flow breakdown*, that happens when the interaction between the rotor wake and the tunnel walls strongly modifies the flow in the vicinity of the rotor due to the formation of recirculating flow regions [1]. In flow breakdown condition the wind tunnel environment is no more representative of the free air environment and the rotor performance cannot be adjusted by means of wall corrections. Therefore, such operating conditions are to be avoided in a test campaign.

Shinoda [2] addressed the study of wall interference effects and the identification of flow breakdown by means of an extensive experimental campaign in the 80×120 foot wind tunnel at NASA Ames. Nonetheless, to the authors knowledge there is no available analysis of the flow breakdown problem for a wind tunnel with open test section. The present work addresses this subject with purely numerical techniques, without the objective of generality, but restricting the attention to the 4×3.8 m open test section of the Politecnico di Milano (PoliMi) large wind tunnel. To this end, a campaign of numerical simulations has been carried out to characterize the flow field in the wind tunnel generated by the AgustaWestland AW139 model rotor. The computations were performed with the CFD code *ROSITA* (ROtorcraft Software ITAly) based on the solution of the Reynolds Averaged Navier-Stokes equations coupled with

the one-equation turbulence model of Spalart-Allmaras. The rotor effect is represented with an actuator disk model. The numerical solutions were obtained for several values of the rotor thrust and of the advance ratio and visualizations of the computed flow field were used to detect the onset of flow breakdown. The analysis permitted to define the flow breakdown boundaries for the considered wind tunnel and also provided some hints about the experimental indicators that may be utilized to recognize the flow breakdown condition during model rotor tests.

The structure of this paper is as follows. Section 2 summarizes the main characteristic of the CFD solver. Section 3 describes the numerical model and the considered flow conditions, while numerical results are discussed in section 4. Some conclusions are drawn in the last section.

2 The flow solver *ROSITA*

The *ROSITA* flow solver [3] numerically integrates the RANS equations, coupled with the one-equation turbulence model of Spalart–Allmaras [4], in systems of moving, overset, multi-block grids. The equations are discretized in space by means of a cell-centred finite-volume implementation of the Roe's scheme [5]. Second order accuracy is obtained through the use of MUSCL extrapolation supplemented with a modified version of the Van Albada limiter introduced by Venkatakrishnan [6]. The viscous terms are computed by the application of the Gauss theorem and using a cell-centred discretization scheme. Time advancement is carried out with a dual-time formulation [7], employing a 2nd order backward differentiation formula to approximate the time derivative and a fully unfactored implicit scheme in pseudo-time. The generalized conjugate gradient (GCG), in conjunction with a block incomplete lower-upper preconditioner, is used to solve the resulting linear system.

The connectivity between the (possibly moving) component grids is computed by means of the Chimera technique. The approach adopted in *ROSITA* is derived from that originally proposed by Chesshire and Henshaw [8], with modifications to further improve robustness and performance. The domain boundaries with solid wall conditions are firstly identified and all points in overlapping grids that fall close to these boundaries are marked as holes (seed points). Then, an iterative algorithm identifies the donor and fringe points and lets the hole points grow from the seeds until they fill entirely the regions outside the computational domain.

The *ROSITA* solver is fully capable of running in parallel on computing clusters. The parallel algorithm is based on the message passing programming paradigm and the parallelization strategy consists in distributing the grid blocks among the available processors.

3 Numerical model

Figure 1 shows the numerical domain used for the simulations of the open test section of the PoliMi large wind tunnel. The Chimera grid system consists of the following components.

- A background mesh which represents the chamber containing part of the wind tunnel circuit and the open test section.
- Four grids representing the flow deflectors placed at the beginning of the wind tunnel return circuit.
- A cylindrical mesh for the actuator disk.

The figure 2 reports a slice of the computational mesh in the symmetry plane of the wind tunnel, where the different component grids can be clearly viewed. A magnified view of the mesh close to the upper deflector is also shown in the same figure and it let appreciate better the cell distrubution in the region of Chimera interpolation between the deflector grid and the wind tunnel grid. It can also to be noted that the all the solid walls (deflectors and wind tunnel walls) have non-zero thickness. In total the mesh counts about 13 million cells. The applied boundary conditions are:

- viscous wall boundary conditions on the wind tunnel walls;

- inviscid wall boundary conditions on the outer chamber walls;
- velocity inlet boundary conditions at the inflow section;
- pressure outlet boundary conditions at the outflow section.

The actuator disk is instead seen as a distribution of linear momentum sources over a disk-shaped grid plane of the cylindrical grid. The actuator disk has been positioned parallel to the wind tunnel floor since the pressure distributions on the disk where extracted from full 3D simulations of the rotor performed with the mast angle equal to zero.

The considered combinations of wind tunnel flow speed and rotor thrust coefficient cover the range $10 < V_\infty < 35$, $0.08 < C_T/\sigma < 0.12$. The steady simulations have been performed with the RANS solver using the Spalart-Allmaras turbulence model. The solver was run in parallel on 64 processors so that 26 hours were needed to complete the 5000 iterations at CFL = 2.5 scheluded for each test case.

4 Analysis of results

All the selected test conditions were simulated with the assumption of a steady flow, despite the fact that unsteady phenomena are intuitively to be expected for the flow under investigation. Nevertheless, the steady solutions are retained here as a good reference to understand the overall qualitative pattern of the flow (note that an unsteady simulation of the flow would cost roughly ten times more in terms of computational time). This assumption is reasonable since we are mainly interested in the trajectory of the rotor wake before its possible impingement on the tunnel walls, where the flow is likely to be locally close to a steady condition. Unsteady flow regions are instead probably present after the wake impact on the tunnel walls or on the chamber floor and within the low-speed recirculation zones in the plenum, but these last mentioned regions are of less interest to us.

Figure 3 shows the convergence of the force coefficients on the lower flow deflector for two opposite operating conditions, that is, respectively, low speed with high disk loading and high speed with low disk loading. It is evident that in both cases the solution is not stationary. In particular, at the lower wind tunnel speed, the behaviour of the force coefficients is quite irregular and this is reasonably a symptom of a strong interaction between the rotor wake and the wind tunnel walls and flow deflectors. At the higher speed, the variation of the force is more regular (nearly periodic): in this regime the wake/walls interaction is weaker, but the flow is not yet sufficiently stable to reach a steady state.

The above considerations become more clear if we look at figure 4, where the streamtraces of the velocity field are plotted in the region surrounding the open test section of the wind tunnel for the same two operating conditions. At the lower wind tunnel speed and higher disk loading the rotor wake goes downward, it partly interacts with the lower flow deflector and then it is convected in the chamber right below the return circuit of the wind tunnel, promoting the onset of regions of unsteady, low speed, recirculating flow. At the higher speed and lower disk loading the wake is instead completely ingested in the return circuit.

The trajectory of the wake for the two operating conditions under analysis is better visualized in figures 5(a)–5(b), showing the normalized Mach number distribution in the vertical simmetry plane of the open test section. The same figure permits to appreciate the effect of the rotor on the wind tunnel nozzle jet. At low wind tunnel speeds the jet is significatively bent downward and, consequently, part of the jet flows outside the wind tunnel circuit and, at the same time, there is a suction of low speed fluid in the upper part of the inlet section of the circuit. It is important to note, however, that the low speed fluid does not interact with the rotor disk, at least in the considered range of wind tunnel speeds. Figure 6 shows the flow in the proximity of the upper deflector, the region where the low speed fluid enters the return circuit of the wind tunnel.

In figures 5(c)–5(d) two other solutions are shown, corresponding to operating conditions laying between the two extreme cases considered

so far.

In some of the previously analyzed figures it is possible to note slight discontinuities of the Mach number within the wake of the rotor. They are not actual discontinuities of the flow field but instead they are caused by the different cell size across the Chimera interface between the actuator disk grid and the wind tunnel grid.

By the analysis of the presented numerical solutions it seems that, for all the considered speeds and rotor thrusts, the flow in the neighborhood of the rotor disk is qualitatively similar to that of the free-flight flow. Indeed, there is no evident interaction of the rotor with its own wake or with the shear layer of the nozzle jet, even for the lowest velocities and the highest disk loading. The only noticeable effect, observed for an intermediate range of velocities and thrust coefficients, is the impact of the wake on the lower deflector, with a possibly consequent ground effect. It is however reasonable to think that the performance measurements gathered in the open test section of the PoliMi wind tunnel might be related to the performance of the real-size rotor in free-flight by a fairly general correction procedure.

Nevertheless, there can be other effects that could potentially alter the wind tunnel measurements and need to be investigated. One example of potentially adverse phenomenon is the following: it has been shown that there are operating conditions where part of the wake is ingested by the wind tunnel and part goes outside. In this situation a small perturbation on the wake trajectory alters the energy content of the flow entering the return circuit and thus modifies the dynamics of the wind tunnel circuit. That leads to a perturbation on the jet leaving the nozzle and flowing around the rotor and, in turn, to a modification of the wake trajectory. It is to be verified if the perturbation of the wake trajectory gets damped or if it leads to a limit cycle for some values of the operating parameters.

The presented numerical results suggest that this unsteady phenomenon *could* appear for $V_\infty < 25\,\mathrm{m/s}$, but cannot help in predicting if it will be effectively observed or not. It can however provide a possible explanation to the nearly periodic perturbations observed during past rotor tests performed at the PoliMi wind tunnel with open test section configuration. A careful monitoring of the wind tunnel state is therefore suggested when operating within the identified critical range.

In order to detect the proximity of such conditions during wind tunnel operation, it is desirable to devise a (simple) means to detect the position of the rotor wake. Pressure probes are among the simplest sensors that can be installed in the wind tunnel and thus we want to rely on pressure measurements in order to estimate the trajectory of the rotor wake. In figures 7 and 8 the pressure coefficient distribution on the wind tunnel walls is plotted beside the distribution of the Mach number in the test section vertical symmetry plane for $V_\infty = 10$–$15\,\mathrm{m/s}$ and for various rotor thrusts. Velocities higher than $25\,\mathrm{m/s}$ are not considered here because at those regimes the wake is completely ingested by the tunnel inlet for all the values of the thrust coefficient.

The analysis of the two figures reveals that when the wake is impacting on the lower deflector one or more pressure peaks can be observed on the deflector surface. The distribution of the pressure in terms of the location where the wake impact on the deflector can be summarized as follows.

- For $V_\infty = 10\,\mathrm{m/s}$ and $C_T/\sigma = 0.12$ the wake escapes almost completely the wind tunnel inlet but there is no reingestion of the wake by the rotor (see figure 7(c)). The effect of the wake impingement on the lower deflector are two symmetric pressure peaks on the leading edge of the deflector, reasonably due to the two main vortices generated by the roll-up of the wake. For lower velocities and higher thrusts it is expected that the wake does not interact any more with the deflector and it is entirely pushed down into the space between the wind tunnel lower wall and the chamber floor.
- For the operating conditions between ($V_\infty = 10\,\mathrm{m/s}$, $C_T/\sigma = 0.10$) and ($V_\infty = 15\,\mathrm{m/s}$, $C_T/\sigma = 0.08$), the two pressure peaks coalesce into a single peak located roughly in the middle (in the y direction) of

the deflector surface. The peak moves from the leading edge to the trailing edge of the deflector as the velocity is increased or as the thrust is decreased. In these operating conditions a ground effect due to the interaction between the wake and the deflector could arise.

- For $V_\infty > 25\,\mathrm{m/s}$ the wake enters completely the return circuit of the tunnel, and only a slight pressure raise distributed over the leading edge of the lower deflector may be noticed.

From the above description stems that it is possible to monitor the rotor wake position by measuring the pressure at some selected locations on the lower deflector. For instance, some rows of pressure sensors could be installed close to its leading edge and along its span in the x direction. This solution is quite cheap and not intrusive (the pressure sensors can be glued on the deflector surface without the need to drill the tunnel walls).

5 Conclusions

The paper described the qualitative analysis of the flow inside the open test section of the Politecnico di Milano large wind tunnel in the presence of rotor effects. The flow has been simulated with the flow solver *ROSITA*, adopting an actuator disk model of the AW139 rotor to account for the rotor effects in the numerical solution.

The computed flow fields in the open test section appeared to be qualitatively similar to the flow field of the rotor in free-flight conditions for all the operating conditions of interest. Nonetheless, quantitative effects on the measured performance are to be expected, due to the interaction of the rotor wake with the wind tunnel walls, especially for an intermediate range of flow speed and rotor thrust values where the wake impinges on the lower deflector.

In addition, a possibly adverse effect on the performance measurements, due to an unsteady interaction between the wake and the wind tunnel walls, has been described and the operating parameters range where the phenomenon could occour have been identified. A pratical means to detect these critical conditions during the actual wind tunnel operations has been given, based on pressure measurement at some selected locations.

References

[1] J.-J. Philippe. Consideration on wind-tunnel testing techniques for rotorcrafts. In *AGARD Special Course on Aerodynamics of Rotorcraft*, 1990.

[2] P. M. Shinoda. Wall interaction effects for a full-scale helicopter rotor in the NASA Ames 80- by 120-foot wind tunnel. In *AGARD Meeting on "Wall Intereference, Support Interference and Flow Field Measurements"*, 1993.

[3] M. Biava, A. Pisoni, A. Saporiti, and L. Vigevano. Efficient rotor aerodynamics predictions with an Euler method. In *29th European Rotorcraft Forum, Friedrichshafen, Germany*, 2003.

[4] P.R. Spalart and S.R. Allmaras. One equation model for aerodynamic flows. In *30th AIAA Aerospace Science Meeting & Exhibit, Reno, NV. AIAA 92-0439*, 1992.

[5] P.L. Roe. Approximate riemann solvers, parameter vectors and difference schemes. *J.Comput. Phys*, 43:357–372, 1981.

[6] V. Venkatakrishnan. On the accuracy of limiters and convergence to steady state solutions. 31st *AIAA Aerospace Sciences Meeting and Exhibit*, AIAA 93-0880, 1993. Reno, Nevada.

[7] A. Jameson. Time dependent calculations using multigrid with applications to unsteady flows past airfoils and wings. In *10th AIAA Computational Fluid Dynamics Conference, Honolulu, HI. AIAA 91-1596*, 1991.

[8] G. Chesshire and W. D. Henshaw. Composite overlapping meshes for the solution of partial differential equations. *J. Comp. Phys.*, 90:1–64, 1990.

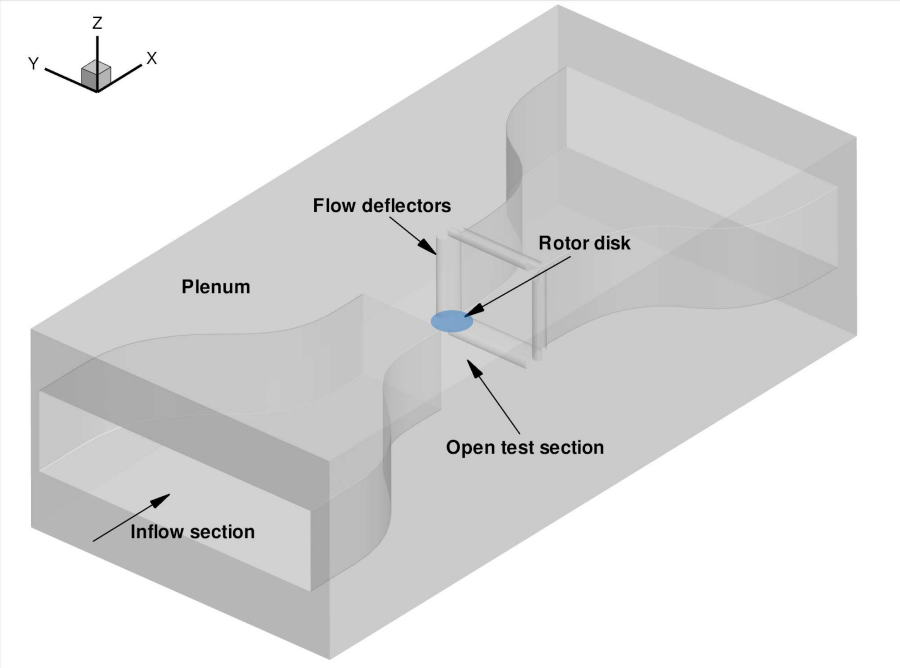
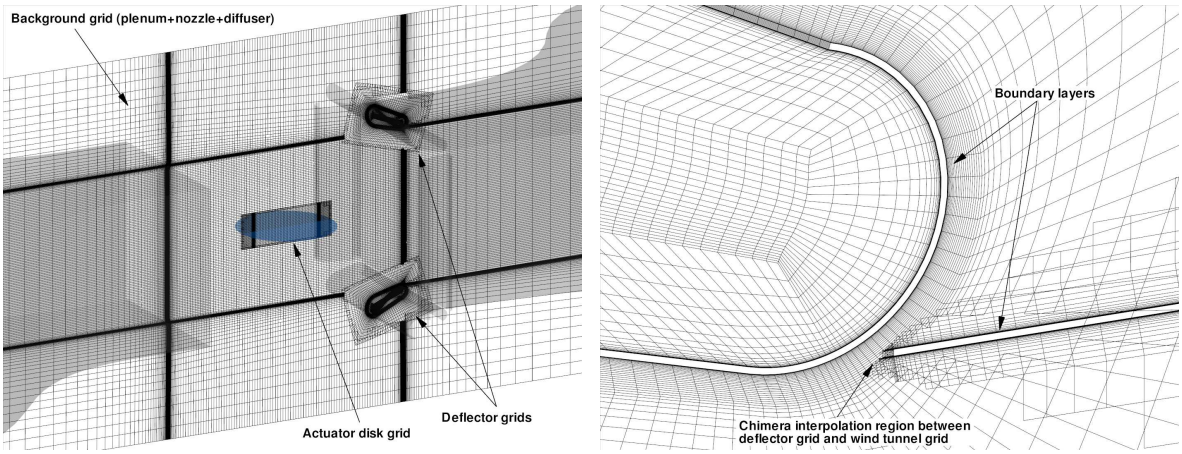
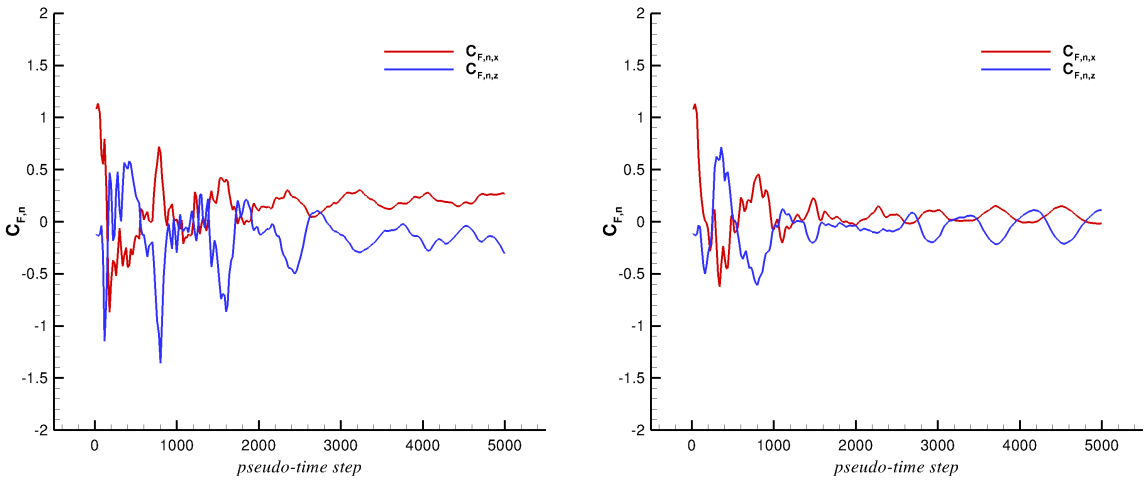


Figure 1: Numerical domain for the open test section simulations.



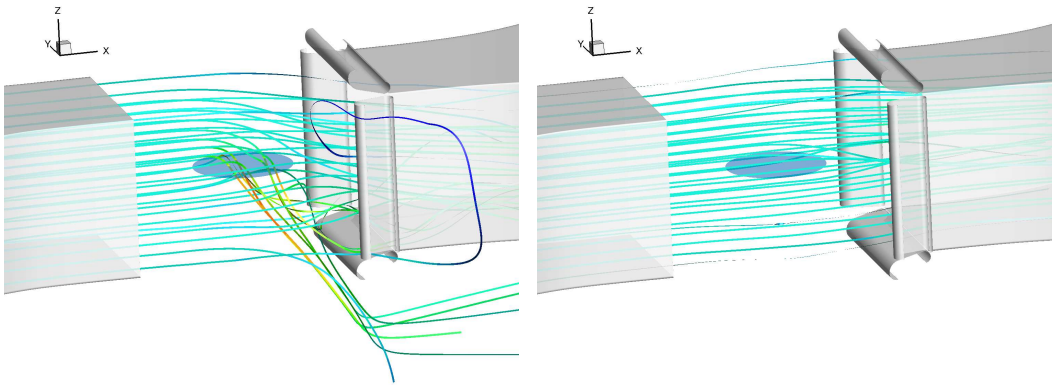
(a) (b)

Figure 2: A slice of the computational grid in the symmetry plane of the wind tunnel (a) and a zoomed view of the slice close to the upper deflector (b).



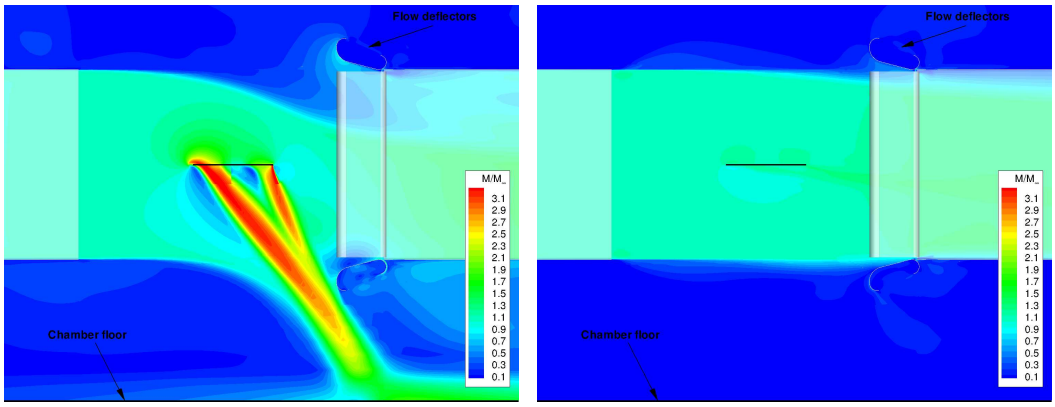
(a) $V_\infty = 10\,\mathrm{m/s}$ – $C_T/\sigma = 0.12$ (b) $V_\infty = 35\,\mathrm{m/s}$ – $C_T/\sigma = 0.08$

Figure 3: Force coefficients on the lower flow deflector.



(a) $V_\infty = 10\,\mathrm{m/s}$ – $C_T/\sigma = 0.12$ (b) $V_\infty = 35\,\mathrm{m/s}$ – $C_T/\sigma = 0.08$

Figure 4: Streamtraces of the velocity field.



(a) $V_\infty = 10\,\mathrm{m/s}$ – $C_T/\sigma = 0.12$ (b) $V_\infty = 35\,\mathrm{m/s}$ – $C_T/\sigma = 0.08$

(c) $V_\infty = 15\,\mathrm{m/s}$ – $C_T/\sigma = 0.10$ (d) $V_\infty = 25\,\mathrm{m/s}$ – $C_T/\sigma = 0.12$

Figure 5: Mach number distribution in the vertical simmetry plane of the open test section (normalized with respect to the Mach number at the exit of the wind tunnel nozzle).

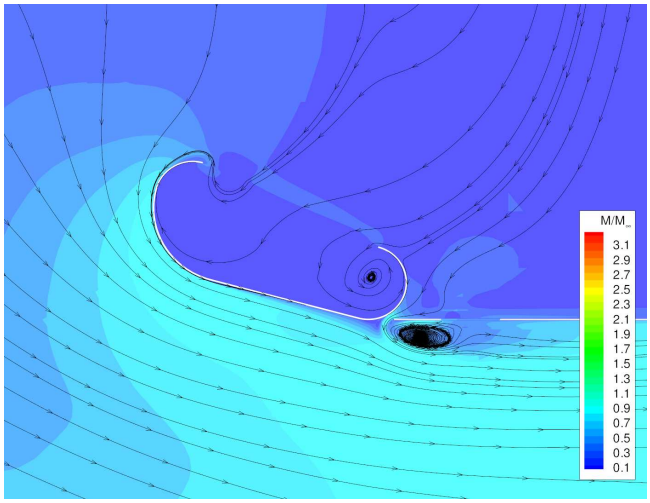
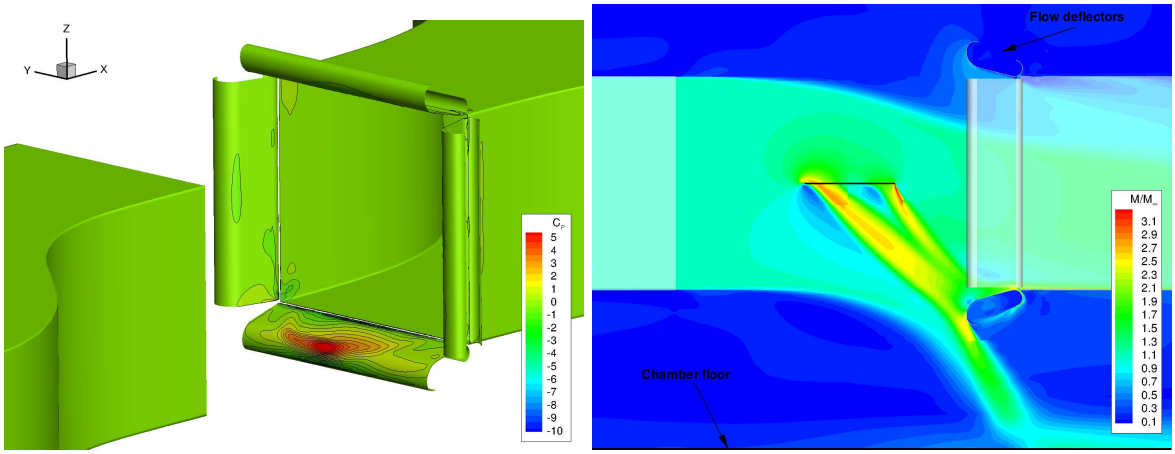
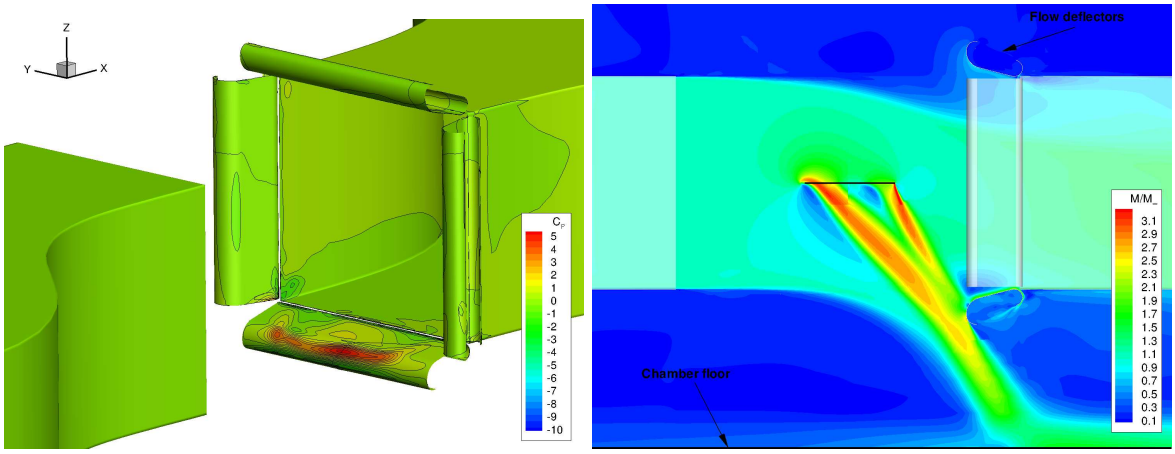


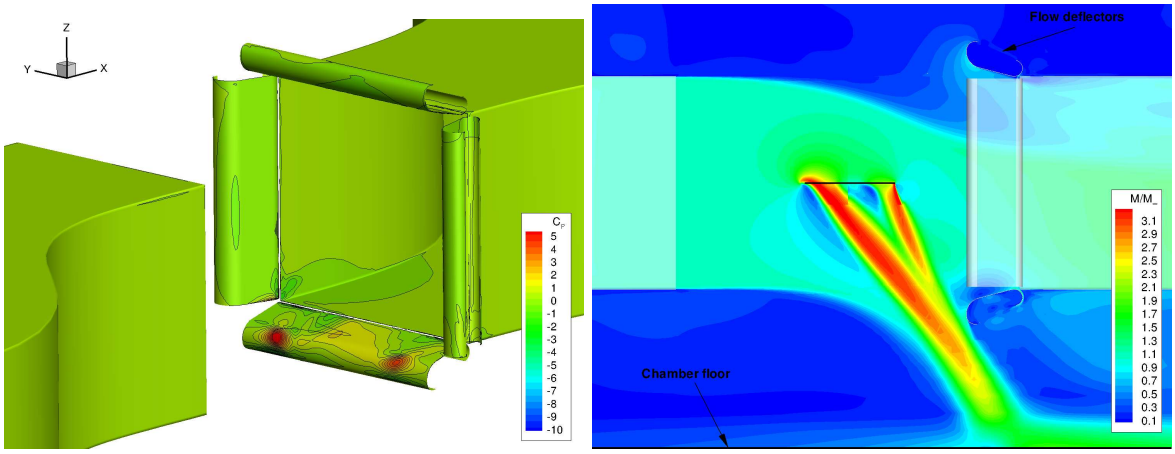
Figure 6: Mach number distribution in the vertical simmetry plane of the open test section (normalized with respect to the Mach number at the exit of the wind tunnel nozzle): zoom on the region surrounding the upper deflector.



(a) $C_T/\sigma = 0.08$

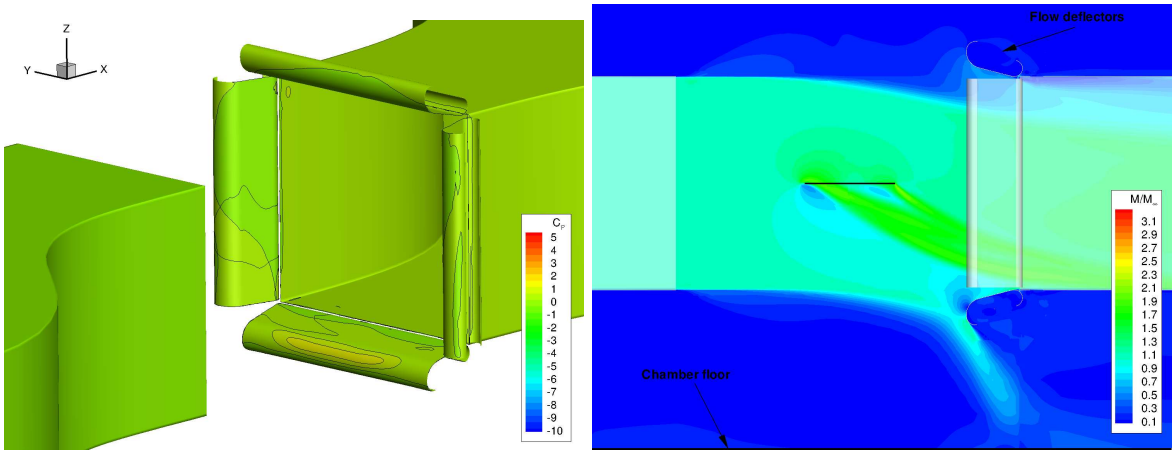


(b) $C_T/\sigma = 0.10$

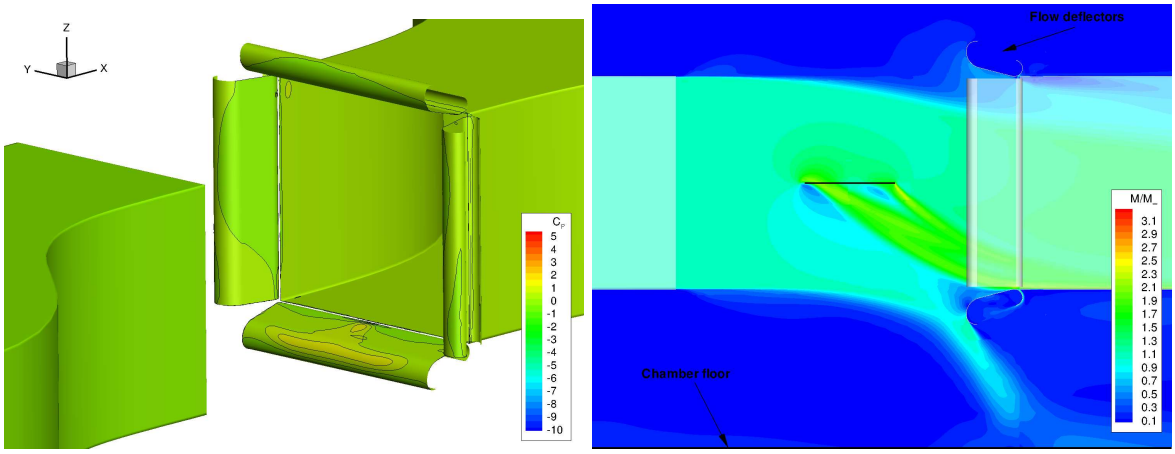


(c) $C_T/\sigma = 0.12$

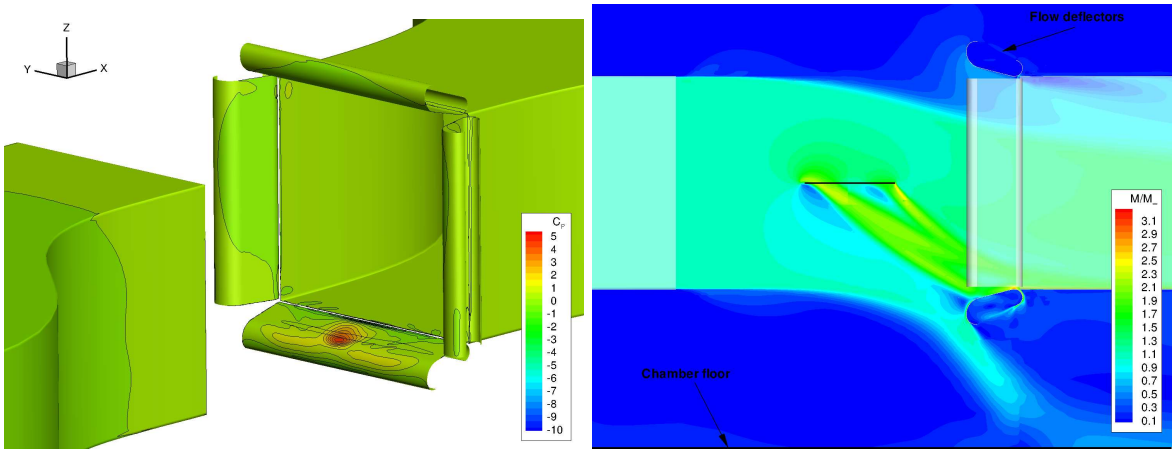
Figure 7: Pressure coefficient distribution on the tunnel walls (left) and Mach number distribution in the vertical simmetry plane of the test section (right) for $V_\infty = 10\,\mathrm{m/s}$.



(a) $C_T/\sigma = 0.08$



(b) $C_T/\sigma = 0.10$



(c) $C_T/\sigma = 0.12$

Figure 8: Pressure coefficient distribution on the tunnel walls (left) and Mach number distribution in the vertical simmetry plane of the test section (right) for $V_\infty = 15\,\mathrm{m/s}$.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Validation of numerical models for the investigation of the dynamic stall phenomenon

R.Nilifard, A.Guardone, G.Quaranta*, A.Zanotti* and G.Gibertini
Dipartimento di Ingegneria Aerospaziale – Politecnico di Milano
Via La Masa 34, 20156, Milano, Italy

e–mail: *quaranta,zanotti@aero.polimi.it

Keywords: *Dynamic stall experiments, Computational fluid dynamics, Turbulence models.*

Abstract

As a first step towards the design of dynamic stall control systems for helicopter performance improvement, a new experimental rig has been designed consisting in a blade section model supported by a motorized strut, which can move it in pitch around its quarter chord. The design and performance of the experimental rig have been purposely conceived to reproduce the deep dynamic stall condition of a full-scale retreating rotor blade section at 75% radius at high forward flight speed. Several measurement techniques, including fast unsteady pressure measurements and PIV, can be employed to completely characterise the time dependent flowfield. The paper describes a preliminary assessment of the dynamic stall phenomenon modeling capabilities, comparing different numerical CFD models against the experimental results. The objective is to obtain reliable numerical models to be used in the following phases of development of control devices for dynamic stall effect mitigation.

1 Introduction

A strong demand for faster helicopter has spurred the research on new rotorcraft design. Several design currently in advanced or preliminary development stage, like the tiltrotor (Bell Boeing V-22, Bell Agusta BA609, ERICA concept) or the compound (Sikorsky X2, or the Eurocopter X3), call for significant changes in the aircraft configuration. One of the problems that these configurations try to overcome is the dynamic stall on the retreating blade that limits the high speed performance of classical helicopter rotor configurations [1]. However, it is possible to mitigate or suppress the effect of dynamic stall by using several passive or active control techniques. Many recent activities analysed the effectiveness of different control systems integrated into the blade section in order to improve helicopter performance and expand the flight envelope and vehicle utility. Currently, improvements to control rotor blade dynamic stall rely upon the optimization of the blade airfoil shape, introducing, for example, a variable droop leading edge [3]. Moreover, the use of blowing devices such as air-jet vortex generators [4] or plasma actuators [5] represents an attractive solution for the reduction of the airloads hysteresis and the suppres-

sion of stall-driven flutter occurring during a blade pitching cycle [6]. In the aim of testing new flow control devices integrated on real helicopter blade sections, a new experimental-numerical activity started at Politecnico di Milano on this topic. The first step of this research, presented in this work, is to evaluate modeling capabilities, comparing different numerical CFD models against the experimental results. Wind tunnel tests have been carried out on a pitching NACA 23012 blade section. The airfoil has been selected since, being a typical helicopter blade airfoil, it has been employed in experimental activities in the past years about the study of the dynamic stall phenomenon on pitching blade sections [7, 8]. The main goal of this activity is to obtain reliable models to be used in the following phases of development of control devices for dynamic stall effect mitigation. For this purpose, several two-dimensional models have been built using EDGE [9], a compressible Navier-Stokes solver developed at FOI, the Swedish Research Agency. Two turbulence models, namely, the model by Hellsten [11, 12] and the $k-\omega$ SST by Menter [10], have been tested, to assess their performances in these peculiar unsteady flows. In the paper, the experimental results for deep dynamic stall conditions are presented together with a thorough comparison with the different numerical results.

2 Experimental set up

The experimental activity was conducted at Politecnico di Milano in the low-speed closed-return wind tunnel of the Aerodynamics Laboratory. The wind tunnel has a rectangular test section with 1.5 m height and 1 m width. The maximum wind velocity is 55 m/s and the turbulence level is less than 0.1%. A dynamic stall rig has been designed consisting in a blade section model supported by a motorized strut that can move it in pitch around its quarter chord, see the layout in Fig.1.

The employed NACA 23012 blade section model, with airfoil chord of 0.3 m and aspect

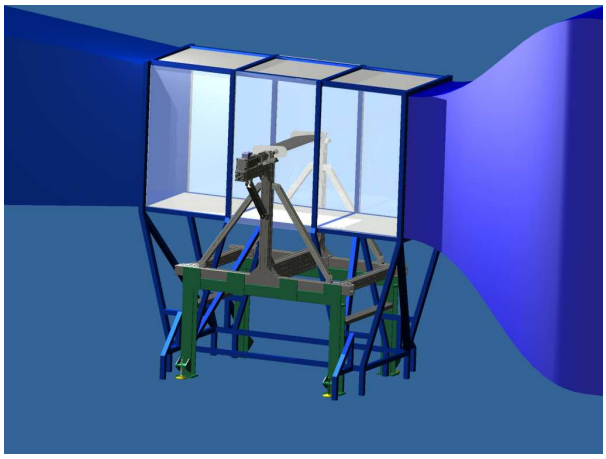


Figure 1: Layout of the dynamic stall experimental rig.

ratio of 3.1, was mounted horizontally in the wind tunnel test section and was pivoted about the quarter-chord position on two tubular steel shafts positioned on two self-aligning bearings, see Fig.2.

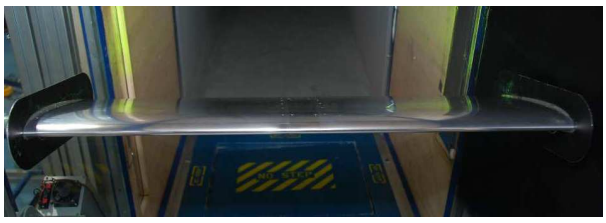


Figure 2: NACA 23012 blade section model in the wind tunnel.

The blade section model is composed by three aluminum machined external sections attached on an internal metallic frame in alminum. The model presents an interchangeable midspan section for the different measurements techniques employed, one for Particle Image Velocimetry (PIV) and another for unsteady pressure measurement equipped with pressure ports positioned along the midspan chord line. End plates were used during the tests to minimize interference effects of the walls boundary layer. The driving mechanism is composed by a brushless servomotor and a 12:1 gear drive, see the particular of the motorized strut in Fig.3. The driving mechanism arrangement is positioned on a cantilevered aluminum profile and provides 120 Nm of contin-

uous torque. The model in connected to the driving mechanism by a torsionally stiff steel laminae coupling between the model tubular shaft and the gear drive shaft. A 4096 imp/rev encoder is used to determine the instantaneous position of the model as well as for feedback control.



Figure 3: Motorized strut of the dynamic stall rig

2.1 Unsteady pressure measurement set up

The midspan chord line of the model central section is instrumented with 21 fast-response pressure transducers with a slight increase in concentration near the leading edge. The particular of the central section of the model instrumented with the pressure transducers is presented in Fig.4. The transducers signals were acquired with 50 kHz simultaneous sampling rate on 21 channels for a time period corresponding to 30 complete pitching cycles. The high sampling rate was suggested to capture the fine detail of the dynamic stall phenomenon characterized by severe unsteadiness conditions, especially at the relatively high reduced frequency tested $k = 0.1$. The integration of the phase averaged pressure signals acquired by the 21 miniature pressure transducers allowed to determine the time history of airloads, lift and pitching moment, in a pitching cycle and to evaluate its hysteresis.

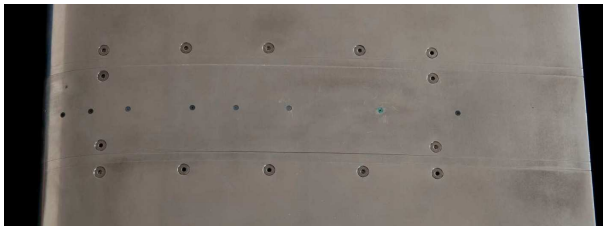


Figure 4: Particular of the model central section instrumented with pressure transducers.

2.2 Particle Image Velocimetry set up

Figure 5 presents the arrangement of the PIV devices in the wind tunnel test section.

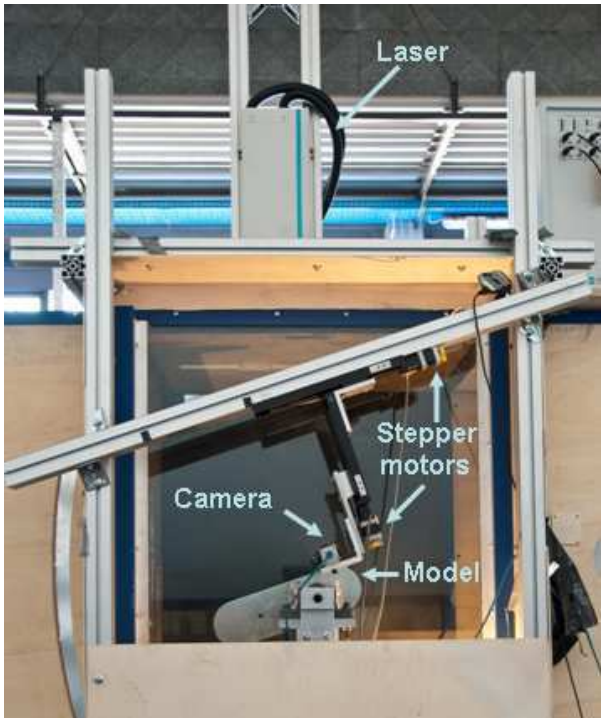


Figure 5: Layout of the PIV set up.

A double shutter CCD camera with a 12 bit, 1280×1024 pixel array and a 55 mm lens was used to acquire the image pairs. In order to

perform the flow survey above the entire airfoil upper surface, the measurement field was composed by 4 measurement windows spanning the airfoil chord direction. The CCD camera was mounted on a dual axis traversing system composed by two stepper motors that allowed to move the measurement window along two orthogonal directions. The PIV system used a Nd:Yag double pulsed laser with 200 mJ output energy and a wavelength of 532 nm. The laser sheet passed through an opening in the wind tunnel roof aligned with the flow and positioned in the midspan of the test section width. The laser was mounted on a single axis traversing system to move the sheet along the wind tunnel flow direction, enabling the use of a smaller width sheet with higher energy centered on the measurement window. The tracer particles injected in the wind tunnel are small oil droplets with diameter in the range of 1-2 μm, generated by means of Laskin nozzles.

3 CFD solver

The CFD solver used for this study is the EDGE code developed by FOI [9]. The code is capable of solving flows with different conditions from inviscid to fully turbulent using various turbulence models. RANS (Reynolds-Averaged Navier-Stokes), DES (Detached-Eddy Simulation) and LES (Large Eddy simulation) models in 2/3 dimensions are considered in this code.

3.1 Numerical methodology

The flow equations for viscous turbulent flows are solved on the structured grid using the node-centered finite volume solver EDGE [9]. Inviscid fluxes are discretized using a central scheme with artificial dissipation. A blend of second and fourth order differences are chosen as artificial dissipation to eliminate the effect of oscillations in vicinity of shock waves or stagnation points. The viscous fluxes are obtained from application of Green-Gauss'formulations. A dual time stepping method as a driver is applied to march implicitly in time. Furthermore, this ensures to has unconditional stability, also the time step size is mostly chosen based on accuracy rather than stability. To accelerate the convergence within each real time step, local time stepping, multigrid and implicit residual smoothing are used. Low Mach number preconditioning is used to solve the incompressoble flows of interest here. To start the unsteady computation, a steady state solution at the mean angle of incidence is used as the initial solution. In both cases, the periodic solutions for lift and moment coefficient are normally obtained after two oscillation cycles. The time step is chosen to have 136 time steps per period in all test cases.

Turbulence models used for both test cases are Menter SST $k-\omega$ [10] and EARSM (Explicit Algebraic Reynolds Stress Model) by Wallin & Johansson coupled to Hellsten $k-\omega$ [11, 12].

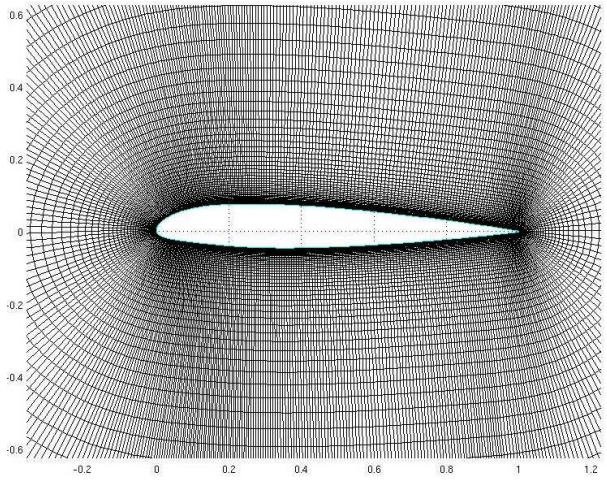


Figure 6: Close view of the used O-type grid.

The structured grid used in all numerical computations is an O-type grid made of 530×160 (84800) quadrilateral elements, with a higher resolution of the grid near leading and trailing edge. The wall-normal grid resolution of the closest internal node from wall is of the order of $y+ = 1$, which gives a distance between the wall and the first layer of elements of about $10^{-5}c$, with c chord length. The outer boundary is at $20c$.

In Fig.6, a close view of computational grid is shown. In both test cases, a sinusoidal pitching motion is imposed on the whole grid which therefore rotates as like a rigid body.

4 Results

The test cases considered in this work are two pitching cycles characterized respectively by a mean angle of attack of 10° and 15°, with a 10° amplitude and reduced frequency equal to 0.1. The experimental tests and the numerical simulations have been carried out at Reynolds number $\mathrm{Re} = 1 \cdot 10^6$ and Mach number Ma= 0.15 in order to reproduce the condition of a full-scale retreating rotor blade section at 75% radius in forward flight. The experimental tests have been carried out also at $Re = 6 \cdot 10^5$ to evaluate the effect of Reynolds number on the airloads in dynamic conditions. Figure 7 presents the comparison of the measured airloads with the results of the simulations carried out using the Hellsten and Menter $k - \omega$ turbulence models.

For the two conditions tested, the numerical and experimental curves are in good agreement. Indeed, the amount of airloads hysteresis over a pitching cycle and the behavior of the airloads curves is similar. For the pitching cycle case with a mean angle of attack of 10°, the numerical results, for both turbulence models, do not reproduce the increase of lift and pitching moment that is observed in the experimental curves during the downstroke motion, for an angle of attack between 18° and 14°. This behavior in the experimental curve is explained by the formation and migration over the airfoil upper surface of a vortex disturbance that the numerical simulations are not capable to reproduce. The formation of this vortex disturbance during the downstroke motion is well observed in Fig.8 that presents the Z-component vorticity field (normalized by the maximum value) obtained from PIV at 16° of angle of attack over the upper surface of the airfoil pitching with the same sinusoidal law. The flow field has been measured at a lower Reynolds number equal to $Re = 6 \cdot 10^5$ but its effect on the dynamic airloads are quite small as the measured airloads curves at the different Reynolds

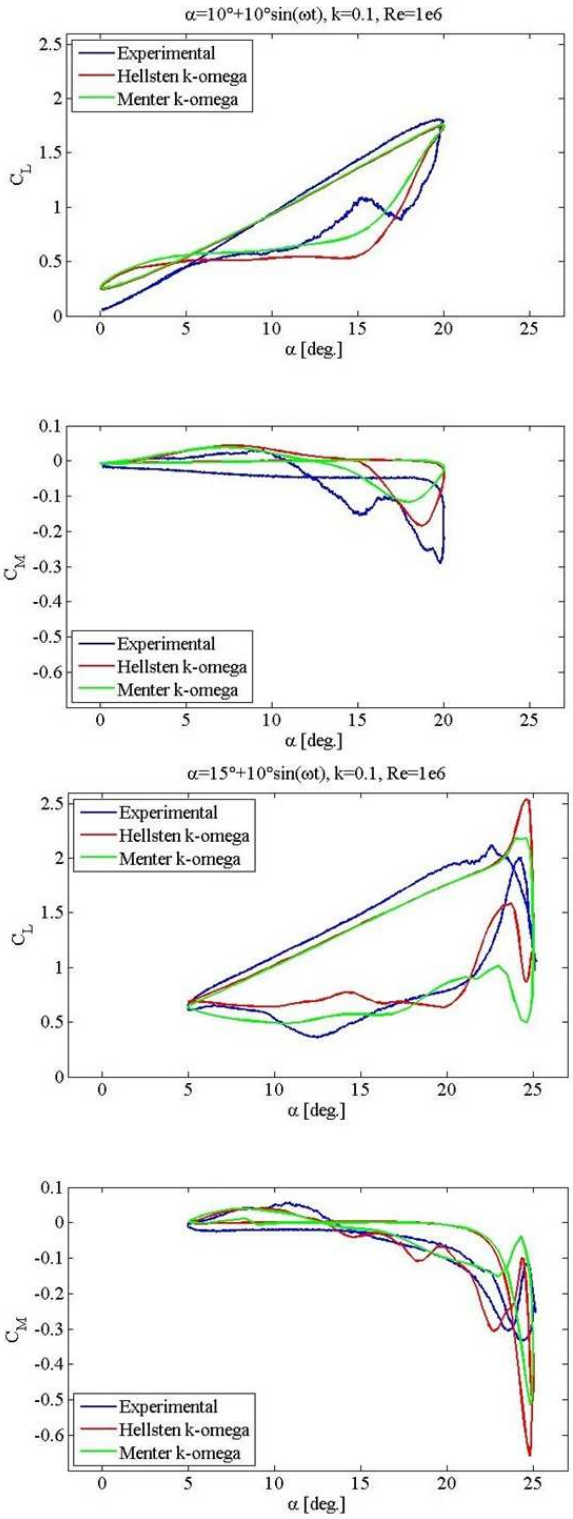


Figure 7: Comparison of experimental and numerical results for pitching airfoil in dynamic stall conditions

numbers reproduce the same characteristic issues.

For the pitching cycle case with 15° of mean angle of attack, the numerical results, for both turbulence models, present a delay with respect to the experimental curve of the non-linear increase of lift during the upstroke motion due to the formation and migration of the Dynamic Stall Vortex [2]. Moreover the numerical curves present a peak of lift and pitching moment during the cycle greater than the measured ones.

5 Conclusions

A test rig for investigating dynamic stall conditions was built at the Aerodynamic Laboratory of the Aerospace Department, Politecnico di Milano. The test rig was designed to reproduce the flow conditions and the airfoil motion representative of a full-scale retreating rotor-blade section at 75% radius at high-speed forward flight. Both high-response pressure measurements and particle image velocimetry were used to measure the flow features. The finite-volume unstructured-grid code EDGE by FOI was used to simulate the experimental conditions, using two different turbulence models. Numerical simulations were found to be in good agreement with the experimental ones, with the only exception of the lift peak observed experimentally during the downstroke phase. In this situation, both the SST k-ω and the EARSM turbulence models failed to capture the formation and propagation of a vortex generated at the airfoil leading edge, a failure that partially explains the disagreement between the experimental and numerical results observed during the first half of the downstroke movement. Current research activities are devoted to the investigation of diverse dynamic stall conditions and to the evaluation of more advanced numerical techniques, such as the Detached Eddy Simulation or the Large Eddy Simulation, to improve the accuracy of the numerical model.

References

[1] McCroskey W.J., "The Phenomenon of Dynamic Stall", NASA TM 81264, 1981.

[2] Leishman J.G., *Principles of helicopter aerodynamics*, Cambridge Aerospace Series, 2000.

[3] Chandrasekhara M., Martin P. and Tung, C., "Compressible Dynamic Stall Control Using a Variable Droop Leading Edge Airfoil", *Journal of Aircraft*, Vol.41, 2004, pp.862-869.

[4] Singh C., Peake, D., Kokkalis A., Khodagolian V., Coton F. and Galbraith R., "Control of Rotorcraft Retreating BladeStall Using Air-Jet Vortex Generators", *Journal of Aircraft*, Vol.43, 2006, pp.1169-1176.

[5] Post M. and Corke T., "Separation Control Using Plasma Actuators: Dynamic Stall Vortex Control on Oscillating Airfoil", *AIAA Journal*, Vol.44, 2006, pp. 3125-3135.

[6] Carta F.O., "An Analysis of the Stall Flutter instability of Helicopter Rotor Blades", *Journal of American Helicopter Society*, Vol.12, 1967, pp.1-8.

[7] Leishman J.G., "Dynamic stall experiments on the NACA 23012 aerofoil", *Experiments in Fluids*, Vol.9, 1990, pp.49-58.

[8] Raffel M., Kompenhans J., Stasicki B., Bretthauer B. and Meier G.E.A., "Velocity measurement of compressible air flows utilizing a high-speed video camera", *Experiments in Fluids*, Vol.18, 1995, pp.204-206.

[9] Eliasson P., "EDGE, a Navier-Stokes solver for unstructured grids". *Proc. to Finite Volumes for Complex Applications III*, Vol.1, 2002, pp.527-534.

[10] Menter F., Kuntz M. and Langtry R., "Ten years of industrial experience with the SST turbulence model", *Turbulence, heat and mass transfer*, Vol.4, 2003, pp.625-632.

[11] Wallin S. , "Engineering turbulence modeling for CFD with a focus on explicit algebraic Reynolds stress models ", *Doctoral Thesis Norsteds Tryckeri AB*, Stockholm, Sweden,2003.

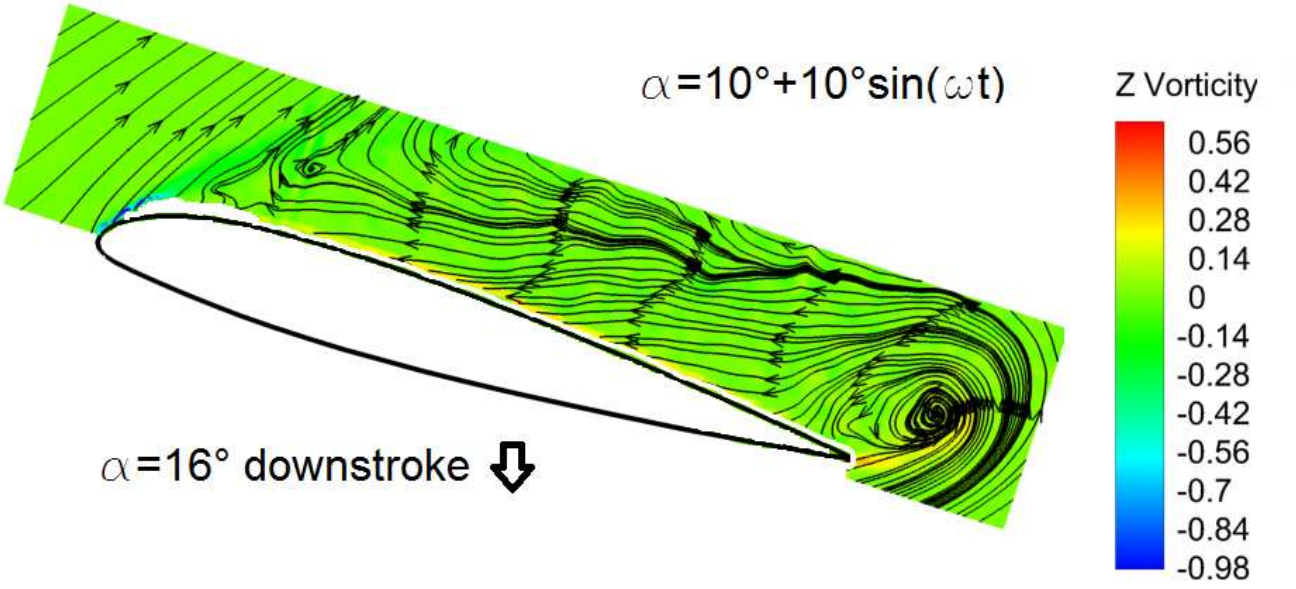


Figure 8: PIV vorticity field over the upper surface of the NACA23012 pitching airfoil at 16° of incidence in downstroke

[12] Hellsten A., "New advanced $k-\omega$ turbulence model for high-lift aerodynamics", *AIAA Journal*, Vol. 43, No.9, 2005, pp.1857-1869.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Experimental-Numerical study of small amplitude oscillations of a pitching airfoil

S. De Bernardi, G. Gibertini, A. Guardone, V. Motta, R. Nilifard, G. Quaranta and A. Zanotti
Politecnico di Milano - Dipartimento di Ingegneria Aerospaziale, Italy

Keywords: Oscillating airfoil, Wind Tunnel, Computational fluid dynamics, Turbulence models.

Abstract

In the aim of research about unsteady airfoil aerodynamics, a new experimental rig has been designed consisting in a blade section model supported by a motorized strut that can move in pitch around its quarter chord. The experimental setup allowed to employ several measurement techniques as fast unsteady pressure measurements and Particle Image Velocimetry to characterise the time dependent flowfield. The present paper describes the comparison of CFD simulations results with the data set produced by experiments carried out on a pitching airfoil with small oscillation amplitudes. The investigation of this phenomenon currently has become a research field of increasing interest due to its application in the design and development of new concept of micro aerial vehicles.
The analysis of the flowfield survey in the wake of the airfoil and the comparisons with the results of the numerical simulations enabled to assess and better tune the experimental and numerical tools.

1 Introduction

The present paper describes the results obtained by an experimental test campaign and numerical simulations on the case of a NACA 23012 airfoil slightly oscillating in pitch. This activity, that was carried out to assess the reliability of the experimental and numerical tools, was preceded by measurements and simulations of the steady flow to allow for a first tune of the numerical simulation.
The experimental setup allowed to use different measurement techniques as, for instance, fast unsteady pressure measurements, hot wire anemometry and PIV. The investigation of small amplitudes pitching airfoils has been object of several numerical [1] and experimental [2] activities due to, for instance, its application in the design and development of new concept of micro aerial vehicles that currently has become a research field of increasing interest. The first step of this research, presented in this work, is to evaluate modeling capabilities, comparing different numerical CFD models against the experimental results. Wind tunnel tests have been carried out on a pitching NACA 23012 blade section. The airfoil has been selected since, being a typical helicopter blade airfoil, it has been employed in experimental activities in the past years about the study of the dynamic stall phenomenon on pitching blade sections [3, 4]. The two-dimensional models have been built using EDGE [5], a compressible Navier-Stokes solver developed by the Swedish research agency FOI. The different measurement techniques employed during the test activity produced a complete set of experimental data to be compared with the numerical simulations results. A complete PIV flowfield survey in the wake of the air-

foil at midspan has been performed to evaluate the flow physics concerning the reproduced unsteady conditions [9]. The analysis of the complete experimental data set and the comparisons with the numerical simulations enabled to assess and better tune the experimental and numerical tools.

2 Experimental set up

The experimental activity was conducted at Politecnico di Milano in the low-speed closed-return wind tunnel of the Aerodynamics Laboratory of the Aerospace Department. This facility has a 1.5 $m \times 1$ m rectangular test section, a maximum wind velocity of 55 m/s and a turbulence level less than 0.1%. The oscillating airfoil rig consists in a blade section model supported by a motorized strut that can move it in pitch around its quarter chord, see the layout in Fig.1.

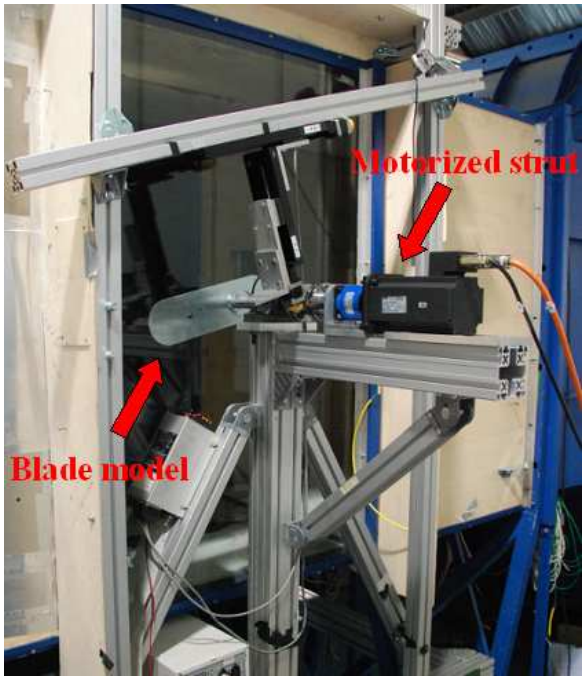


Figure 1: The experimental rig.

The aluminium machined airfoil model had NACA 23012 section, the chord $c = 0.3$ m and span $b = 0.93$ m (so that the aspect ratio was 3.1) was instrumented with 21 fast-response pressure transducers with a slight increase in concentration near the leading edge. The trasducers signals were acquired with 50 kHz simultaneous sampling rate on 21 channels for a time period corresponding to 30 complete pitching cycles. Lift and pitching moment have been obtained by integration of the measured pressures.

Wake flow field has been surveyed by means of a Particle Image Velocimetry (PIV) system. A double shutter 1.3 Mpx CCD camera with a 55 mm lens was used to acquire the image pairs. In order to perform the flow survey along the airfoil wake, the measurement field was composed by two aligned 104 $mm \times 83$ mm measurement windows (the camera was mounted on a motorized guide) with a 10 mm overlapping. The PIV system used a Nd:Yag double pulsed laser with 200 mJ output energy and a wavelength of $532nm$.

In order to filter out the effect of turbulent oscillations the results, both the aerodynamic coefficents and the flow field, have been phase averaged.

3 CFD model

3.1 Governing Equations

The flow equations in an Arbitrary Lagrangian Eulerian (ALE) framework read

$$\frac{d}{dt}\int_v \mathbf{u} dv = -\int_s [F_I(\mathbf{u}) + F_V(\mathbf{u}) - \mathbf{u}\dot{\mathbf{x}}_\mathbf{f}]\, ds \tag{1}$$

where $\mathbf{u} = [\rho, m, E^t]$ is the vector of the conservative variables density, momentum and total energy per unit volume, and $F_I(\mathbf{u})$ and $F_V(\mathbf{u})$ are the inviscid and viscous flux matrices, respectively. In the above equations, v_i is the control volume, s is its surface and $\dot{\mathbf{x}}_f(\mathbf{s}, t)$ is the local velocity of the surface. As it is well known, the Reynolds Averaged Navier-Stokes (RANS) equations (1) are obtained by time averaging the Navier–Stokes system, with a first order closure, based on Boussinesq's assumption.

In the present work, the dynamic viscosity is kept constant because the temperature differences are small within the flowfield. The Spalart-Allmaras turbulence model is used to compute the turbulent viscosity. The laminar

and turbulent thermal conductivity are computed by assuming a constant Prandtl and turbulent Prandtl number, respectively. Air is threated a calorically perfect gas since the gas temperature is low enough so that the vibrational and electronic modes are not excited.

3.2 Numerical method description

The finite volume discretization of the RANS system is obtained by applying the integral formulation of the governing equations to a finite number of control volumes, each defined around a given node of the triangulation. Defining $F = F_I + F_V$ as the total flux the discrete form of the expression (1) can be written as:

$$\frac{d(\Omega_i \mathbf{u}_i)}{dt} = \sum_{k \epsilon T_i} \Phi(\mathbf{u}_i, \mathbf{u}_k, S_{ik}, \mu_{ik}) + \Phi^{\partial}(\mathbf{u}_i, S_i, \mu_i), \tag{2}$$

where T_i is the set of nodes k sharing a portion of their boundary with the volume of node i, node i excluded. Furthermore S_{ik} is the integrated outward normal of the surface portion associated with the couple of nodes ik,$S_{ik} = \int_{\partial\Omega_{ik}} dS_i(s,t)$ and $\mu_{ik} = \int_{\partial\Omega_{ik}} \dot{\mathbf{x}}_f dS_i(s,t)$ is the integrated interface velocity. In order to have a numerical scheme to be stable the grid velocities μ_{ik} and the and the grid dependent geometric quantities , such as for example the volume associated with the finite volume of the integrated outward normals, are to be computed according to suitable consistency conditions, usually referred as "Geometric Conservation Law" [12]. It is possible to define either a central scheme or an upwind scheme to discretize the convective terms of equations (1). According to a piecewise representation of finite volumes, the term Φ^{∂}, accounting for the flux boundary contribution on the frontiers of the computational domain is

$$\Phi^{\partial}(\mathbf{u}_i, S_i, \mu_i) = F(\mathbf{u}_i) S_i - \mathbf{u}_i \mu_i. \tag{3}$$

Finally, the flow equations are integrated explicitly using a multi-stage Runge-Kutta scheme with convergence acceleration by agglomeration multigrid and implicit residual smoothing. A central scheme plus a Jameson type diffusion is used for the spatial discretization. Since the multigrid approach (even if not the full multigrid one) has been used, on coarser elements a simplified form of the artificial dissipation operator is used to save computational time. This one is characterized by a dissipation coefficient whose standard value corresponds to $vis_z = 0.15$. The second order dissipation term that is active in the neighborhood of shocks ($vis_{II} = 0.5$). The fourth order dissipation term is used to remove high-order oscillations ($vis_{IV} = 0.001$).

Time-accurate solutions are obtained using an implicit dual time-stepping technique with explicit sub-iterations and a three-point backward differences formula

$$\frac{d(\Omega_i^* \mathbf{u}_i^*)}{d\tau} = \tilde{\Phi}(\mathbf{u}^*, S_{ik}, \mu_{ik}) + Q, \tag{4}$$

with the unsteady residual

$$\begin{aligned} \tilde{\Phi}(\mathbf{u}^*, S_{ik}^*, \dot{x^*}_{ik}) = & \frac{3\Omega_i^* u*_i}{2\triangle t} \\ & + \sum_{k \epsilon T_i} \Phi(\mathbf{u}_i, \mathbf{u}_k, S_{ik}, \mu_{ik}) \\ & + \Phi^{\partial}(\mathbf{u}_i, S_i, \mu_i) \end{aligned} \tag{5}$$

and the source term reads

$$Q = \frac{4\Omega_i^n \mathbf{u}_i^n - \Omega_i^{n-1} \mathbf{u}^{n-1}}{2\Delta t}. \tag{6}$$

The stationary solution of the relationship (4) corresponds to the flow variables at the new time level, i.e. $\mathbf{u}^* = \mathbf{u}^{n+1}$. Since the residual is driven to zero at steady state in pseudo time, hence $\frac{d}{d\tau} = 0$ the equation 2 is fullfilled.

4 Results

Before the dynamic tests, static measurements and simulations have been carried out on the NACA 23012 airfoil at $Re = 1 \cdot 10^6$. The comparison of the experimental and numerical $C_L - \alpha$ curves presents a very good agreements as can be observed in Fig. 2.

The dynamic tested condition consists in pitching oscillation of the NACA 23012 around the zero-lift angle of attack (i.e. $\alpha_0 \cong -1°$) with

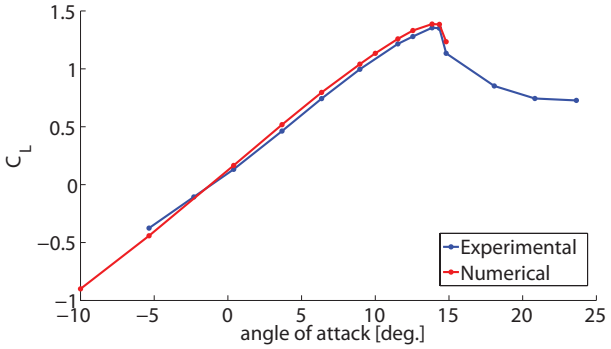


Figure 2: Steady $C_L - \alpha$ curve at $Re = 1 \cdot 10^6$.

2° of oscillation amplitude. The reduced frequencies reproduced experimentally and numerically are $k = 0.1$ and $k = 0.2$. The reduced frequencies have been obtained with model frequency respectively of $f = 3.18\ Hz$ and $f = 6.37\ Hz$ at the flow velocity $V = 30\ m/s$ (corresponding to $Ma = 0.09$ and $Re = 6 \cdot 10^5$).
For both the two tested reduced frequencies an hysteretic behavior of lift and moment coefficient was found by the experiments and by the simulations as well although the amplitude of the histeresys is not the same. In Fig. 3 and 4 these results are compared with the Theodorsen theory: the measured data result to be in quite good agreement with the theory except for the lift coefficiente at $k = 0.1$, while the numerical results present sensible differences that probably are due to a rather low resolution of the mesh in the far wake region.

Figures 5-8 present the comparison of the velocity profiles in the wake of the oscillating airfoil at maximum, minimum and zero-lift angle of attack both in upstroke and downstroke. In particular the velocity profiles are plotted at a distance from the airfoil trailing edge in the range between 5% and 20% of the chord.

The comparison of the velocity profiles demonstrates very good results as the numerical curves reproduce very well the behavior of the experimental curves derived from the PIV flow surveys carried out for the same pitching cycle conditions. In fact, the numerical and experimental flow field present a very good agreement as can be seen from figures 9-12 that illustrate the contours of the adimensional velocity mag-

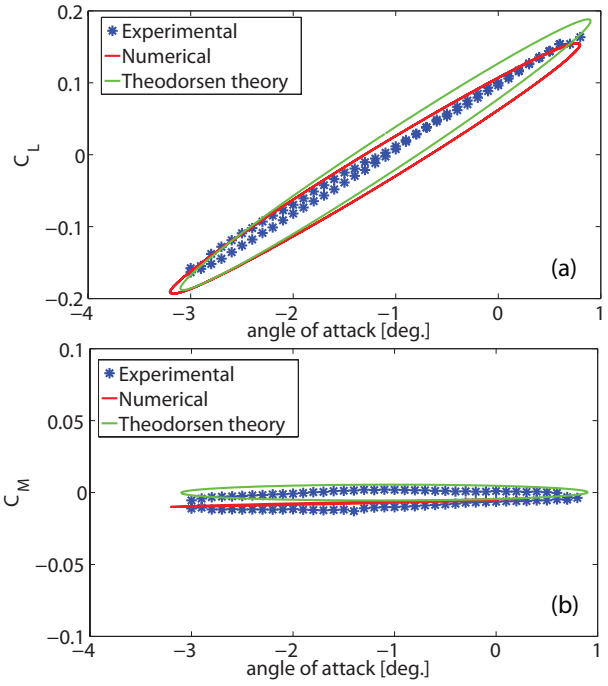


Figure 3: Airloads coefficient cycle for $k = 0.1$: (a) lift coefficient, (b) pitching moment coefficient.

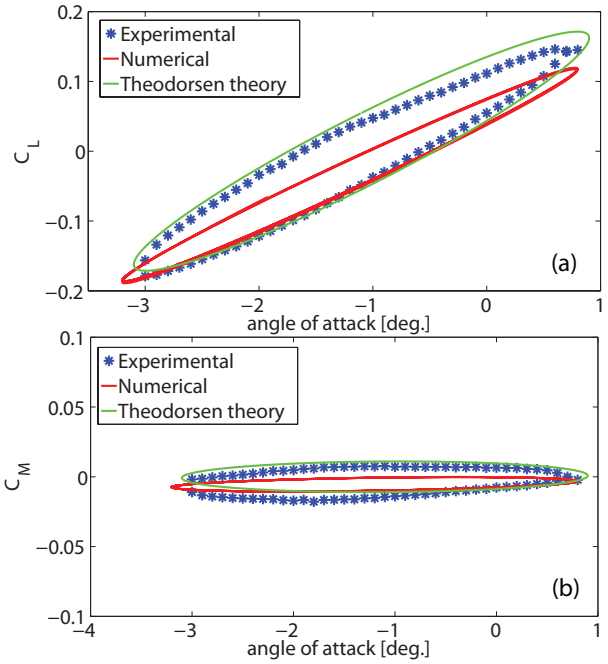


Figure 4: Airloads coefficient cycle for $k = 0.2$: (a) lift coefficient, (b) pitching moment coefficient.

nitude in the airfoil near field at maximum, minimum and zero-lift angle of attack both in up-

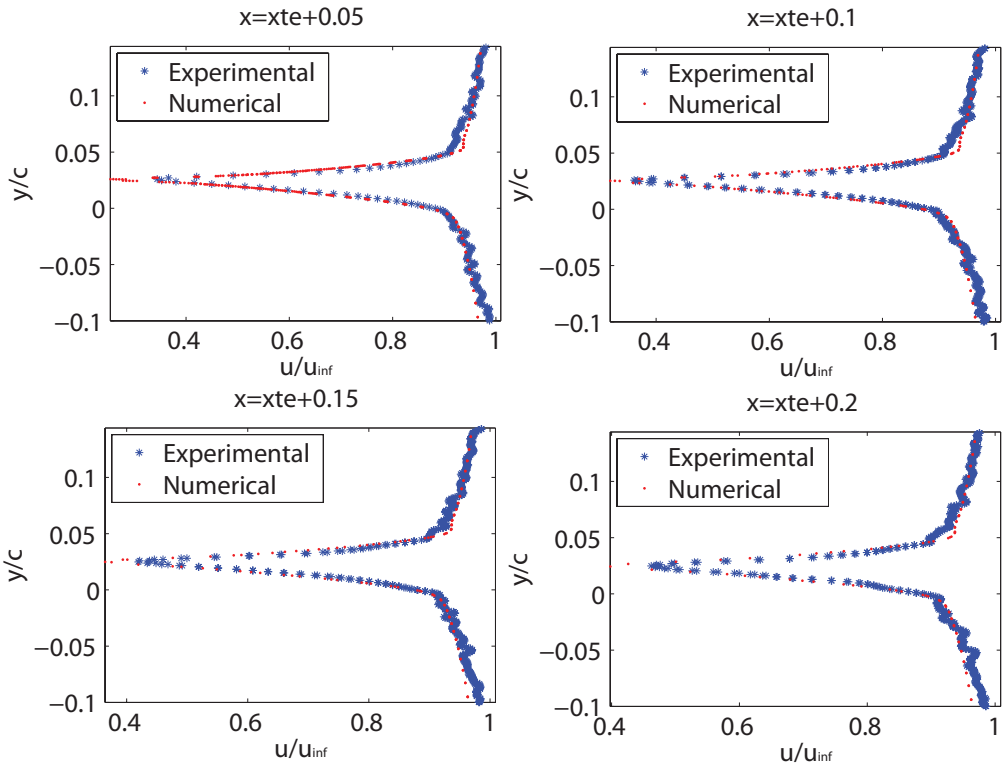


Figure 5: Velocity profiles in the wake of the oscillating airfoil at $\alpha_0 - 2°$ for $k = 0.2$.

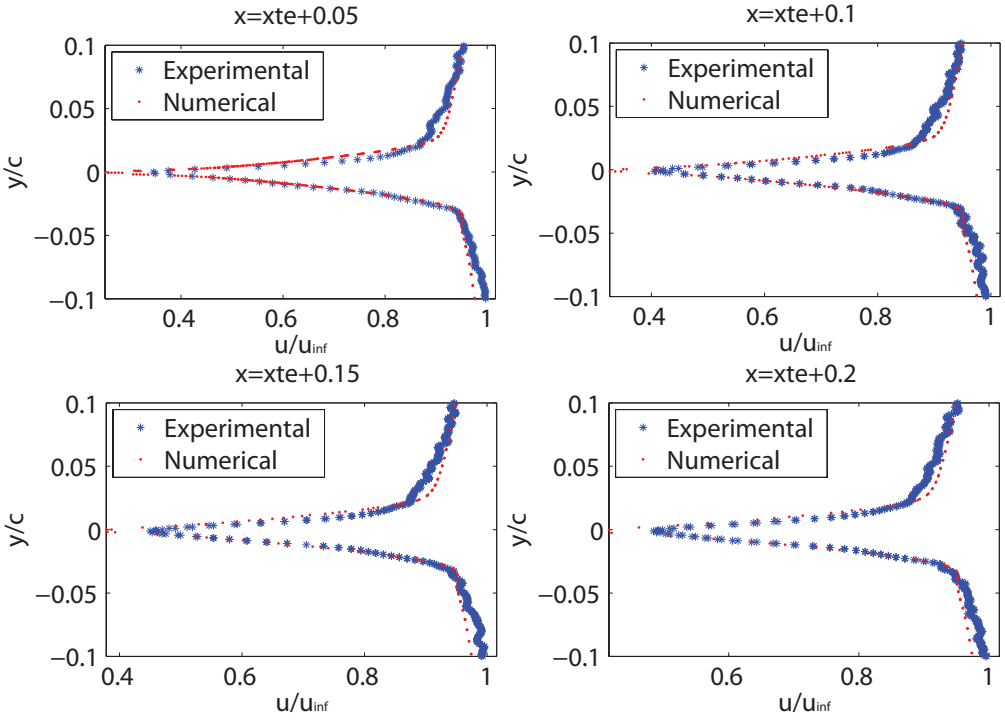


Figure 6: Velocity profiles in the wake of the oscillating airfoil at α_0 in upstroke for $k = 0.2$.

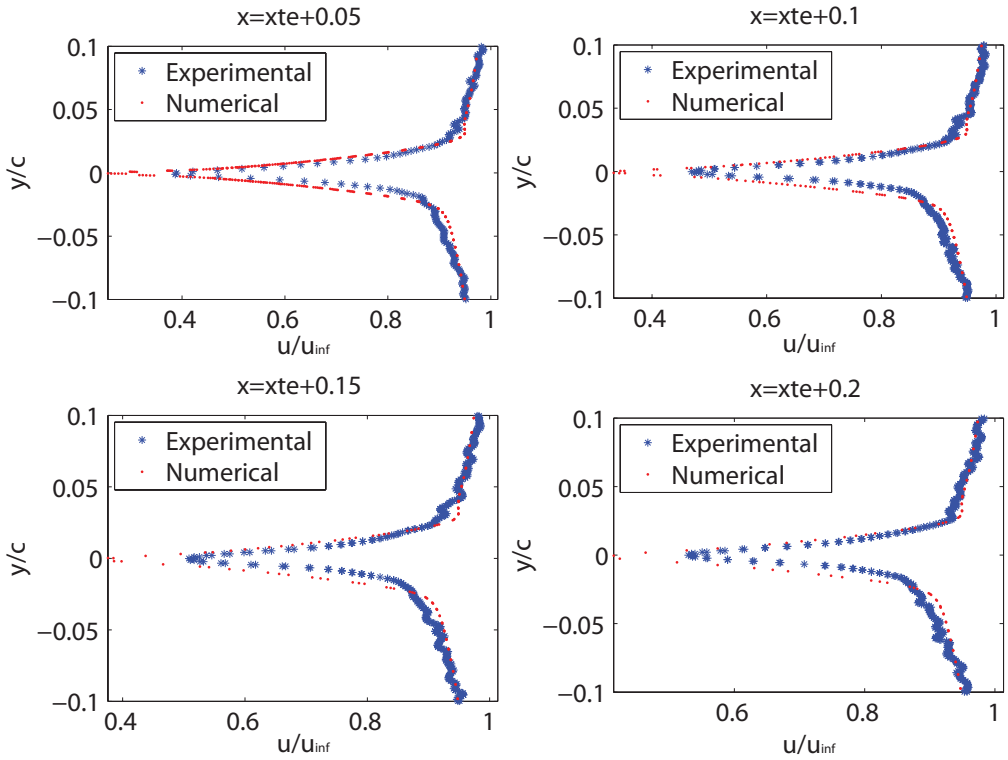


Figure 7: Velocity profiles in the wake of the oscillating airfoil at α_0 in downstroke for $k = 0.2$.

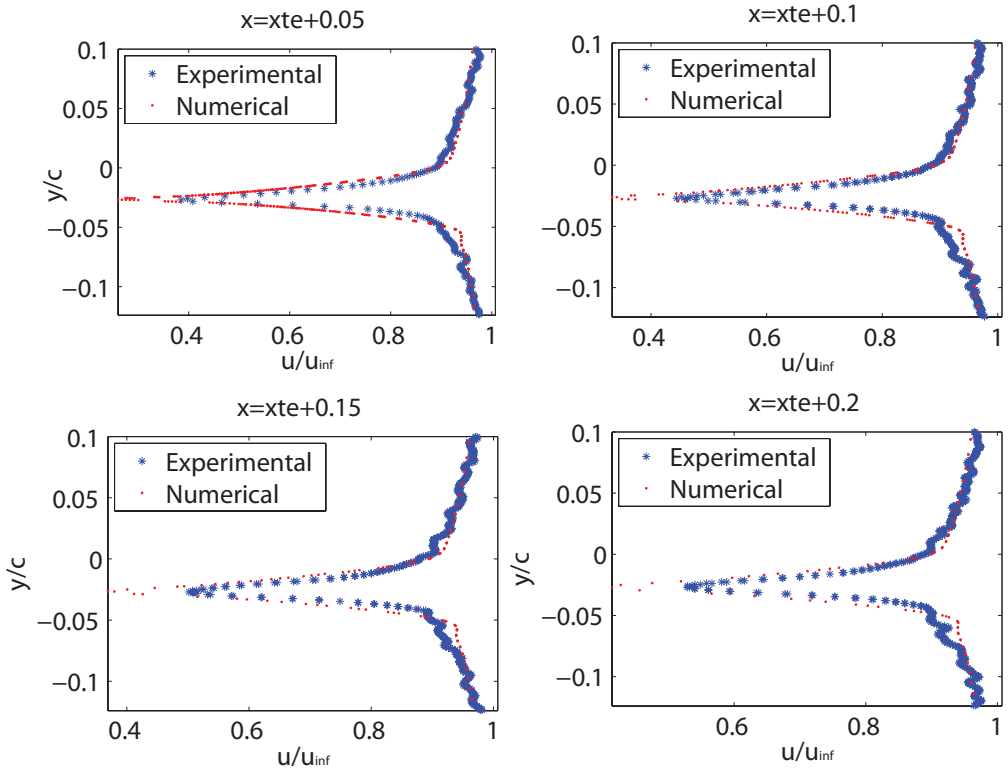


Figure 8: Velocity profiles in the wake of the oscillating airfoil at $\alpha_0 + 2°$ for $k = 0.2$.

stroke and downstroke. The velocity step across the wake at α_0, due to the circulation issued from the trailing edge, is clearly visible from the measurements and accurately predicted by the simulations.

5 Conclusions

A new experimental rig has been designed for testing oscillating airfoil in pitch. The experimental rig allowed to employ fast unsteady pressure measurements and Particle Image Velocimetry to completely characterise the time dependent flowfield. The experimental rig has been involved for an experimental campaign on a NACA 23012 airfoil section pitching around the zero-lift angle of attack with small amplitudes at different reduced frequency. The tests results have been compared with simulations using the Navier-Stokes solver EDGE by FOI reproducing the same experimental conditions. The comparison of the results enabled to assess and better tune the experimental and numerical tools at low reduced frequency. In particular, both the velocity profiles and the flow field in the airfoil near wake present a very good correlation, while improvements in the employed CFD model are necessary for better capturing the hysteresis of the airloads cycle measured experimentally. In this aim, future developments of the present work will focus on simulations and tests of the same airfoil pitching with small oscillation amplitudes at a higher reduced frequency $k = 1$, in order to study the flow field characterized by severe condition of unsteadiness.

References

[1] Chandar D. D. J. and Damodaran M., "Computational Study of the Free Flight of a Flapping Wing at Low Reynolds Numbers", 46th AIAA Aerospace Sciences Meeting and Exhibit, 7 - 10 January 2008, Reno, Nevada.

[2] Kuo C.H. and Hsie J.K.., "Unsteady flow structure and vorticity convection over the airfoil oscillating at high reduced frequency, *Experimental Thermal and Fluid Sciences*, Vol.24, 2001, pp.117-129.

[3] Leishman J.G., "Dynamic stall experiments on the NACA 23012 aerofoil", *Experiments in Fluids*, Vol.9, 1990, pp.49-58.

[4] Raffel M., Kompenhans J., Stasicki B., Bretthauer B. and Meier G.E.A., "Velocity measurement of compressible air flows utilizing a high-speed video camera", *Experiments in Fluids*, Vol.18, 1995, pp.204-206.

[5] Eliasson P., "EDGE, a Navier-Stokes solver for unstructured grids". *Proc. to Finite Volumes for Complex Applications III*, Vol.1, 2002, pp.527-534.

[6] Menter F., Kuntz M. and Langtry R., "Ten years of industrial experience with the SST turbulence model", *Turbulence, heat and mass transfer*, Vol.4, 2003, pp.625-632.

[7] Wallin S. , "Engineering turbulence modeling for CFD with a focus on explicit algebraic Reynolds stress models ", *Doctoral Thesis Norsteds Tryckeri AB*, Stockholm, Sweden,2003.

[8] Hellsten A., "New advanced $k - \omega$ turbulence model for high-lift aerodynamics", *AIAA Journal*, Vol. 43, No.9, 2005, pp.1857-1869.

[9] Ramaprian B. R. and Zheng Y., "Near Field of the Tip Vortex Behind an Oscillating Rectangular Wing", *AIAA Journal*, Vol.36, 1998, pp.1263-1269.

[10] Stanek M.J. and Visbal, M.R., "'Study of the Vortical Wake An Oscillating Airfoil"', *AIAA Paper*, 89-0554, 1989.

[11] Ko S. and McCroskey W.J., "'Computations of unsteady separating flows over an oscillating airfoil"', *AIAA Paper*, 95-0312, 1995.

[12] Thomas P.D. and Lombard C.K., "'Geometric conservation law and its application to flow computations on moving grids"', *AIAA Journal*, Vol. 17, pp.1030-1037, 1979.

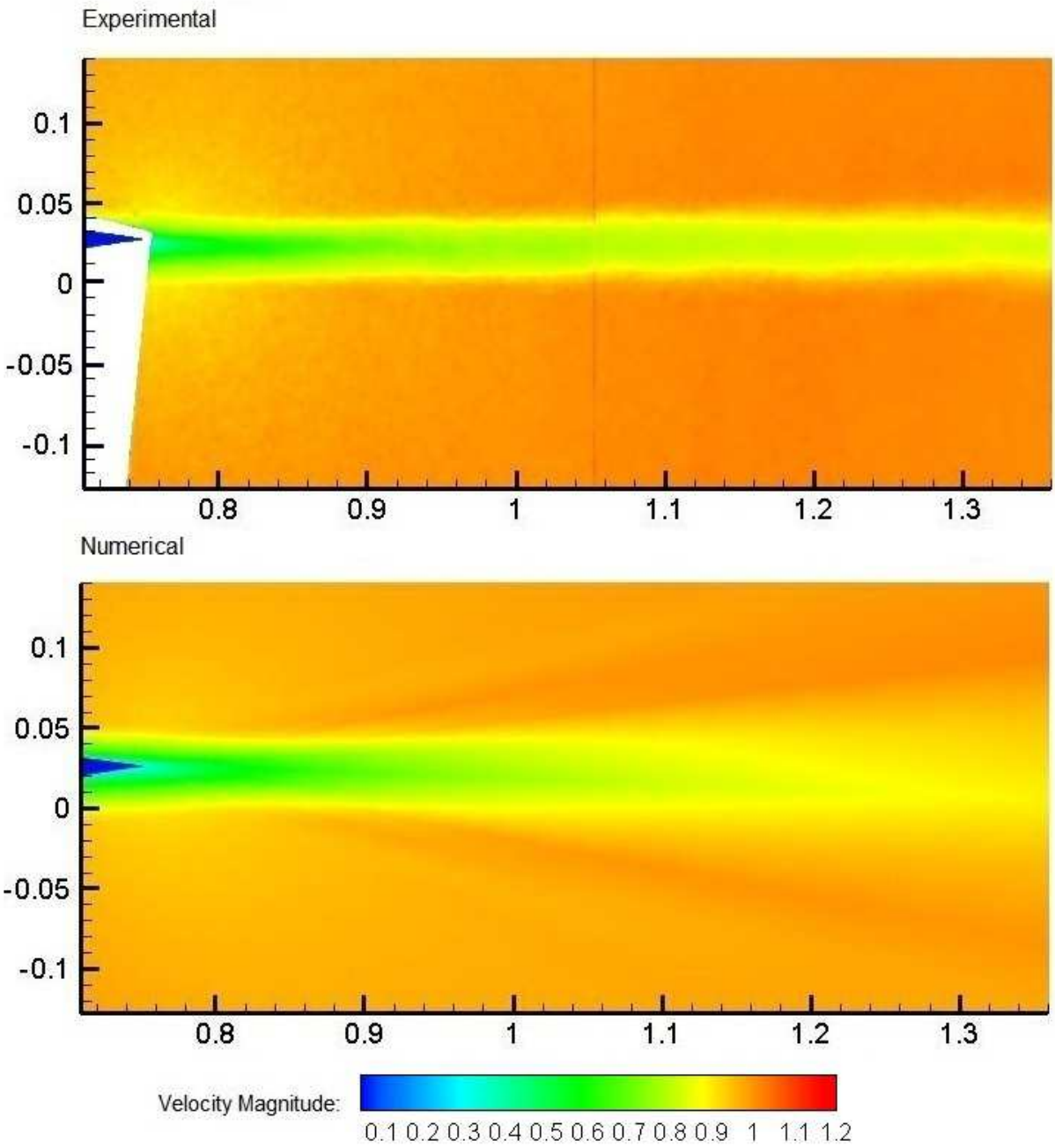


Figure 9: Contours of the adimensional velocity magnitude in the near wake of the oscillating airfoil at $\alpha_0 - 2°$ for $k = 0.2$.

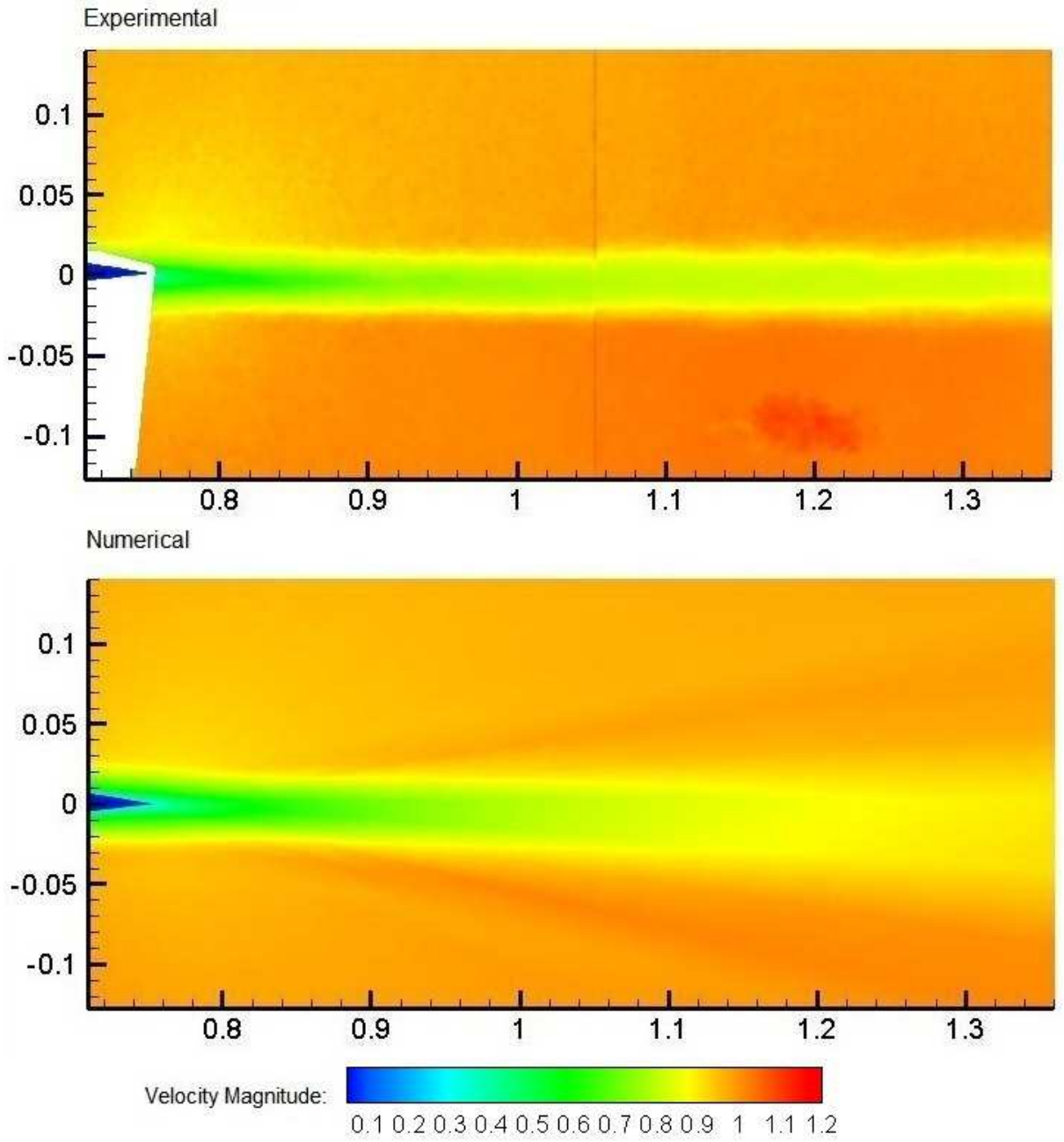


Figure 10: Contours of the adimensional velocity magnitude in the near wake of the oscillating airfoil at α_0 in upstroke for $k = 0.2$.

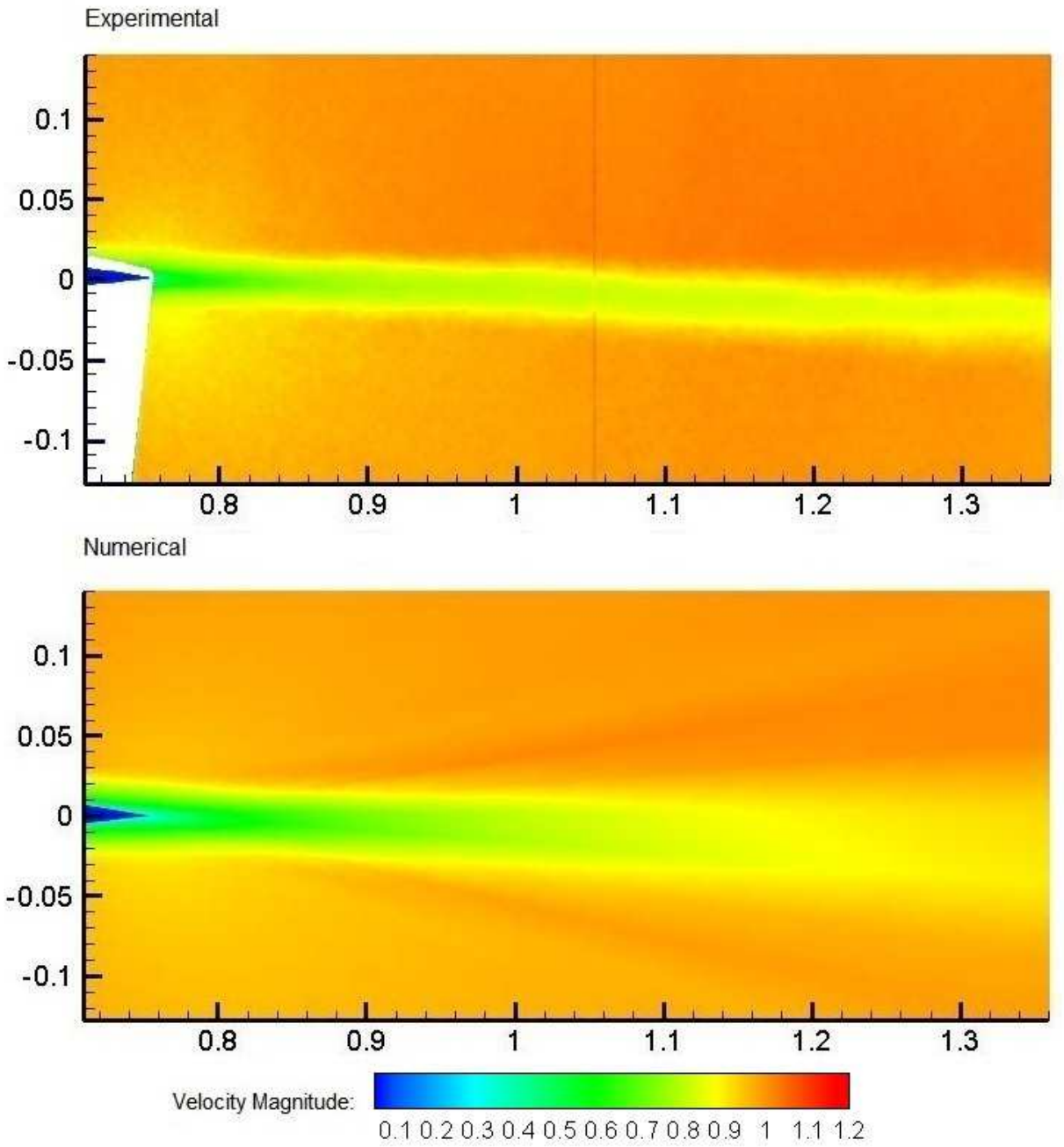


Figure 11: Contours of the adimensional velocity magnitude in the near wake of the oscillating airfoil at α_0 in downstroke for $k = 0.2$.

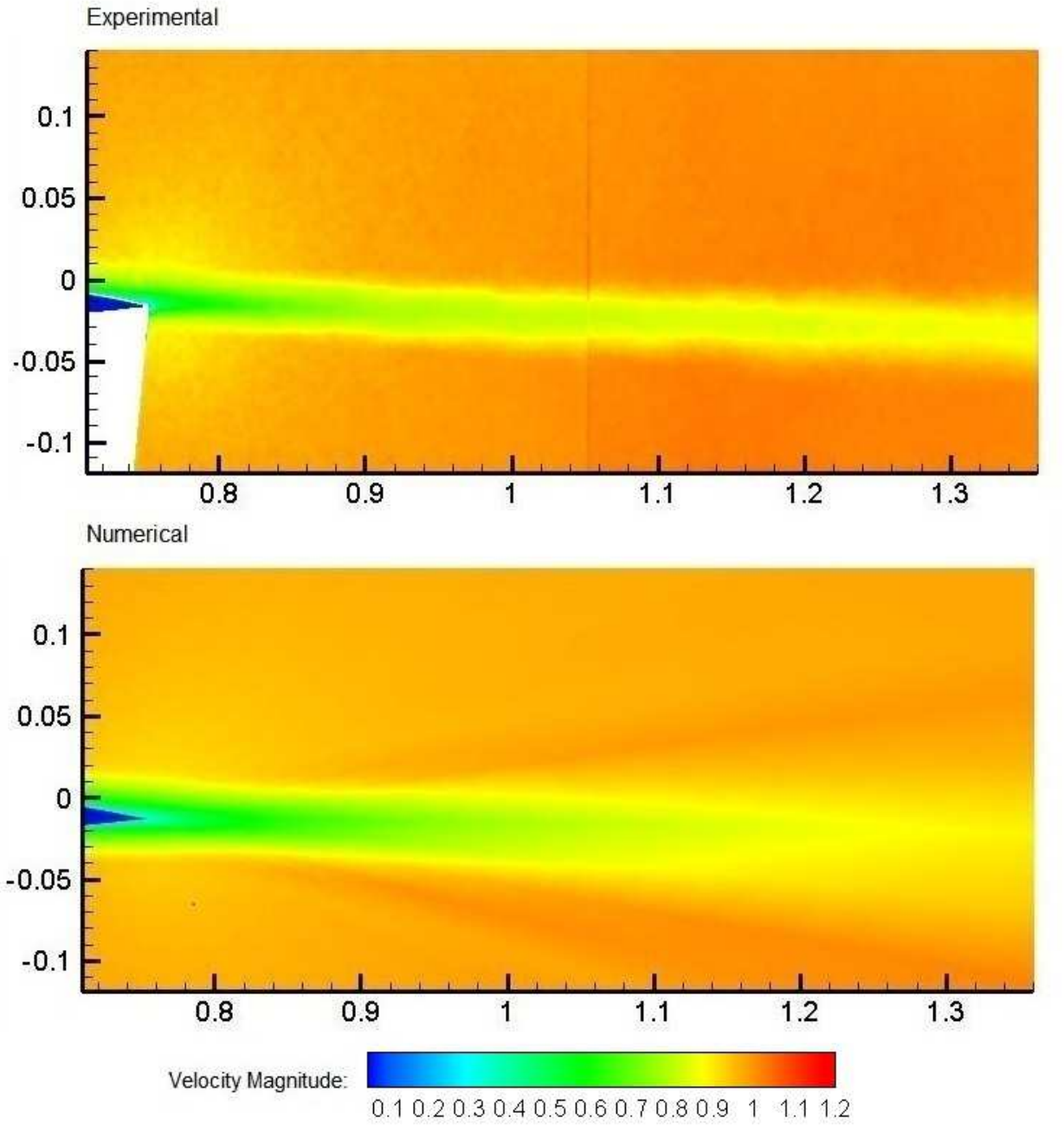


Figure 12: Contours of the adimensional velocity magnitude in the near wake of the oscillating airfoil at $\alpha_0 + 2°$ for $k = 0.2$.

Aerodynamics of the Coaxial Rotors of a rotary wing MAV: Numerical and Experimental Investigation

Davide Ferrari and Maurizio Boffadossi
Dipartimento di Ingegneria Aerospaziale - Politecnico di Milano,
Via La Masa 34, 20156 – Milano – Italy

Keywords: *Micro aerial vehicle, helicopter, coaxial rotor, blade element theory,*

Abstract

The purpose of the present work is to understand and to implement a possible way for modeling the counter rotating coaxial rotors of a rotary wing MAV. Micro aerial vehicle (MAV) is a remotely operated aircraft with a maximum linear dimension of approximately six-eight inches. MAV can be employed for several missions, both military and civil in potentially hostile environments where human life would be at risk. MAVs necessitate highly stable and controlled flight. Coaxial rotors configuration seem to be a good solution for small scale aircraft, especially because of their hover stability and compactness. This paper describes the design and the assemblage of a micro coaxial helicopter at Dipartimento di Ingegneria Aerospaziale. Numerical investigations are based on Blade Element Momentum Theory because of its cost effectiveness.. This is clearly true for large scale rotors but not enough demonstrated for the MAV scale rotors, where different aerodynamic phenomena are involved and predominant. Therefore experimental investigation was also performed, to validate theoretical approach and numerical results.

1 Introduction

A serious effort to design aircraft that are as small as possible for special, limited-duration missions began in the early 1990s. The concept of Micro Air Vehicles entered in the technical vocabulary in 1992, when for the first time the word MAV was used during a Workshop called “Future Technology – Driven Revolutions in Military Operation”.

The event was organized by DARPA, the government agency, within the US Department of Defense, whose mission is R&D in the field of new technologies for military scopes.

Defined at the beginning as micro-UAVs, MAVs have represented a further development of Unmanned Air Vehicles in the way of a progressive reduction of dimensions. After 1992, the interest in these devices increased, leading to a series of feasibility studies, conducted in particular by MIT and the US Naval Research Laboratory.

The positive feedback of these preliminary studies convinced DARPA to fund an ambitious project of several billion dollars to develop flying robots with a maximum size of 6 inches. It quickly became a challenge that involved teams of research within universities,

private companies and government laboratories not only in the US.

Since those years, thanks to the increasing availability of low cost and reliable electrical and mechanical micro components, especially from the RC model market, hundreds of configurations have been created. Some of these flew with success.

Many scientific papers have been published. In 1997, an annual MAVs competition was established in the US, hosted for the first time by the University of Florida. In 2000, it was given the name "International Micro Air Vehicle Competition" (IMAV).

Some years later, the European version was also born. IMAV and EMAV still continue, involving teams from many countries all over the world, and they are always followed by workshops and exhibitions to take stock of the state of the art.

MAVs are nowadays a growing reality. They carry visual, acoustic or chemical sensors and, for this reason, can be piloted without the eye-contact or can follow a predetermined path in autonomous navigation.

MAVs can be employed for several missions, both military and civil, such as: over-the-hill and around-the-corner reconnaissance in war context; indoor and outdoor video surveillance also in the dark; detection of chemical, bacteriological or radioactive agents; rescue operations; traffic monitoring; and information collection in urban settings.

The purpose of this work is to understand and to implement a possible way for modeling the counter rotating coaxial rotors of a rotary wing MAV.

Coaxial rotors have not found large application in the Helicopter Industry. Companies that have adopted them, like the Russian Kamov, are rare.

This is mainly due to the mechanical complexity they need in order to work, particularly while transmitting the motion from the motors to the blades.

Furthermore it is not a very efficient system, mainly because of the lower rotor aerodynamic losses.

However, coaxial rotors seem to be a good solution for small scale aircraft, especially because of their hover stability and compactness, due to the absence of the tail rotor and the presence of two main rotors which provide the lifting thrust, instead of only one.

2 Numerical method

In order to investigate coaxial rotor MAV *Blade Element Momentum Theory* was chosen because of its cost effectiveness: it has seemed fairly simple but thought to be able to provide good results according to experimental measures as demonstrated in ref. [1, 2].

This was clearly true for large scale rotors but not enough demonstrated for the MAV scale rotors where different aerodynamic phenomena are involved and low Reynolds number aerodynamics is predominant.

The BEMT is extensively illustrated by Leishman in [1] and it will be only summarized in the following. This theory combines the basic principles from both the blade element and the momentum approaches, considering the equivalence between the circulation and momentum theories of lift, and the conservation laws to an annulus of the rotor disk.

The incremental thrust on each annulus is computed using momentum theory with the assumption that successive rotor annuli have no mutual effects on each other and taking into consideration the inflow velocity v.

Let R the rotor radius R, $A=\pi R^2$ the rotor disk area and Ω the rotational speed of the rotor. The *non-dimensional induced velocity* can be defined as:

$$\lambda = \frac{v}{\Omega R}$$

Let $\bar{r}$ the non-dimensional radial distance along the blade, the corresponding *non-dimensional induced velocity* of the *Upper Rotor* can be written as in reference [3]:

$$\lambda_{UR} = \sqrt{\left(\frac{\sigma_{lUR} C_{L_\alpha}}{16F} - \frac{\lambda_c}{2}\right)^2 + \frac{\sigma_{lUR} C_{L_\alpha}}{8F}(\theta_{UR} - \alpha_{0L})\bar{r}} - \left(\frac{\sigma_{lUR} C_{L_\alpha}}{16F} - \frac{\lambda_c}{2}\right)$$

where:

$\lambda_c = V/_{R\Omega}$ is the non-dimensional induced velocity due to *climb velocity V;*

$\sigma_l(\bar{r}) = \frac{N_b c(\bar{r})}{\pi R}$ is the "local rotor solidity", based on the local blade chord *c*, and the number of blades, N_b ;

C_{L_α} is the lift curve slope;

$\theta = \theta(\bar{r})$ is the blade pitch angle;

F is the *Prandtl tip-loss function* [1]:

$$F = \left(\frac{2}{\pi}\right)\cos^{-1}(\exp(-f))$$

with

$$f = \frac{N_b}{2}\left(\frac{1-\bar{r}}{\lambda}\right)$$

F represents a *correction factor for the blade tip effects* and is an implicit function of λ itself.

Therefore the corrected local value $F(\bar{r})$ has to be determined by an iterative procedure, starting from *F*=1 (which is the case of an ideal rotor, with infinite number of blades N_b).

The BEMT flow model for the coaxial rotor assumes that the lower rotor operates partly in the fully developed slipstream ($r \leq r_c$) of the upper rotor. Consequently it is necessary to consider a suitable *radial contraction transformation* as illustrated by Leishman in ref. [3]. The non-dimensional induced velocity of the *Lower Rotor* can be written as:

$$\lambda_{LR} = \sqrt{\left(\frac{\sigma_{lLR} C_{L_\alpha}}{16F} - \frac{\lambda_c + (A/A_c)\lambda_{UR}}{2}\right)^2 + \frac{\sigma_{lLR} C_{L_\alpha}}{8F}(\theta_{LR} - \alpha_{0L})\bar{r}} - \left(\frac{\sigma_{lLR} C_{L_\alpha}}{16F} - \frac{\lambda_c + (A/A_c)\lambda_{UR}}{2}\right)$$

if $r \leq r_c$, and

$$\lambda_{LR} = \sqrt{\left(\frac{\sigma_{lLR} C_{L_\alpha}}{16F} - \frac{\lambda_c}{2}\right)^2 + \frac{\sigma_{lLR} C_{L_\alpha}}{8F}(\theta_{LR} - \alpha_{0L})\bar{r}} - \left(\frac{\sigma_{lLR} C_{L_\alpha}}{16F} - \frac{\lambda_c}{2}\right)$$

if $r > r_c$.

After computing the radial distribution of the non dimensional induced velocity for each rotor, the elementary contribution to thrust coefficient can be calculated:

$$\Delta C_T(\bar{r}_n) = \frac{\sigma_l(\bar{r}_n) C_{L_\alpha}}{2}\left(\theta(\bar{r}_n) - \frac{\lambda(\bar{r}_n)}{\bar{r}} - \alpha_{0L}\right)\bar{r}_n^{\,2}\Delta\bar{r}$$

A code based on BEMT was developed and fully validated with numerical results available in [2] and experimental data from Harrington [4] as it is shown in Fig. 1.

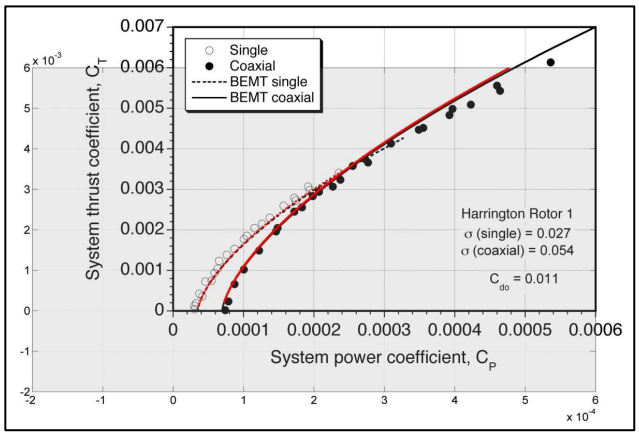


Fig. 1 Validation of BEMT code for single and coaxial rotors against computations [2] and measurements [4] of thrust and power for Harrington Rotor 1.

Red lines (– and ····) denote the present calculation.

3 MAV coaxial rotor

The application of BEMT to MAV rotors encounters different problems and difficulties.

In order to choose the most simple and reliable system for the MAV, an architecture without the necessity of changing both the collective and cyclic pitch for each rotor was investigated. It requires one engine for each rotor and a suitable control system.

Lifting thrust T_{OB} can be provided by differentially changing the two counter-rotating rotors angular velocities, Ω_{UR} and Ω_{LR}, with respect to the torques' equilibrium:

$$\rho A(\Omega_{UR}R)^2 C_{T_{UR}} + \rho A(\Omega_{LR}R)^2 C_{T_{LR}} = T_{OB}$$

$$\rho AR(\Omega_{UR}R)^2 C_{Q_{UR}} - \rho AR(\Omega_{LR}R)^2 C_{Q_{LR}} = 0$$

Using commercial available components by Esky "LamaV3 Co-Axial Helicopter" a flying prototype of a rotary wing MAV was assembled as illustrated in [5]. The final configuration is shown in Fig. 2.



Fig. 2 The micro helicopter assembled using RC commercial components. Rotor radius is 172 mm and the distance between the two rotors is 64 mm.

The actual geometry of the blade and the different airfoil sections were experimental measured, cutting progressively the blade from the tip and accurately digitalizing each section. The procedure is shown in Fig. 3.

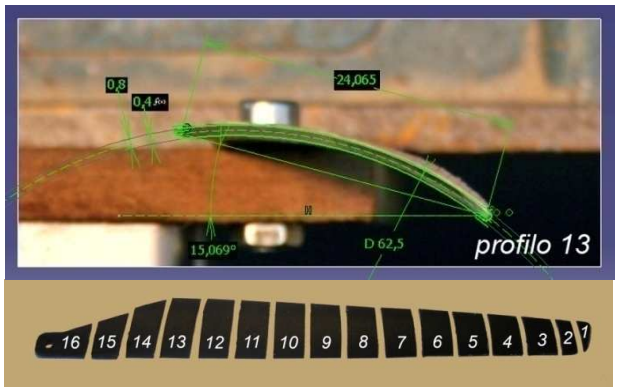


Fig. 3 Measurement of pitch angle and airfoil section along the blade.

The thickness ratio *TR* was found variable from 3.17 to 7.15 %, the camber ratio *CR* from 2.57 to 10.02 % and the pitch angle from 4.9° to 15.8°. It was established the possibility to accurately represent each section by a circular-arc airfoil. In this way the geometric characteristics of blade sections are similar to 2D airfoil computations of Bohorquez [6].

The lift curve (as shown in Fig. 4) are considerably not linear and strongly Reynolds number dependent. Consequently the convergence of λ distribution along the rotor radius was found very critical. The details of this numerical procedure are reported in [7].

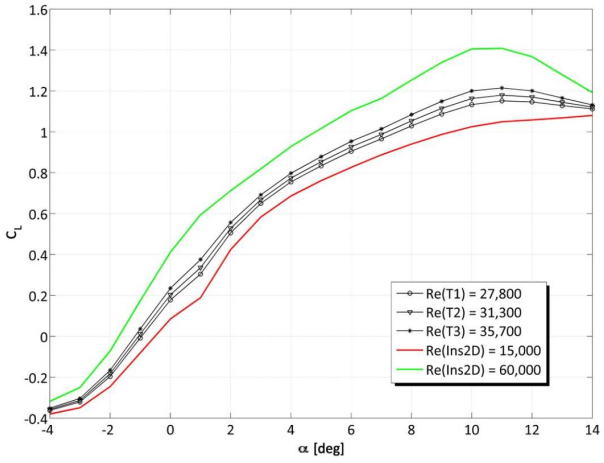


Fig. 4 Non linear lift curve of blade airfoil for different Reynolds numbers. Computation by Bohorquez [6]

4 Experimental investigation

Starting from the micro-sized coaxial model shown in Fig 2, an experimental apparatus was set to measure the static thrust and the velocity in the wake of small scale rotors, as illustrated in Fig. 5.

The pitch angle of each blade was carefully fixed to avoid uncertainties and to allow comparison with numerical computations.

Further details of instruments and experimental procedures can be find in [7].

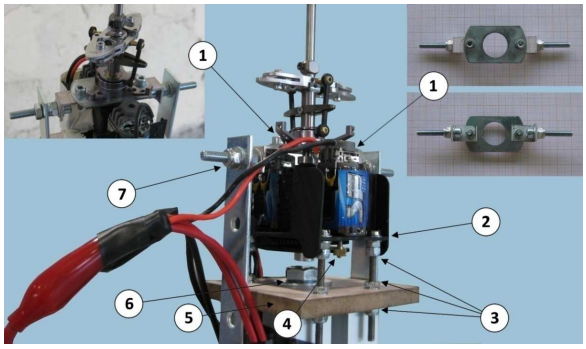


Fig. 5 The fixed-pitch model used in experimental investigation of coaxial rotor.

Three different test cases were considered to compare experimental and numerical data. Note that it was only possible to measure the total thrust produced by the coaxial rotor system, but using an optical tachometer the rotational velocity of each rotor was really detected and considered for the numerical computation.



Fig. 6 Measurement of velocity flowfield in the wake of the lower rotor.

Using a Schiltknecht MiniAir 60, vane anemometer (see Fig. 6) the velocity flowfield in the wakes of single and coaxial rotor configuration was also detected. The comparison between measures and numerical results are reported in Fig. 7 and Fig. 8.

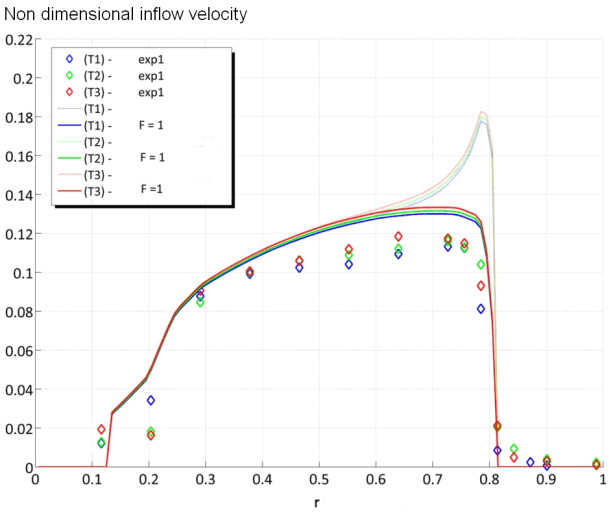


Fig. 7 Comparison between experimental induced velocity and BEMT computation of for a single (upper) rotor. The distance between the measurement plane and the rotor disk was 64 mm.

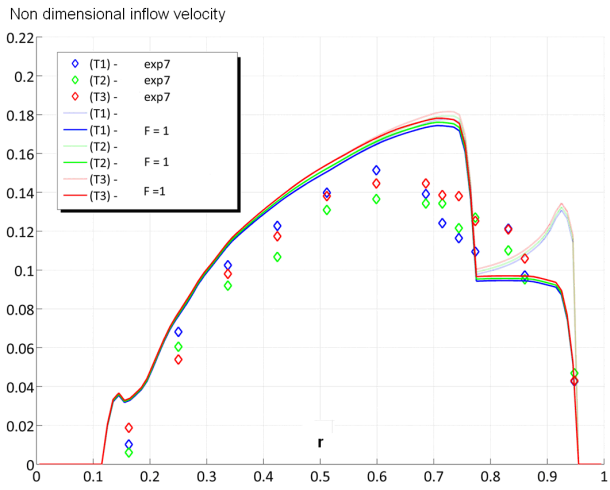


Fig. 8 Comparison between experimental induced velocity and BEMT computation for coaxial rotor configuration. The distance of the measurement plane from the lower rotor disk was 20 mm.

The induced inflow velocity is referred to rotor disk plane itself, while the measurements are provided with some distance from it. To better compare the two velocity distributions the BEMT computation consider the case with

Prandtl function F=1. The agreement is reasonable, both for single and coaxial configuration. Notice that the wake of helicopter rotor in hover condition is highly unsteady and three-dimensional [8].

The radial contraction of the upper rotor outflow is experimentally proved in Fig. 7 by the abrupt variation of the radial distribution of the induced velocity for $r=r_c=0.8$.

The measurements of Fig. 8 display that the inflow at the lower rotor is affected by the impinging wake (outflow) from the upper rotor with a more gradual spreading then BEMT computation, which is based on the ideal assumption of *vena contracta*.

Satisfactory prediction of thrust was obtained. The following table compares measured values and numerical results: the relative differences are of the order of few percent.

Test case		1	2	3
EXP measure	Ω_{UR} [rpm]	1600	1758	1967
	Ω_{LR} [rpm]	1562	1760	2086
	$T_{COAX_{EXP}}$ [N]	**1.77**	**2.26**	**2.94**
numerical data	T_{NUM} [N]	**1.702**	**2.135**	**2.864**
	Δ num-exp	***3.5%***	***5.3%***	***2.6%***
	T_{NUM} UR	107.5 [62%]	132.4 [61%]	169.9 [58%]
	T_{NUM} LR	66.1 [38%]	85.2 [39%]	122.2 [42%]
non dim. numerical data	C_T UR	0.0112 [61%]	0.0114 [61%]	0.0117 [61%]
	C_T LR	0.00719 [39%]	0.00731 [39%]	0.00746 [39%]
	C_T COAX	0.0183	0.0187	0.0191
	$C_Q - C_F$ UR	0.00165 [54%]	0.00166 [54%]	0.00166 [53%]
	$C_Q - C_F$ LR	0.00143 [46%]	0.00144 [46%]	0.00145 [47%]
	$C_Q - C_F$ COAX	0.00308	0.00309	0.00311
	Figure. of Merit	**0.52**	**0.53**	**0.55**

The experiments didn't simulate the real operative conditions of the MAV. In fact, during the flight, the control system manages the power to the engines with two different voltages to ensure the trim of the coaxial rotors. The preset trim curves are followed. While, during the tests, the two engines were fed with the same voltage.

Experimental conditions are described in detail in reference [7] where the full data are reported.

The last row of the table reports the computation of Figure of Merit (FM) as defined by Leishman for coaxial rotors in [3]. This is a non-dimensional efficiency parameter that provides a basis to conduct a relative comparison of rotor performance in hover. The FM considers "actual" power required to hover against the "ideal" power required to hover (calculated using the moment theory). The calculation of FM≈0.5 is a typical estimation for MAV rotors.

Blades of different geometries, materials and roughness were also considered. The experimental and numerical investigations of these alternative configurations can be found in reference [7].

In order to apply the BEMT code to more realistic condition a trim condition was also considered, balancing the torques produced by the two counter-rotating rotors.

The experimental data were obtained during the test of the flying MAV prototype shown in fig. 2. The MAV was equipped with a suitable flight control system as described by Tacca in reference [5].

The comparison between experimental data and numerical computation is reported in Fig. 9 for different thrust values. The agreement is satisfactory. Note that, in order to balance the torque generated by the upper rotor, the lower one has higher rotational speed.

The knowledge of these characteristic curves (thrust versus angular velocity) is a valuable tool for design the MAV flight control system.

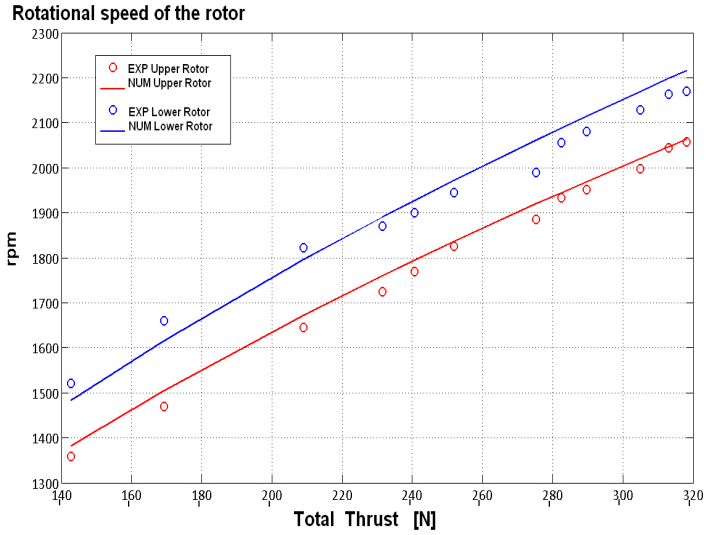


Fig. 9 Measurement of total thrust and rotational velocities of upper and lower rotors, compared with the numerical results in a trim condition.

In order to measure the actual wake contraction r_c and the diffusion of tip vortex, an experiment of flow visualization was included in the investigation.

Smoke was injected into the flow just above the tip rotor, as shown in Fig. 10. The image confirms the measurement of Fig. 7.

5 Conclusion

The aim of this work was to understand and to implement a possible way for modeling the counter rotating coaxial rotors of a rotary wing MAV.

A numerical code based on BEMT was developed and validated over available results of coaxial rotors. The code was improved to account for low Reynolds number aerodynamics.

Using available commercial RC components a prototype of coaxial rotor MAV system was assembled and tested.

An experimental apparatus was set to measure the static thrust and the velocity in the wake shed from the rotors, and the comparison between the measurements and the numerical forecast was found satisfactory.

The problem of the rotor trim was also investigated in order to provide the characteristic curves (thrust versus angular velocity) useful in the control of the MAV flight.

The code can be used to solve both the direct problem - predicting thrust of a given coaxial rotors geometry - and the inverse problem - design the optimum rotor.

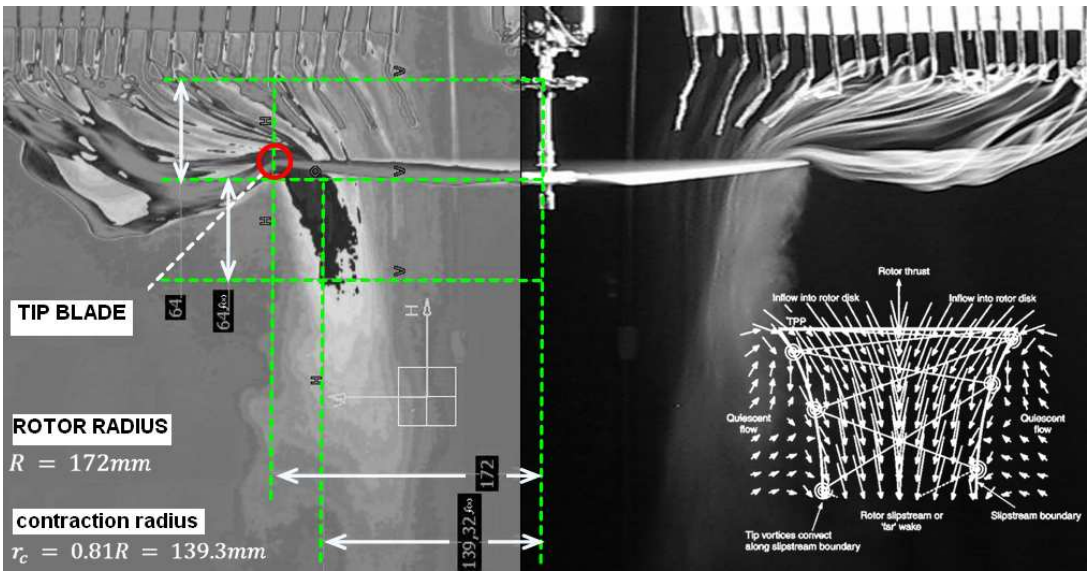


Fig. 10 Flow visualization image and measurement of the wake contraction of a single "upper" rotor, in a plane corresponding to lower rotor disk.

References

[1] Leishman, J.G., *Principles of Helicopter Aerodynamics*, Cambridge University Press, 2006.

[2] Syal, M., "Contributions to the Aerodynamic Optimization of a Coaxial Rotor System", Department of Aerospace Engineering, University of Maryland, 2008.

[3] Leishman, J.G. & Ananthan, S., "Aerodynamic Optimization of a Coaxial Proprotor", AHS International 62nd Annual Forum Proceedings, Phoenix AZ, 2006.

[4] Harrington, R. D., "Full-Scale Tunnel Investigation of the Static Thrust Performance of a Coaxial Helicopter Rotor", NACA Technical Note 2318, 1951.

[5] Tacca, S., "Progettazione e sviluppo di un MV ad ala rotante in configurazione alternativa", Tesi di Laurea in Ingegneria Aerospaziale Politecnico di Milano, 2008.

[6] Bohorquez, F., "Rotor Hover Performance and System Design of an Efficient Coaxial Rotary Wing Micro Air Vehicle", Department of Aerospace University of Maryland, 2007.

[7] Ferrari D., "Aerodinamica del rotore coassiale di un Mav ad ala rotante: indagine numerico sperimentale", Tesi di Laurea in Ingegneria Aerospaziale, Politecnico di Milano, 2009.

[8] Boffadossi M., Crosta G., "Experimental Investigation and Numerical Study of a Helicopter Rotor Model", *XX Congresso nazionale AIDAA*, Milano 2009.

CEAS 2011

Crack Propagation and Crashworthiness

3rd CEAS Air&Space Conference
21st AIDAA Congress

Hierarchical BEM for dynamic analysis of anisotropic 3-D cracked solids

A. Milazzo and I. Benedetti
Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale, Università di Palermo, Viale delle Scienze, 90128 Palermo - Italy

M. H. Aliabadi
Department of Aeronautics, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

Keywords: *Anisotropic crack elastodynamics, Fast BEM solvers*

Abstract

A hierarchical BEM solver for the analysis of three-dimensional anisotropic dynamic crack problems is presented. The solution of three-dimensional anisotropic dynamic problems by the boundary element method is computationally intensive and time consuming. In this framework, the use of the hierarchical matrices and iterative solvers represent an effective technique for speeding up the solution procedure and reducing the required memory storage. The results presented show the effectiveness of the approach and its accuracy.

1 Introduction

The complete characterization of a mechanical structure involves the solution of different problems, like the analysis of the static and dynamic behavior, the design of damage assessment and repair procedures, the design of non destructive inspection techniques, etc. In some cases, the inherent anisotropic behavior of the employed materials, and the subsequent complexity arising in the mathematical modelling, make the use of numerical methods mandatory. Today, the finite element method (FEM) is the most widely employed numerical method for structural analysis, due to its computational effectiveness related to the generation of banded symmetric resolving system, differently from the boundary element method (BEM), that, on the contrary, generates a fully populated and non-symmetric resolving system. However, there are fields of investigation where the application of the BEM is very appealing. Elastodynamics and Fracture Mechanics are the fields in which boundary element formulations have been established as an accurate and efficient tool for modelling and analysis [1, 2].

In the framework of dynamic fracture mechanics BEMs, a great deal of work has been done for isotropic solids in both 2*D* and 3*D* [1]. On the other hand a small number of boundary element papers has been dedicated to anisotropic materials. Actually, it is the lack of explicit closed-form fundamental solutions that makes the formulation and implementation of boundary element methods difficult for anisotropic materials. The most frequently used method to obtain the required Green's functions is the Fourier transform technique, which allows to obtain a closed form inverse transform when applied to the isotropic case whereas, for the anisotropic case, leads to a solution expressed in terms of infinite integrals. Non-explicit

fundamental solutions have been proposed but these are generally cumbersome to deal with in boundary elements implementations. Wang and Achenbach [3, 4] proposed an expression of the time-harmonic anisotropic fundamental solutions in terms of an integral defined over the surface of a unit radius sphere for 3*D* problems and over the circumference of a unit radius circle for 2*D* problems. This fundamental solution has been investigated in terms of its properties, numerical implementation and efficiency by different authors [5-9]. Different boundary element formulations have been presented for the analysis of 2*D* wave propagation and fracture mechanics in anisotropic solids. 3*D* BEM for anisotropic dynamics have been mainly proposed for transversely isotropic solids [5, 10, 11] whereas very few papers are available in the literature for general anisotropic media [12].

Although the BEM offers appealing modeling features, its extensive application is hindered by some limitations related to its non-symmetric and fully-populated resolving system which leads to intensive use of computational resources in terms of solution time and memory storage. For anisotropic dynamic problems these limitations are worsened by the poor efficiency in the influence matrix assembly, related to the need of numerically computing the fundamental solution kernels. The above mentioned issues become crucial when large scale models need to be analyzed for engineering purposes. Many investigations have been carried out to improve the computational efficiency of boundary element methods and different techniques have been proposed such as the fast multipole method [13], the panel clustering method [14], the mosaic-skeleton approximation [15] and the methods based on hierarchical matrices [16]. The FMMs and panel clustering tackle the problem from an analytical point of view and require the knowledge of some kernel expansion in advance to carry out the integration, proving very efficient for cases in which closed form expressions of the involved kernel are available, like in 3*D* isotropic elastodynamics [17]. On the other hand, hierarchical matrices provide purely algebraic tools for the approximation of boundary element matrices, thus resulting suitable for problems in which analytic closed-form expressions of the kernels are not available. Their application to 3*D* anisotropic fracture mechanics problems in the framework of the Dual Boundary Element Method has proven effective, showing reduction in the required computational time for both assembly and solution together with memory storage savings in case of large systems [18].

In the present paper, a fast BEM for three-dimensional anisotropic dynamic Fracture Mechanics, based on the hierarchical matrices and the Adaptive Cross Approximation (ACA) algorithm, is presented addressing the peculiarities and behaviour of the blocks approximation algorithm for the considered problem. Finally numerical experiments are performed to validate the approach and assess the advantages of the proposed computational strategy.

2 Time-harmonic Dual Boundary Element method

The modelling of fracture mechanics problems by the boundary element method can be accomplished through different formulations among which the Dual Boundary Element Method (DBEM) represents a very effective single-domain approach [1]. In the following, the involved integral equations are briefly reviewed for the case of time-harmonic loading, zero body forces and free crack surfaces.

2.1 Fundamental solutions

The problem fundamental solutions represent the elastic response of the unbounded domain subjected to time-harmonic unit point loads directed along the reference axes and applied at the source point $\boldsymbol{x}_0$. *F*or the 3*D* general anisotropic solid, it was derived by Wang and Achenbach [4]. They employed the Radon transform to reduce the governing equations to a one dimensional system of ordinary differential equations, which can be uncoupled by its eigenvalues and eigenvectors. After the solution of the transformed problem the subsequent

application of the inverse Radon transform allows to obtain the fundamental solution kernels which are given in terms of an integral over the surface of a unit sphere.

The time-harmonic fundamental solution kernels can be decomposed as the sum of a term depending from the position vector $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{x}_0$ and the frequency ω, which goes to zero as ω does, and a term depending on $\boldsymbol{r}$ only, which corresponds to the static anisotropic fundamental solution. Denoting by the superscript R and S the frequency dependent and the static part of the time-harmonic fundamental solution one has the following form of its kernels $\mathcal{K}_{ij(k)}$

$$\mathcal{K}^{*}_{ij(k)}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) = \mathcal{K}^{R}_{ij(k)}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) + \mathcal{K}^{S}_{ij(k)}(\boldsymbol{x}_0, \boldsymbol{x}) \tag{1}$$

The symbol $\mathcal{K}$ with two subscripts stands for the fundamental solution displacements U and tractions T, needed for the displacement boundary integral representation. The same symbol $\mathcal{K}$ with three subscripts stands for the derivative kernels D and S involved in the stress and tractions boundary integral representation [1]. The frequency independent part of the fundamental solution kernels U^{S}_{ij}, T^{S}_{ij}, D^{S}_{ijk} and S^{S}_{ijk}, which corresponds to the static fundamental solution [4], presents singularities of order $|r|^{-1}$, $|r|^{-2}$, $|r|^{-2}$, $|r|^{-3}$, respectively. Their expressions can be found in reference [18]. The frequency dependent part of the displacements kernel is given by

$$U^{R}_{ij}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) = \frac{\mathrm{i}}{4\pi^2} \int_{\substack{|\boldsymbol{q}|=1 \\ \boldsymbol{q}\cdot\boldsymbol{r}>0}} \sum_{m=1}^{3} \frac{\kappa_m E_{im} E_{jm}}{2\rho c_m^2} \, e^{\mathrm{i}\kappa_m |\boldsymbol{q}\cdot\boldsymbol{r}|} \, dS(\boldsymbol{q}) \tag{2}$$

where ρ is the material density, $\boldsymbol{q}$ is the external unit normal to the integration half a unit sphere centered at the field point $\boldsymbol{x}$ and extending on the half-space $\boldsymbol{q}\cdot\boldsymbol{r} > 0$. The other terms appearing in Eq. (2) are defined in terms of the eigenvalues λ_m and eigenvectors E_{lm} of the Christoffel matrix $\Lambda_{ij}(q) = C_{irjs} q_r q_s$, where C_{irjs} are the components of the elasticity tensor. Accordingly, one defines the phase velocity $c_m = \sqrt{\lambda_m/\rho}$ and the corresponding wave number $\kappa_m = \omega/c_m$. The expression of the frequency dependent contribution of the traction kernel is given by

$$T^{R}_{ij}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) = -\frac{C_{jsvz}}{4\pi^2} n_s \int_{\substack{|\boldsymbol{q}|=1 \\ \boldsymbol{q}\cdot\boldsymbol{r}>0}} \sum_{m=1}^{3} q_z \frac{\kappa_m^2 E_{vm} E_{im}}{2\rho c_m^2} \, e^{\mathrm{i}\kappa_m |\boldsymbol{q}\cdot\boldsymbol{r}|} \, dS(\boldsymbol{q}) \tag{3}$$

For the derivative kernels one has the following expression of the frequency dependent part

$$D^{R}_{ijk}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) = \frac{C_{ijvz}}{4\pi^2} \int_{\substack{|\boldsymbol{q}|=1 \\ \boldsymbol{q}\cdot\boldsymbol{r}>0}} \sum_{m=1}^{3} q_z \frac{\kappa_m^2 E_{vm} E_{km}}{2\rho c_m^2} \, e^{\mathrm{i}\kappa_m |\boldsymbol{q}\cdot\boldsymbol{r}|} \, dS(\boldsymbol{q}) \tag{4}$$

$$S^{R}_{ijk}(\boldsymbol{x}_0, \boldsymbol{x}, \omega) = \frac{\mathrm{i}\, C_{ijgh}\, C_{ksvz}}{4\pi^2} \, n_s \left\{ \int_{\substack{|\boldsymbol{q}|=1 \\ \boldsymbol{q}\cdot\boldsymbol{r}>0}} \sum_{m=1}^{3} q_h q_z \frac{\kappa_m^3 E_{vm} E_{sm}}{2\rho c_m^2} \, e^{\mathrm{i}\kappa_m |\boldsymbol{q}\cdot\boldsymbol{r}|} \, dS(\boldsymbol{q}) - \frac{\mathrm{i}}{|\boldsymbol{r}|} \int_{\substack{|\boldsymbol{q}|=1 \\ \boldsymbol{q}\cdot\boldsymbol{r}=0}} \sum_{m=1}^{3} q_h q_z \frac{\kappa_m^2 E_{vm} E_{sm}}{\rho c_m^2} \, dL(\boldsymbol{q}) \right\} \tag{5}$$

where the last integral is performed on the unit radius circle centered at the field point $\boldsymbol{x}$, lying in the plane $\boldsymbol{q}\cdot\boldsymbol{r} = 0$. It should be noted that the terms U^{R}_{ij}, T^{R}_{ij}, D^{R}_{ijk} exhibit regular behavior with respect to $|\boldsymbol{r}|$, whereas the kernel S^{R}_{ijk} presents a weak singularity of order $|r|^{-1}$.

The computation of the fundamental solution requires to perform numerically the integrations appearing in both the static, see reference [18], and the frequency dependent parts of the kernels, see Eqs. (7), (11), (12) and (13). The integration scheme proposed in reference [4] and based on Gauss quadrature formulas appears to be the more efficient alternative for general anisotropic media. It should be noted

that, in order to obtain the desired accuracy, the order of the fundamental solution integration formulas needs to be accurately set, as the complex exponential term exhibits highly oscillatory behavior for large values of $|\boldsymbol{r}|$. This makes the computation of the fundamental solution kernels a computationally intensive and time-consuming task.

2.2 Boundary integral equations and numerical model

In the DBEM the needed boundary integral equations are obtained by: i) collocating the boundary integral representation for the displacements on the external boundary of the body and on one of the crack surfaces, ii) collocating the boundary integral representation for the tractions on the other crack surfaces [1].

Accounting for the decomposition of the time-harmonic fundamental solution into static and frequency dependent contributions, the displacement boundary integral equation for a point $\boldsymbol{x}_0$ belonging to the external boundary of the body subjected to a load with circular frequency ω is written as [4]

$$c_{ij}\, u_j(\boldsymbol{x_0}) + \cancel{\int}_\Gamma T_{ij}^S\, u_j\, d\Gamma - \int_\Gamma U_{ij}^S\, t_j\, d\Gamma + \int_\Gamma T_{ij}^R\, u_j\, d\Gamma - \int_\Gamma U_{ij}^R\, t_j\, d\Gamma = 0 \tag{6}$$

where c_{ij} are coefficients depending on the boundary geometry and computed through rigid body considerations and the symbol $\cancel{\int}$ stands for Cauchy principal value integral due to the $|\boldsymbol{r}|^{-2}$ singularity of the of the T_{ij}^S kernel.

For a point belonging to one of the crack surfaces, namely the lower surface, the displacement integral equation specializes as

$$c_{ij}^-\, u_j(\boldsymbol{x}_0^-) + c_{ij}^+\, u_j(\boldsymbol{x}_0^+) + \cancel{\int}_\Gamma T_{ij}^S\, u_j\, d\Gamma - \int_\Gamma U_{ij}^S\, t_j\, d\Gamma + \int_\Gamma T_{ij}^R\, u_j\, d\Gamma - \int_\Gamma U_{ij}^R\, t_j\, d\Gamma = 0 \tag{7}$$

where $\boldsymbol{x}_0^-$ and $\boldsymbol{x}_0^+$ are the two coincident crack points belonging to the lower and upper crack surfaces. For smooth crack surfaces at $\boldsymbol{x}_0^-$ it holds $c_{ij}^- = c_{ij}^+ = 0.5\,\delta_{ij}$ where δ_{ij} is the Kronecker delta.

The traction integral equation collocated at a point $\boldsymbol{x}_0^+$ belonging to the other crack surface, namely the upper one, where strain continuity is assumed, is written as

$$\frac{1}{2}t_j(\boldsymbol{x}_0^+) - \frac{1}{2}t_j(\boldsymbol{x}_0^-) + \cancel{\cancel{\int}}_\Gamma S_{ijk}^S\, n_j\, u_k\, d\Gamma - \cancel{\int}_\Gamma D_{ijk}^S\, n_j\, t_k\, d\Gamma + \int_\Gamma S_{ijk}^R\, n_j\, u_k\, d\Gamma - \int_\Gamma D_{ijk}^R\, n_j\, t_k\, d\Gamma = 0 \tag{8}$$

where n_j are the components of the outward normal at the collocation point, again the symbol $\cancel{\int}$ denotes the Cauchy principal value integral associated with the $|\boldsymbol{r}|^{-2}$ singularity of the of the D_{ijk}^S kernel and $\cancel{\cancel{\int}}$ stands for Hadamard principal value integral, originating from the presence of the $|\boldsymbol{r}|^{-3}$ singularity of the of the S_{ijk}^S kernel.

An important application of the time-harmonic boundary element formulation is wave scattering. Denoting by u_j^{in} and t_j^{in} the displacement and traction components of the known incident field and assuming that the crack boundaries Γ are traction free, Eqs. (6), (7) and (8) specializes as

$$c_{ij}\, u_j(\boldsymbol{x}_0) + \cancel{\int}_\Gamma T_{ij}^S\, u_j\, d\Gamma + \int_\Gamma T_{ij}^R\, u_j\, d\Gamma = u_j^{in}(\boldsymbol{x_0}) \tag{9}$$

$$c_{ij}^-\, u_j(\boldsymbol{x}_0^-) + c_{ij}^+\, u_j(\boldsymbol{x}_0^+) + \cancel{\int}_\Gamma T_{ij}^S\, u_j\, d\Gamma + \int_\Gamma T_{ij}^R\, u_j\, d\Gamma - \int_\Gamma U_{ij}^R\, t_j\, d\Gamma = u_j^{in}(\boldsymbol{x}_0^-) \tag{10}$$

$$\frac{1}{2}t_j(\boldsymbol{x}_0^+) - \frac{1}{2}t_j(\boldsymbol{x}_0^-) + \cancel{\cancel{\int}}_\Gamma S_{ijk}^S\, n_j\, u_k\, d\Gamma + + \int_\Gamma S_{ijk}^R\, n_j\, u_k\, d\Gamma = t_j^{in}(\boldsymbol{x}_0^+) \tag{11}$$

For the numerical solution of the boundary integral equation model, the external boundary and the crack surfaces are sub-divided into a set

of boundary elements over which the displacements and the tractions, as well as the geometry, are expressed by means of suitable shape functions and nodal values. Due to the presence of singularities in the boundary integral equations kernels, the boundary elements employed in the discretization have to fulfill the conditions for the existence of the singular integrals of Eqs. (6), (7) and (8). In the present work, the external boundary of the body is modeled by using eight-node super-parametric constant elements with central collocation node, whereas iso-parametric discontinuous eight-node quadratic elements are used for modeling both the crack surfaces. Once the discretized forms of Eqs. (6), (7) and (8) are obtained, they are collocated at the pertaining nodes introduced by the discretization process. For the computation of the resolving system coefficients, the regular integrals are evaluated by standard Gaussian quadrature. Suitable techniques are employed for weakly-singular, strongly-singular and hyper-singular integrations [19]. Finally, the application of the boundary conditions leads to a resolving system of linear equations of the form

$$\boldsymbol{A}\,\boldsymbol{X} = \boldsymbol{Y} \qquad (12)$$

where the vector $\boldsymbol{X}$ contains the unknown boundary displacements or tractions and the crack unknown displacements and the right-hand-side $\boldsymbol{Y}$ is obtained considering the prescribed displacements and tractions.

3 Hierarchical dual boundary element method for anisotropic time-harmonic elastodynamics

In this section the strategy for the fast hierarchical boundary element solution is described.

3.1 Basic principles

Integral operators can be efficiently approximated by using separable expansions of the corresponding kernels when the condition of *asymptotic smoothness* is met [16]. Actually, in the collocation of the boundary integral equations, i.e Eqs (6), (7) and (8), at the node $\boldsymbol{x}_0$, considering a suitable collection Ω_e of contiguous boundary elements which are sufficiently far from the collocation point according to a set geometric criterion, the kernel approximation renders into a *low-rank approximation* of some blocks of the original matrix, corresponding to collocation over points close to $\boldsymbol{x}_0$ and elements belonging to Ω_e. Grouping the collocation points and the sets Ω_e in a hierarchical way leads to schemes based on low-rank approximations of blocks of the matrix, the so-called hierarchical matrix scheme, which allows to build a data-sparse approximation $\boldsymbol{\mathcal{A}}$ of the DBEM resolving matrix $\boldsymbol{A}$ given in Eq. (12) [16, 20].

Therefore, the hierarchical representation $\boldsymbol{\mathcal{A}}$ of a boundary element matrix consists of a collection of hierarchically ordered matrix *blocks*. Some blocks, corresponding to sets of collocation nodes and integration elements which are sufficiently far apart from each other, admit a special compressed representation.

The subdivision of the matrix into blocks is based on a *hierarchical partition* of the mesh nodes aimed at grouping subsets of nodes corresponding to contiguous collocation nodes and integration elements, on the basis of a geometrical criterion. This partition has to take into account that, in the problem under consideration, different boundaries (external and cracks) and different integral operators are involved (displacements and tractions integral equation) [18]. The partition is stored in a tree structure of index subsets, called *cluster tree*.

The cluster tree is used to form pairs of clusters recursively so to define the *block tree*, which provide the hierarchical block subdivision of the matrix. The blocks generated from the hierarchical partition can be classified into: i) *non-admissible blocks*, which need to be represented entirely; ii) *admissible blocks*, which admit a compressed low-rank approximation. The admissibility of a *block* is checked against the following geometrical condition [20]

$$\min(diam\,\Omega_{\boldsymbol{x}_0}, diam\,\Omega_{\boldsymbol{x}}) \leq \eta \cdot dist(\Omega_{\boldsymbol{x}_0}, \Omega_{\boldsymbol{x}}) \qquad (13)$$

where Ω_{x_0} denotes the cluster of collocation nodes, Ω_x the cluster of elements over which the integration is carried out to compute the block coefficients and $\eta > 0$ is a suitable parameter, which controls the trade-off between the number of admissible blocks and the the quality of the approximation. The low-rank representation of an admissible $m \times n$ block $\boldsymbol{C}$ is given by

$$\boldsymbol{C} \approx \boldsymbol{C}_k = \boldsymbol{V} \cdot \boldsymbol{W}^{\mathrm{T}} = \sum_{i=1}^{k} \boldsymbol{v}_i \cdot \boldsymbol{w}_i^T \qquad (14)$$

where the matrices $\boldsymbol{V}$ and $\boldsymbol{W}$ are of order $m \times k$ and $n \times k$, respectively, being k the *rank* of the new representation. The approximating block $\boldsymbol{C}_k$ satisfies the relation $|\boldsymbol{C} - \boldsymbol{C}_k|_2 \leq \varepsilon_c |\boldsymbol{C}|_2$, where $|\cdot|_2$ is the *2-norm* of a complex matrix and ε_c is the accuracy required to the approximation. The low-rank approximation of admissible blocks is computed by the Adaptive Cross Approximation algorithm (ACA), which appears to be very efficient [21]. The low-rank representation of the admissible blocks provides remarkable computation and storage memory savings, because only few entries of the original matrix blocks need to be computed and stored. Moreover, the hierarchical representation of the resolving BEM system accelerates the matrix-vector product, which is the core operation when iterative solvers are used. This provides significant reduction in the solution time.

3.2 Fast DBEM for time-harmonic anisoropic elastodynamics

In the present paper, the computational scheme employed for the hierarchical representation of the resolving system matrix $\boldsymbol{A}$ is based on the discretization scheme presented in section 2.2.

The binary *cluster tree* is built by bisection, and the hierarchical representation of the matrix is stored in the associated quaternary *block tree*. The block admissibility condition is implemented by using *bounding boxes* for the node and element clusters [20]. Finally, the admissible blocks are approximated employing the Adaptive Cross Approximation algorithm that allows to compute adaptively the approximation, according to an intrinsic stopping criterion related to the pre-set required accuracy ε_c [21]. Indeed, the ACA algorithm computes each low-rank block by using only some of the entries of the original block, which are obtained by integrating the fundamental solution kernels. These are computed according to the scheme proposed by Wang and Achenbach [4], where the integration domain, given by the surface of a unit sphere is mapped on the domain $[0,1] \times [0,2\pi]$, over which the standard Gauss quadrature of order $M \times N$ is applied. The low-rank representation obtained through ACA is further optimised by a recompression procedure based on the reduced Singular Value Decomposition (SVD). Once the basic hierarchical representation of the collocation matrix has been set up, the system solution is efficiently computed by using a preconditioned *GMRES* iterative solver. An *LU* hierarchical preconditioner, built from a coarse approximation with accuracy ε_p of the original collocation matrix, is used. Further details on the computational scheme can be found in [18].

When the fundamental solutions are not available in a closed analytical form, the application of ACA to the approximation of low-rank blocks can be affected by the computational scheme employed for their calculation. As the complex exponential term appearing in the time harmonic fundamental solutions exhibits oscillatory behaviour and the quadrature scheme employed for their computation influences the accuracy, this topic deserves particular attention in the present approach. Figures 1 and 2 show some representative results for the ACA approximation of low rank blocks, stemming from the different kernels involved in the DBEM, with respect to the accuracy of the fundamental solution for different non-dimensional circular frequencies $\overline{\omega} = \omega L/c$, where c and L are reference phase velocity and length, respectively, taken as the corresponding maximum in the considered problem. Results in terms of block rank and memory storage percentage with respect to full rank are given for two blocks, stemming from the clusters of nodes

and elements shown in the figures; the material employed for the computation is graphite-epoxy [22]. The investigated frequencies are chosen in the range for which the considered mesh can provide accurate representation of the wave form. The convergence study is carried out in terms of block rank and memory ratio (low-rank memory storage over full rank memory storage). Results similar to those shown in Figs. 1 and 2 have been founds for different anisotropic materials. The analysis shows that the block approximation convergence threshold depends on: i) the circular frequency, as the higher the frequency the higher the convergence rank and memory storage; ii) the accuracy required to ACA, as the higher the accuracy the higher the convergence rank and memory storage. Additionaly, the analyses performed for different materials (not given here for the sake of coinciseness) show the influence of the degree of anisotropy, as the convergence rank and memory storage grows with respect to it. The presented study refers to the single block, but similar behaviour was found for each low-rank block and then for the overall hierarchical matrix. In conclusion, the accuracy in the numerical computation of the fundamental solution kernels involved represents a crucial issue, as it affects the low-rank approximation and consequently the overall performance of the approach in terms of computational time and memory storage. The effects of the admissibility parameter η and of the *cluster tree* cardinality have been also investigated and the results obtained are similar to those presented and discussed in ref. [18].

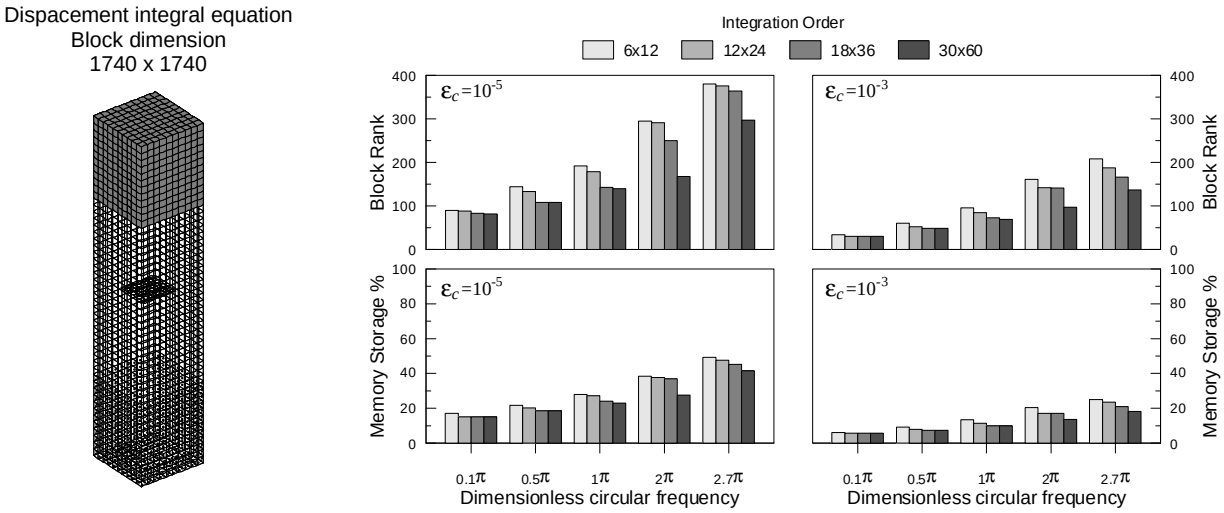


Fig. 1 Low rank blocks performance analysis for displacement boundary integral operator

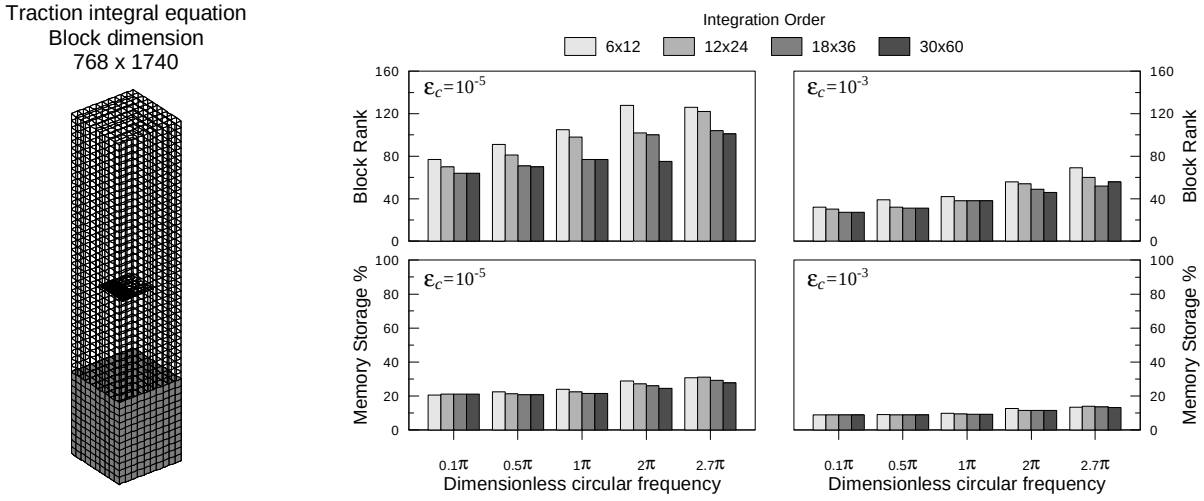


Fig. 2 Low rank blocks performance analysis for traction boundary integral operator

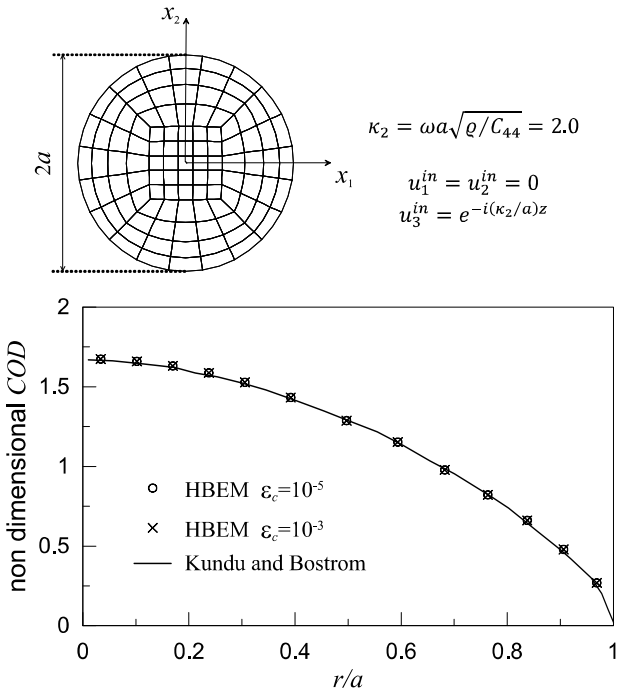


Fig. 3 Crack opening displacement for a penny-crack in graphite/epoxy unbounded domain

4 Validation and numerical performances

To validate the proposed hierarchical time-harmonic BEM and to check its accuracy, the solution for the scattering of an incident longitudinal wave by a penny-shaped crack embedded in a graphite-epoxy unbounded domain was computed and compared with the solution presented by Kundu and Bostrom [22], to which the reader is referred for the details on the problem definition. The results obtained are shown in Fig. 3 and they show the accuracy of the present solution for different values of the ε_c parameter. The results are in very good agreement with the anlytical solution showing the reliability of the proposed approach.

The numerical performances of the method are studied by considering the bar shown in Fig. 4 for which both the uncracked and cracked configuration have been investigated. Graphite/epoxy material properties have been considered and geometry parameters have been set as $L = 1m$, $W = H = 0.2m$ and $a = 0.1m$. The exicitation frequency has been chosen as $\overline{\omega} = 0.4\,\pi$. For both the uncracked and cracked configuration three meshes with increasing number of degrees of freedom have been analyzed. The data of the meshes employed in the computations, the used HBEM parameters and the corresponding hierarchical matrix features are given in Table 1. The method performances are assessed in terms of time reduction, both in assembling and solution, and of memory savings. The following performance indices are then introduced: i) assembly speed up, defined as the ratio between the assembly time of hierarchical BEM and that of standard collocation BEM; ii) solution speed up, defined as the ratio between the solution time of hierarchical BEM with iterative preconditioned GMRES and that of the standard Gaussian elimination applied to the standard collocation BEM; iii) memory storage, defined as the percentage of the storage memory needed by hierarchical BEM with respect to that required to store the standard collocation BEM matrix.

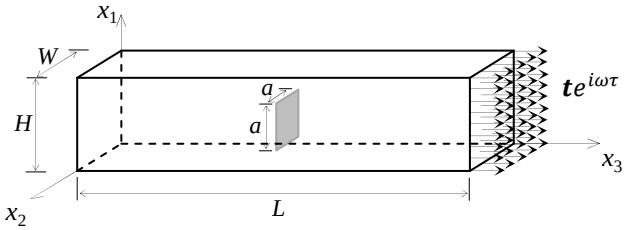


Fig. 4 Clamped bar geometrical and load scheme

	Uncracked Bar			Cracked Bar		
	MESH A	*MESH B*	*MESH C*	*MESH A*	*MESH B*	*MESH C*
Number of elements	2200	4950	8800	2232	5022	9824
Degrees of freedom	6600	14850	26400	7368	16578	29472
Cluster tree cardinality	72	72	72	72	72	72
Admissibility parameter	2.0	2.0	2.0	2.0	2.0	2.0
Number of blocks	604	2965	7252	652	3109	7792
ACA matrix %	53.7	69.5	80.9	43.1	65.8	75.1
LU preconditioner accuracy	10^{-1}	10^{-1}	10^{-1}	10^{-1}	10^{-1}	10^{-1}

Table 1 Mesh characteristics and HBEM parameters

Fig. 5 shows the results for the HBEM performance indices obtained for the each of the three meshes used to analyze the uncracked and cracked bar considering two values of the ACA accuracy parameter ε_c. It is observed that the HBEM reduces both assembly and solution time as well as memory storage for both the considered configurations. Its efficiency improves with the problem size, namely the number of degrees of freedom. Obviously the observed reduction is greater for lower accuracies, when ACA uses a lower number of the collocation matrix coefficients to approximate the low-rank blocks. It is worth to note that the efficiency of the method appears better for the uncracked configuration. This is related to the use in the uncracked case of constant elements only. Actually the crack discretization scheme employs eight node discontinuous elements, which can affect the hierarchical matrix construction efficiency because some nodes of an element can belong to different blocks. Here results are given for one value of the circular frequency and for one material, namely graphite-epoxy. However, similar results and trends have also been obtained for other values of the circular frequency and for other materials with different degree of anisotropy. It is observed that the excitation frequency slightly influence the method performances. In particular, as the excitation frequency grows the HBEM performance indexes grows as a consequence of the ACA behavior. The same considerations hold for the influence of the material anisotropy.

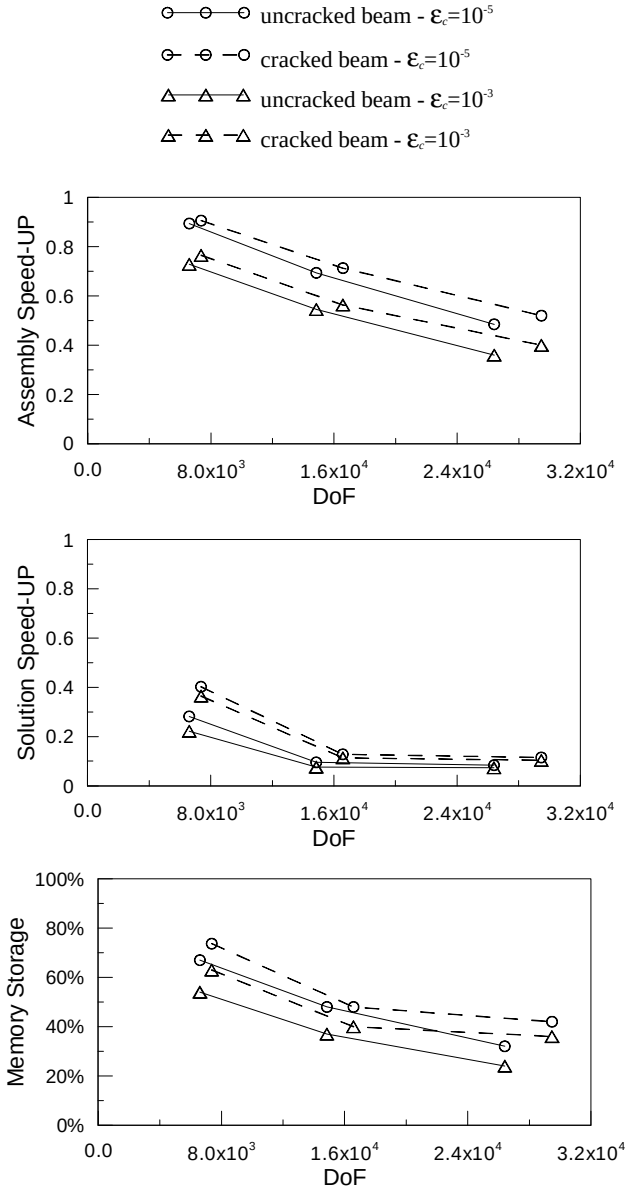


Fig. 5 Hierarchical BEM performance indexes

5 Conclusion

In the present work a hierarchical BEM solver for three-dimensional anisotropic dynamic crack problems has been developed.

It has been shown that the hierarchical format can be used, in conjunction with an iterative solver, to speed up the solution of this class of problems. The effect of the accuracy of the numerical scheme chosen for the computation of the fundamental solutions has been investigated. The effect of various parameters on the method performances has been also analyzed. It has been found out that the performance of the technique slightly depends on the frequency analyzed and on the degree of anisotropy of the considered material, in addition to the hierarchical preset accuracy.

In conclusion, the technique allows to remarkably reduce the computational requirements of the analysis, in terms of time and memory storage, preserving the accuracy of the solution for both uncracked and cracked configurations.

References

[1] Aliabadi M.H, “Boundary element formulations in fracture mechanics”,

Applied Mechanics Review, Vol. 50, No 2, 1997, pp. 83-96.

[2] Manolis G.D. and Beskos D.E., *Boundary Element Methods in Elastodynamics*, Unwin-Hyman, Berlin, 1988.

[3] Wang C.Y. and Achenbach J.D, "Elastodynamic fundamental solution for anisotropic solids", *Geophysical Journal International*, Vol. 118, No 2 , 1994, pp. 384-392.

[4] Wang C.Y. and Achenbach J.D, "Three-Dimensional Time-Harmonic Elasto-dynamic Green's Functions for Anisotropic Solids", *Proceedings of the Royal Society A*, Vol. 449, No 1937, 1995, pp. 441-458.

[5] Sáez A. and Domínguez J., "BEM analysis of wave scattering in transversely isotropic solids", *International Journal for Numerical Methods in Engineering*, Vol. 44, No 9, 1999, pp. 1283-1300.

[6] Dravinski, M. and Zheng T., "Numerical evaluation of three-dimensional time-harmonic Green's functions for a nonisotropic full-space" *Wave Motion*, Vol. 32, No 2, 2000, pp. 141-151.

[7] Dravinski M. and Niu Y., "Three-dimensional time-harmonic Green's functions for a triclinic full-space using a symbolic computation system", *International Journal for Numerical Methods in Engineering*, Vol. 53, No 2, 2002, pp.445-472.

[8] Chen Z.and Dravinski M., "Numerical evaluation of harmonic Green's functions for triclinic half-space with embedded source. Part I: a 2D model", *International Journal for Numerical Methods in Engineering*, Vol. 69, No 2, 2007, pp.:347-366.

[9] Chen Z.and Dravinski M., "Numerical evaluation of harmonic Green's functions for triclinic half-space with embedded sources—Part II: a 3D model", *International Journal for Numerical Methods in Engineering*, Vol. 69, No 2, 2007, pp.367-389.

[10] Sáez A. and Domínguez J. "Dynamic crack problems in three-dimensional transversely isotropic solids" *Engineering Analysis with Boundary Elements*, Vol. 25, No 3, 2001, pp. 203-210.

[11] Ariza M.P. and Dominguez J. "Dynamic BE analysis of 3-D cracks in transversely isotropic solids", *Computer Methods in Applied Mechanics and Engineering*, Vol. 193, No 9-11, 2004, pp. 765-779.

[12] Niu Y. and Dravinski M. "Three-dimensional BEM for scattering of elastic waves in general anisotropic media". *International Journal for Numerical Methods in Engineering*, Vol. 58, No 7, 2003, pp. 979-998.

[13] Popov V. and Power H., "An O(N) Taylor series multipole boundary element method for three-dimensional elasticity problems", *Engineering Analysis with Boundary Elements*, Vol. 25, No 1, 2001, pp.7-18.

[14] Hackbusch W. and Nowak Z.P., "On the fast matrix multiplication in the boundary element method by panel clustering", *Numerische Mathematik*, Vol. 54, No 4, 1989, pp. 463-491.

[15] Tyrtyshnikov E. "Mosaic Skeleton approximations", Calcolo, Vol. 33, No 1, 1996, pp. 47-57.

[16] Hackbusch W.A "Sparse Matrix Arithmetic Based on H-Matrices. Part I: Introduction to H-Matrices", *Computing*, Vol. 62, No 2, 1999, pp. 89-108.

[17] Nishimura N., "Fast multipole accelerated boundary integral equation methods", *Applied Mechanics Reviews*, Vol. 55, No 4, 2002, pp. 299-324.

[18] Benedetti I., Milazzo A. and Aliabadi M.H., "A fast dual boundary element method for 3D anisotropic crack problems", *International Journal for Numerical Methods in Engineering*, Vol. 80, No 10, 2009, pp.1356--1378.

[19] Wrobel L.C. and Aliabadi M. H., *The Boundary Element method*, Wiley New Jersey, 2002.

[20] Borm S., Grasedyck L. and Hackbusch W., "Introduction to hierarchical matrices with applications", *Engineering Analysis with Boundary Elements*, Vol. 27, No 18, 2003, pp. 405-422.

[21] Bebendorf M. and Grzhibovskis R. "Accelerating Galerkin BEM for linear elasticity using Adaptive Cross Approximation", *Mathematical Methods in the Applied Sciences*,Vol 29, No 14, 2006, pp. 1721-1747

[22] Kundu T. and Boström A., "Axisymmetric Scattering of a Plane Longitudinal Wave by a Circular Crack in a Transversely Isotropic Solid", *Journal of Applied Mechanics*, Vol. 58, No 3, 1991, pp. 695-702.

Experimental and numerical correlation of an helicopter subfloor crashworthiness

M. Anghileri*, A. Milanese, A.Surini***, M. Nicolotti*****
Politecnico di Milano, Department of Aerospace Engineering,
Via La Masa 34, 20156, Milan, Italy

F. Invernizzi[+]
AgustaWestland, Via G. Agusta 520, Cascina Costa di Samarate, Varese, Italy

Keywords: *crashworthiness, helicopter subfloor, intersections, finite element model, experimental tests.*

Abstract

The following paper presents the study of passive safety behavior of a modern helicopter subfloor conducted in collaboration with AgustaWestland. The staged work started with the experimental tests on a single intersection ending on a whole section of subfloor and a complete vertical test with a helicopter seat and AT dummy. The parallel development of a robust numerical model was conducted to reduce the number of tests required for validation of different design solutions. The numerical models were developed and analyzed with the explicit finite element solver LS-DYNA. LS-DYNA is a multidisciplinary finite element code (both explicit and implicit time integration) used to simulate and analyze highly nonlinear physical phenomena usually characterized by large displacements, high deformation rate and contacts, such as crashes, drop tests and impacts. LS-DYNA is worldwide recognized as a standard in passive safety analysis for its accuracy, stability and run-time optimization. This article presents the work related to the tests on individual intersections and four walled cells of the subfloor.

Introduction

To improve the conceptual design of a helicopter, able to fulfill both the structural and impact resistance requirements, a research program to study the energy absorption capacity of the subfloor was started. In particular crash tests have been performed on the structure of the subfloor and on the intersection elements, being the parts of the structure that can introduce a high peak of deceleration at the floor of the cabin causing harm to the occupants. Beside the above mentioned requirements, the design of a helicopter should prevent risk of fire, cockpit collapse as well as the collision between the occupants and the

* Associate Professor: marco.anghileri@polimi.it
** Graduated Laboratory Technician: andrea.milanese@polimi.it
*** Graduated Engineer: angelo.surini@polimi.it
[+] fabio.invernizzi@agustawestland.com

furniture. This is commonly obtained through the use of energy-absorbing elements incorporated in the landing gear, the subfloor and in the seats [2]. The structure of the subfloor must thus be designed to limit, through structural deformation, the deceleration of the cockpit. Particular attention should be given to the design of the intersections between the beams and bulkheads, which makes an essential contribution to the crash behavior of the subfloor [5].Commonly the subfloor is made of light aluminum alloy and absorbs the impact energy through plastic deformation. The research and studies that have taken place are oriented in three directions:

- the study of materials used to absorb energy: in the past only light metal alloys were used while today honeycomb, foam and especially fiber-based composites are adopted [9-11-13], to combine a high energy absorption capacity with relatively low weights;

- the study and optimization of new high-efficiency geometries [3-4-7-8-10]. To control the buckling phenomenon and prevent the total loss of post-crash resistance two solutions were studied: the first is to reinforce the panel with other materials such as foam, honeycomb, while the second is to model the panel so that it has a non-planar geometry;

- the correlation between the results of experimental tests [1-6] and numerical analysis to reduce development cost and to enable numerical optimization of the design.

This article gives special coverage to the latter, aiming to obtain a robust and efficient model that can be used in the future to reduce the number of trials needed in the design and validation phase.

1 Experimental tests

1.1 Tests on a single intersection

Twelve dynamic compression tests were performed on four types of intersection with different thickness, rivet size and stiffener cross section type (T and J). The specimen was crushed on a vertical drop-test setup with a rail driven mass of 110 [kg] hitting the structures at a velocity of 8.4 [m/s]. During the tests the acceleration was measured by two piezoresistive accelerometers placed on the impacting mass. Data acquisition system consisted of a 16 channels Pacific Instrumentation with programmable gain, power and sample rate up to 100 [kHz] per channel. Data were filtered through a digital filter type SAE CFC 180. Tests were also captured through a high-speed camera Photomark V5.1 with up to 100 [kfps]. The displacement was obtained through double numerical integration while the force was recovered from the mass-acceleration product. These results where then plotted to obtain the required force – displacement curve. The work presented here refers to the testing of the type-J stiffener specimens.

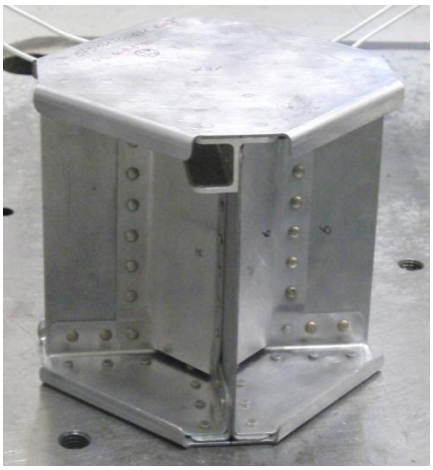


Fig. 1 - Intersection cruciform element before test

For each experimental test different parameters [3] were calculated. These parameters are able to fully characterize strength, absorption and efficiency of the tested specimen.



Fig. 2 - Intersection cruciform element after test

These parameters provide valuable information on the overall design quality of the absorber and are of critical importance to assess an all-round correlation with numerical models, thus improving the common force-displacement correlation. Besides the classical parameters of absorbed energy and mean force, obtained as the ratio of energy absorption and crushing distance, other meaningful quantities were computed like the distance of maximum efficiency (d max), over which the strength is too high compared to the first peak. The ratio between d max and the initial length d0 is a spatial efficiency index called "stroke efficiency" Se. All parameters were also computed at a common distance s to establish an estimate of the numerical-experimental correlation at a given point on the force-displacement curve.

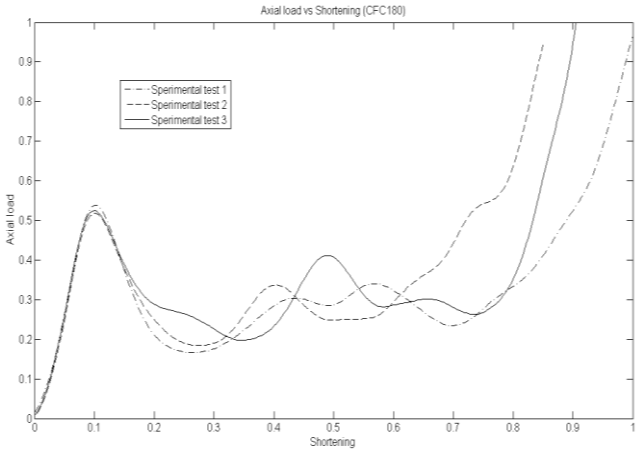


Fig. 3- Intersection cruciform element: experimental curve Axial load vs Shortening

The Table 1 shows the energy and force rating relative to a crushing length of s and d max , as well as the two highest force peaks before complete crush:

	d_{max}	s	$E(d_{max})$	E(s)	Se
Sper1	0.907	0.803	0.267	0.223	0.565
Sper2	0.770	0.803	0.239	0.257	0.482
Sper3	0.835	0.803	0.254	0.239	0.522

	$Favg(d_{max})$	Favg(s)	F(pk1)	F(pk2)
Sper1	0.295	0.278	0.539	0.341
Sper2	0.310	0.320	0.519	0.337
Sper3	0.304	0.298	0.525	0.412

Table 1– Experimental parameters

1.2 Subfloor cell experimental tests

Nine dynamic compression tests were then performed on three types of subfloor cell with different intersection, stiffeners and spar arrangement. The characterization tests were performed on a pneumatically driven horizontal sled of 640 [kg] mass at an impact velocity of 6 [m/s]. The specimen was tied to a rigid base, which was fixed perpendicularly to the lab floor. During the tests, the acceleration was measured by two piezoresistive accelerometers positioned at the center of the sled.



Fig. 4- Subfloor cell before test

The data acquisition system consisted of an IOTECH STRAINBOOK 616 with 8 input channels and a sample rate of 12500 [Hz] per channel. Data were filtered through a type SAE

CFC 180 digital filter. Tests were captured through the same high-speed camera Photomark V5.1 with up to 100 [kfps]. Filtered acceleration data and impacting mass ware then used to recover the crush force vs. time graph, as required by AgustaWestland.



Fig. 5- Subfloor cell after test

Our work focuses on the cell with the J-type stiffener for further numerical and experimental comparison.

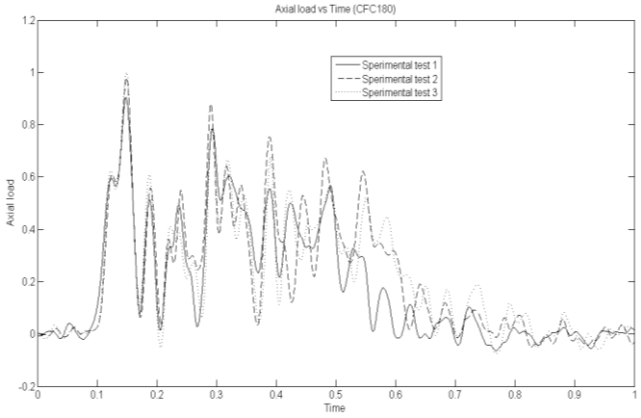


Fig. 6 - Subfloor cell: experimental curve Axial load vs Time

As stated above the accelerometer was placed on the sled. The measurement data, filtered with the SAE CFC 180, still contains part of the structural vibrations of the sled. Choosing a more conservative filter like a CFC 60 produced a very smooth curve, but it also dampened the first peak. We decided to use the CFC180 filter to get a better evaluation of the first force peak while maintaining a fair estimation of the average value.

2 Finite Elements Models

2.1 Numerical tests on single intersection model

The model used in the analysis of the intersection was obtained from the CAD and is therefore a faithful reproduction of the three intersections tested experimentally. The model is mainly composed of shell elements and beam elements. Since the beginning of the work rivets were considered key elements for the validation of the intersection's crash behavior and special attention was made on their modeling. Two node beam elements were favored over hexahedral ones as the high linear density made transmitted torque negligible. The rivets were modeled as point welds. Card *SECTION_BEAM was used to express chosen beam formulation and its cross section. Rivets material (Al 2117-T4) was implemented using the specialized Ls-Dyna's formulation (MAT-100). The rivets and the shell elements were connected by a CONTACT_SPOTWELD card as nodal connection would tie the mesh size to the number of rivets, thus limiting the use of the model in design optimization techniques. The evolution of the model has seen the change of many parameters that allowed us to obtain the results shown in the graph in Fig. 8. At the beginning a series of tests were carried out to identify the correct numerical friction coefficients. Although the very first results of strength and energy were comparable, there was a remarkable difference between the deformation mode of the numerical model and the one of the experimental tests. In particular the spar and the stiffener separated after the failure of the rivets (which does not occur in tests). This has led to further analysis of the elastic and plastic behavior of rivets s which revealed the following possible causes of such a different behavior:

- un-modeled static friction between the plates due to the pre-compression prior to riveting. This pre-compression results in a transfer of tangential stresses effort in the area interested thus reducing the load on the joints;
- the excessive rigidity of the contact with the impacting mass.

To limit this phenomenon and thus prevent the failure of the topmost rivets the yield strength was selectively increased. A subsequent stress analysis revealed that the axial and cutting forces acting on the other rivets were considerably higher than the typical strength of the elements. This was leading to another earlier break in the upper part of the specimen and then a low correlation in the crushing mode of the column. To contain this problem a global mesh refinement was adopted, bringing the typical mesh size of the spars and diagonal walls from 3 [mm] to 1.5 [mm] (Fig.7). The increased computational cost is then balanced by a better stress distribution in the rivet area, which finally led to a satisfying correlation of the numerical model.

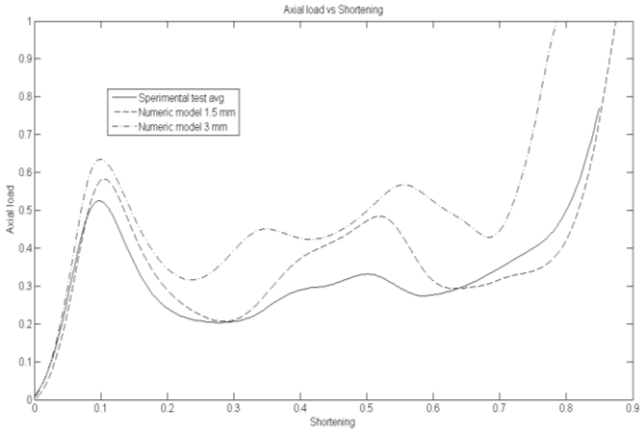


Fig. 7- Influence of the mesh size on FE simulations

A better correlation in the crushing mode was then finally achieved changing the main material model, from a simple and widely used piecewise elastic-plastic representation to a modern incremental-damage Johnson-Cook (JC) model. The finite element solver used already implemented the JC model in the material library, so data from a recent experimental work was used [12] . For the reader's convenience the data is presented in Table 2.

Young's Modulus [MPa]	Poisson's Modulus	A [MPa]	B [MPa]	n	C
72000	0.313	265	426	0.34	0.015
m	D_1	D_2	D_3	D_4	D_5
1	0.13	0.13	-1.5	0.011	0

Table 2- Johnson-Cook model properties for 2024-T3 Aluminum Alloy

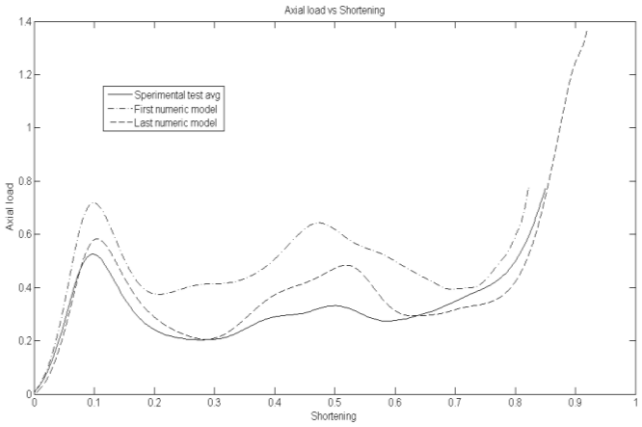


Fig. 8- Numerical and experimental load-shortening diagrams of the intersection cruciform element

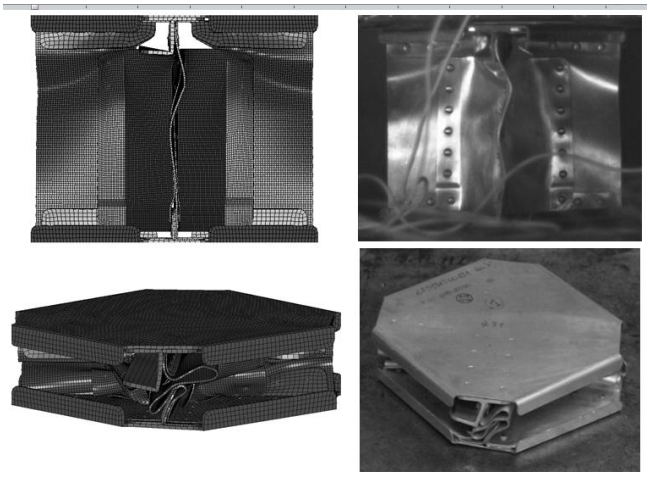


Fig. 9– Comparison between experimental and numerical deformations of the intersection cruciform element

For each test, numerical and experimental, several parameters were calculated to summarize the strength, absorption and efficiency behavior of the columns.

$$E(d)=\int_0^d F(x)dx \qquad (1)$$

$$F_{avg}(d)=\frac{E(d)}{d} \qquad (2)$$

$$A_E(d)=\frac{F_{avg}(d)}{F_{\max}(d)} \qquad (3)$$

$$T_E(d)=\frac{E(d)}{F_{\max}(d)\cdot l} \qquad (4)$$

These parameters, in addition to providing valuable information on the actual goodness of design of the absorber in its entirety, are of fundamental importance to assess a degree of correlation that does not stop at the, still extremely valuable, force-displacement curve. The following table presents the correlation indexes calculated: the correlation indices "r" and "corr" represent in turn the average error and the average correlation with the experimental evidence.

	corr_fo	corr_en	corr_F_{avg}	corr_A_E	corr_T_E
First numeric model	0.641	0.996	0.829	0.535	0.997
Last numeric model	0.780	0.997	0.918	0.730	0.993

	r[E(d_{max})]	r[E(s)]	r[Se]	r[Fpk1]	r[Fpk2]
First numeric model	0.467	0.418	0.940	0.638	0.215
Last numeric model	0.867	0.861	0.945	0.895	0.655

Table 3- Correlation parameters of the intersection element

As previously mentioned, despite an excellent energy correlation of the first models we chose to further refine the model to fix several deformation mode differences between the numerical model and the hi-speed video capture of the experimental tests. In particular, the failure of almost all of the rivets, the rotation of the specimen during crushing and overall incoherence of the deformation led us to investigate the causes and possible solutions to them. The decision to invest time in building a more satisfactory model has been rewarded by the identification of the main problems of the numerical model used initially.

2.2 Testing of the subfloor cell model

Development of the sub-floor cell numerical model took a minor amount of time thanks to the good correlation of the column model. As the same solutions developed for the single junction were adopted in the complete cell a good correlation in terms of average forces was immediately obtained. That can be seen as a good index of the robustness of the model previously developed. The curve obtained is shown in Fig. 10.

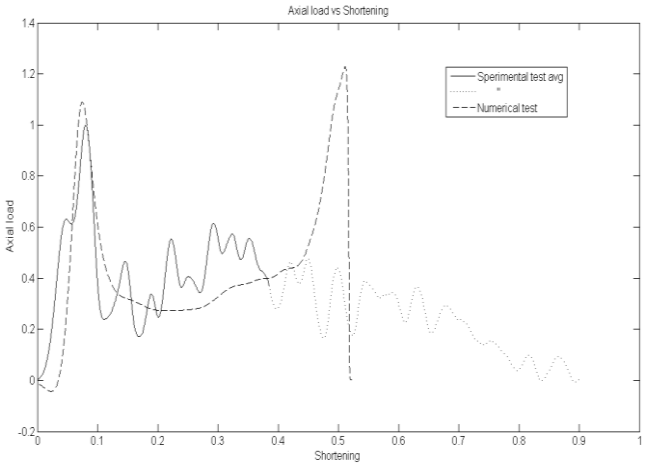


Fig. 10- Numerical and experimental load-shortening diagrams of the subfloor cell

The graph shows a good correlation in the initial phase of the test, while in middle of the crush length the numerical test managed to capture only the average force, probably due to the vibrations introduced by the impact in the experimental test. The mismatch of the final part can be considered acceptable because it is associated with low velocity and therefore energy absorption, although great care must be taken in making assumptions about the max crushing length of the cell. Two conclusions can be made from this evidence:

- at a preliminary project level one must consider the overestimation of the absorbed energy by the numerical model;
- at the certification phase the model is conservative due to the overestimation of the loads coming to the occupants, but its use could be problematic if absolute weight optimization is required.

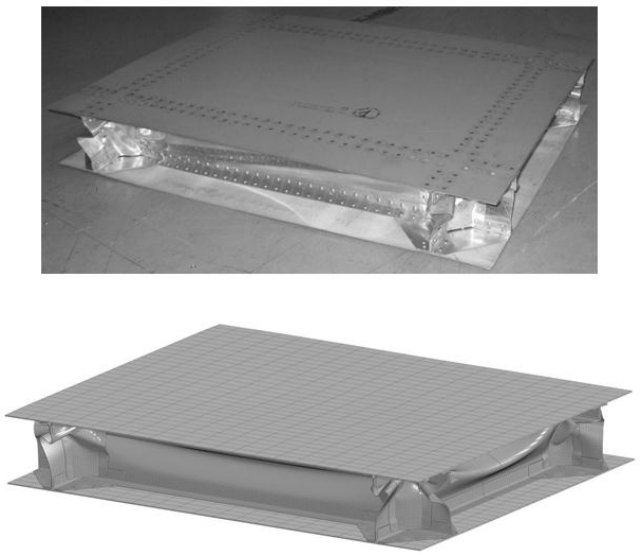


Fig. 11- Comparison between experimental and numerical deformations of the subfloor cell

3 Conclusions

The article presented the work done in collaboration with AgustaWestland on the study and development of a numerical model of a modern helicopter subfloor. In particular, the work focused on the optimization of correlation indexes between tests performed on individual intersections and complete four-walled cells of the subfloor. Three factors were identified as of critical importance for the robust implementation of such a model:

- an incremental damage material;
- the mesh size around the joints;
- the friction and viscous friction coefficients.

Although results on the complete cell are not completely correlated to those obtained experimentally we can be satisfied with the model developed for the intersection in that, as demonstrated by the correlation parameters used, it faithfully represents the evolution of energy absorption, crash efficiency and collapse mode. Good results have led us to implement this model for the simulation of a complete test which includes the sub-floor cell, a helicopter seat and an anthropomorphic crash test dummy. This test will be described in a later publication.

References

[1] Alguadich, S. (s.d.). Full scale crash test for cabin safety research. *Italian aerospace research center - (CIRA)* , 1-13.

[2] Bisagni, C. (2002). Crashworthiness of helicopter subfloor structures. *International Journal of Impact Engineering* , 1-16.

[3] Bisagni, C. (s.d.). Optimization of an helicopter subfloor under crash conditions. 1-16.

[4] Bisagni, C. (n.d.). Retrofit design of a crashworthy subfloor for a commuter aircraft. 1-13.

[5] Giavotto, V. (2000). The Significance of Crashworthiness. *Proceedings of the CEAS FORUM on "Crash Questions", Capua, Italy* , 1-11.

[6] K. Hughes, R. V. (2007). Experimental observations of an 8m/s drop test of a metallic helicopter underfloor structure onto a hard surface: part 1. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering - Vol 221, Number 5* , 661-678.

[7] L. Lanzi, A. A. (2006). Optimisation of energy absorbing subsystems for helicopter vertical crashes. *25th International Congress of the aeronautical sciences (ICAS)* , 1-10.

[8] L. Lanzi, C. B. (2004). Crashworthiness optimization of helicopter subfloor based on decomposition and global approximation. *Industrial applications* , 401-410.

[9] L. Lanzi, C. M.-L. (2004). Crashworthiness shape optimization of helicopter subfloor intersection elements. *24th International Congress of the Aeronautical Sciences (ICAS)* , 1-10.

[10] Lee, P. H. (1997). Topological optimization of rotorcraft subfloor structures for crashworthiness

considerations. *Computers and Structures Vol.64* , 65-76.

[11] M. A. McCarthy, J. F. (2001). Numerical investigation of a crash test of a composite helicopter subfloor structure. *Composite Structures* , 345-359.

[12] Research, O. o. (2000). Experimental investigation of material models for Ti-6Al-4V Titanium and 2024-T3 Aluminum. *U.S. Department of Transportation* , 17.

[13] S. T. Taher, E. M. (2006). A new composite energy absorbing system for aircraft and helicopter. *Composite Structures* , 14-23.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Linear Delamination Growth Initiation Prediction Tool for Stiffened Composite Panels

F. Di Caprio✉, A. Riccio, L. Iuspa and F. Scaramuzzino
Dipartimento di Ingegneria Aerospaziale e Meccanica
Seconda Università degli Studi di Napoli, Via Roma 29, 81031 – Aversa (CE), Italy
✉francesco.dicaprio@unina2.it

Keywords: *delamination, damage tolerance, composites, linear approach, FEM.*

Abstract

The delamination growth in composite panels is usually simulated with non liner analysis, but in this context, an alternative methodology which uses only linear analysis is proposed. Such an approach will allow a relevant reduction in the computational cost associated to damage tolerance analyses especially in the preliminary stages of the design process of complex composite structures. The proposed linear approach is compared here to the classical non-linear methodologies implemented in commercial FEM software and experimental results, and the most important parameters are investigated. A test case, consisting in a stiffened composite panel with an embedded bay delamination, has been considered. The delamination growth initiation threshold will provide indications on the damage tolerance capability of the panel in the presence of low velocity impact induced delaminations.

1 Introduction

This paper is focused on Carbon Fibres Reinforced Plastics (CFRP) composites. Due to their high specific strength and stiffness, CFRP are commonly considered particularly suitable for aerospace structural applications. However, their failure mechanisms are not completely predictable and for this reason the CFRP integration in the aerospace Industry has been generally slow down in the last twenty years.

Further, the lack of robust numerical tools able to take into account the damage tolerance of composite structures, has led to over-conservative designs, not fully realising the promised economic benefits of composites.

The mechanical behaviour of composite structures with delaminations has been widely investigated in the literature by performing experimental tests and by developing predictive numerical tools [1-3].

The compressive behaviour of panels with embedded delaminations is simulated in [6] by adopting three-dimensional models.

Complex numerical models able to simulate the growth of delaminations are presented in [7-13] where the Virtual Crack Closure Technique (VCCT), the Modified Virtual Crack Closure Technique (MVCCT) and the Cohesive Zone Models (CZM) are adopted for the calculation of the Energy Release Rate (ERR) by means of proper interface elements: suitable criteria are then used to identify the growth/no-growth status.

Both Virtual Crack Closure techniques and Cohesive Zone based approaches are progressively being implemented into commercial Finite Element codes which allow for a detailed simulation of delamination initiation and growth [14].

In [15-18], numerical and experimental activities on stiffened composite panels containing delaminations have been performed and the influence of the delamination position (in bay or under stringer) on its growth has been deeply investigated. In [17] a global-local approach is adopted to connect parts of the model with different mesh densities in order to reduce the computational cost associated to the geometrical non linear analysis required for the application of the VCCT and CZM approaches.

However, the use of non-linear techniques for delamination growth simulations is still not applicable to the preliminary design and optimisation of stiffened composite panels, even when accompanied by the global-local coupling. Hence newer fast numerical methodologies for delamination growth simulations (possibly based on linear calculations) are desirable in order to support the effective "damage tolerant" design of composite structures.

In the present paper a linear delamination growth initiation approach (alternative to VCCT and CZM approaches) for the preliminary design of delaminated stiffened composite panels is proposed. The presented methodology is derived by the approach introduced and validated in [19] enabling the simulation of delamination growth by means linear analyses. Under certain hypotheses, which can be considered acceptable in the preliminary design and optimisation phases, the approach presented in [19] and here revisited and deeply tested, can provide the delamination growth initiation load and location in delaminated composite structures.

2 A fast proposed procedure for delamination growth initiation and FEM implementation

Before presenting the procedure and its FEM implementation, it is important to review some phenomenological aspects about delamination, in order to opportunely clarify the hypotheses and the approximations made.

It is well known that, under compression load, composite structures and, in particular, delaminated stiffened panels [15-16], can exhibit local buckling phenomena. In particular, for bay delaminations of appropriate sizes and localized at relatively small depths, two main scenarios apply:

- the thinner sub-laminate buckles locally while the thicker sub-laminate remain unaffected (Fig. 1-a);
- both sub-laminates buckle exhibiting a sort of mixed buckling mode: such a mode can often be seen as a local mode with respect to the rest of the surrounding structure, however it involves the whole delaminated area (Fig. 1-b).

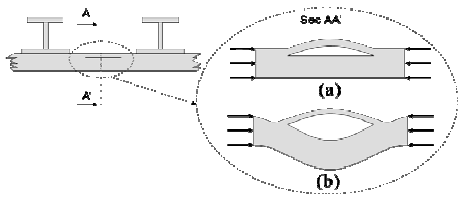


Fig. 1 Schematic representation of possible damage scenarios in composite structures: (a) local buckling of the thinner sub-laminate; (b) "global" buckling of the delaminated area

In general, under compressive load, in stiffened composite panels with thin bay delaminations, the first scenario *may evolve*, due to *delamination growth*, and it may lead to the occurrence of a second-like scenario, where the entity of damage is much more critical triggering more dangerous forms of buckling/failure phenomena.

The first-like delamination scenario could be induced by accidental sources of damage, such as impacts or manufacturing defects. Therefore, in order to design damage tolerant stiffened composite panels, it is extremely important to evaluate and to predict when, starting from a first-like delamination scenario, bay delaminations may propagate leading to a second-like delamination scenario.

Theoretically, delaminations can be handled like cracks in brittle materials; hence the

principles of linear elastic fracture mechanics for brittle materials apply. According to these principles, the propagation initiation of a delamination at a generic location along the delamination front, schematically shown in Fig. 2, is described by Eq. 1 where the ERR is defined as the variation of the potential energy ΔE associated to an increase ΔA of the delaminated area A. When this Released Energy reaches a critical value G_c, which depends only upon the material properties, the delamination is supposed to grow.

$$G = -\frac{\Delta E}{\Delta A} = G_c \qquad (1)$$

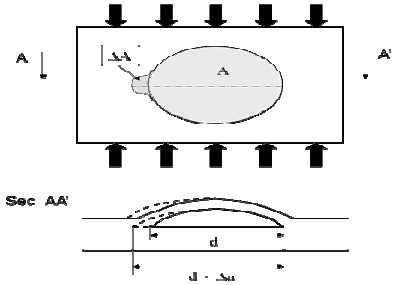


Fig. 2 Propagation of the delamination front (first scenario) under compression load

It is possible to distinguish among three different ERR contributions associated to the three basic fracture modes: mode I, mode II and mode III schematically shown in Fig. 3.

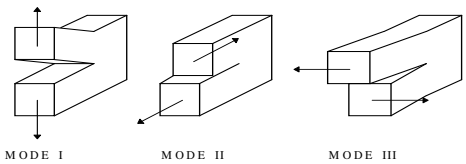


Fig. 3 Basic fracture modes

For the first delamination scenario, when the thickness of the buckled sub-laminate is much smaller than the total laminate thickness, the contributions of fracture Mode II and Mode III can be neglected.

The calculation of the energy released due to an increment in the delamination size and its relation with the applied load determines when the delamination advances. The ERR distribution along the embedded delamination front is non-uniform mainly due to the orthotropy of the materials, the non-uniformity of load distribution and the complexity of the surrounding structure. In particular, the ERR can locally exceed its critical value leading to a non-uniform progression of the delamination front. However, for the first-like delamination scenario the ERR can be assumed mostly influenced by the out-of-plane displacements along the delamination front [9, 16] on first delamination buckling.

For complex structural and material configurations, such as stiffened composite panels, analytical approaches for the determination of the ERR are clearly not easily applicable. Therefore numerical technique based on non-linear algorithms are required: as a consequence, the computational costs associated to such analyses may be quite relevant and time consuming.

Here, a new linear approach, applicable to the first delamination scenario under the previously cited assumptions, is widely investigated. This procedure, can result extremely useful when designing stiffened composite panels capable to avoid the delamination growth event by predicting both the external load which causes the progression of an initial given delamination and growth location along the delamination front.

Indeed, considering only the first delamination scenario under the two main assumptions previously introduced (Mode I predominance and ERR distribution mostly influenced by the fist delamination buckling out-of-plane displacements distribution along the delamination front), the linear method can be successfully applied to stiffened composite panels with thin bay delaminations. Moreover, it should be noted that thin bay delaminations are of major concern in aerospace design with composite materials since they represent one of the most frequently occurring damage typology due to low velocity impact treats. Hence, the proposed linear approach can be properly considered as finalized to the improvement of preliminary design and optimization of stiffened

composite panels tolerant to low velocity impacts induced damage.

The new proposed methodology starts with a linearized buckling analysis to calculate the first buckling load and mode.

Considering the overall structure subjected to an external compressive load F, when the elastic stability threshold is reached, the thinner sub-laminate will buckle. An eigenvalue analysis (linearized buckling analysis) can give the delamination buckling load (eigenvalue) and mode (eigenvector).

Beyond the delamination buckling event, characterized by the critical load F^{cr}, the residual non-linear contribution of the buckled thinner sub-laminate to the global stiffness can be neglected when the thickness of the buckled sub-laminate is smaller than the total laminate thickness. In this situation, even long after the first buckling event, the non-linear stiffness contribution due to large displacements and rotations of the buckled sub-laminate is extremely small when compared to the global laminate stiffness. Thus, the stiffness K^A of the overall post-buckled structure with a delamination size A can be approximately calculated by removing the thinner sub-laminate. The same considerations can be repeated when the delamination size is $A+\Delta A$ (see Fig. 2); hence a new delamination buckling load $F^{cr'}$ and a new stiffness of the post-buckled structure $K^{A+\Delta A}$ can be defined.

In Fig. 4 a graphical representation of the global structural stiffness evolution is given.

Starting from the pre-buckling global stiffness value K^0 , the delamination buckling for the two analyzed delamination sizes reduces the global stiffness respectively to K^A and $K^{A+\Delta A}$ opening two distinct equilibrium paths. With reference to Fig. 4, the progressive quasi-static increase of the delamination size, from A to $A+\Delta A$, for a structural element subject to compressive displacements, can be reduced to a jump from the former to the latter equilibrium path [11].

Hence, for each displacement value u^*, the overall energy loss $\Delta E(u^*)$ (area $A'ACB$ in Fig. 4), related to the delamination size increment ΔA, is given by Eq. 2.

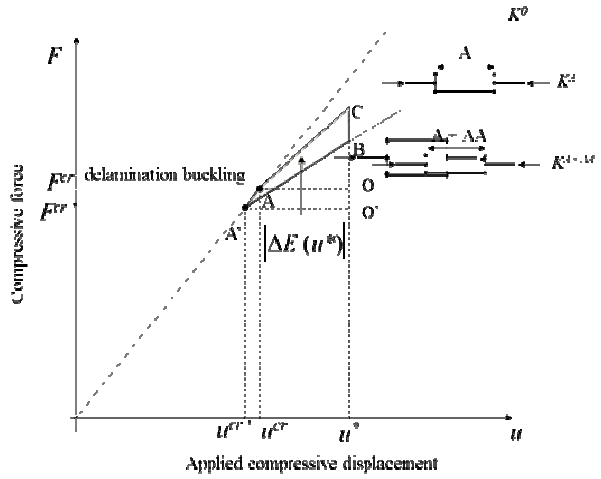


Fig. 4 Stiffness evolution of the stiffened composite panel under compression; before, after the delamination buckling and after the delamination propagation

$$\begin{aligned}
&\Delta E(u^*) = E^A(u^*) - E^{A+\Delta A}(u^*) = \\
&= area(A'ACO') - area(A'BO') = \\
&= area(ACO) + area(A'AOO') - area(A'BO') = \\
&= \frac{1}{2}K^A\left(u^* - u^{cr}\right)^2 + \frac{1}{2}K^0\left(2u^* - u^{cr'} - u^{cr}\right)\left(u^{cr} - u^{cr'}\right) \\
&- \frac{1}{2}K^{A+\Delta A}\left(u^* - u^{cr'}\right)^2
\end{aligned} \quad (2)$$

where $E^A(u^*)$ and $E^{A+\Delta A}(u^*)$ are the elastic energies absorbed as a consequence of the applied displacement u^*, by the structures respectively with the delamination size A and $A+\Delta A$.

Equation 2 represents the total amount of energy released during the delamination propagation as a quadratic function of the applied displacement. Considering that the delamination growth initiation is a local phenomenon occurring only at the location of delamination front, where the ERR is maximum, the correct evaluation of the quantities involved in Eq. 2, necessarily requires the determination of the location characterized by the maximum ERR.

According to the assumptions made in the previous section, it can be assumed that the ERR distribution, related to the delamination size increment ΔA has the same distribution as the square of the first delamination buckling out-of-plane displacements.

By means of this consideration, the location m where the ERR reaches the maximum can be easily determined. The delamination growth will initiate at this location when the local ERR will reach the critical value G_{IC}. Indeed, at growth initiation, the following relation must hold:

$$\frac{G^m(u^*)}{G_{IC}} = \frac{\Delta E(u^*)}{\Delta A^m \cdot G_{IC}} = 1 \qquad (3)$$

where ΔA^m is the increment of delamination size at the maximum ERR location.

By solving Eq. 3 for the applied compressive displacement u^*, the critical displacement u^{del} at which the *initiation of delamination growth* is expected, can be found.

The last step in the proposed methodology consists of the evaluation of the applied external compressive load required to trigger the delamination growth; this load value can be computed by using Eq. 4.

$$F^{del} = F^{cr} + K^A\left(u^{del} - u^{cr}\right) \qquad (4)$$

The methodology presented above has been implemented in the ANSYS FEM code [20] by using a macro written in the Ansys Parametric Design Language (APDL).

In Fig. 5-a, the shell finite element model of a stiffened panel with a bay delamination is presented, while, in Fig. 5-b the different parts of delamination model are shown.

Eight-node Shell99 layered elements are used to model the skin, the stringers and the delaminated sub-laminates. Full DOFs rigid constraints are used to connect the stringer feet to the skin. Other constraints, uniformly applied along the delamination front at the delamination tip, are used to connect the sub-laminates to the rest of the skin. MPC (Multi-Point Constraints) elements are adopted to connect the delaminated area to the rest of panel without transition mesh.

In the ANSYS environment all the quantities shown in Eq. 1-4 can be easily determined by combining linearized buckling analyses and linear static analyses by means of an APDL macro. Hence, the values of the applied displacements and the loads at growth initiation can be determined.

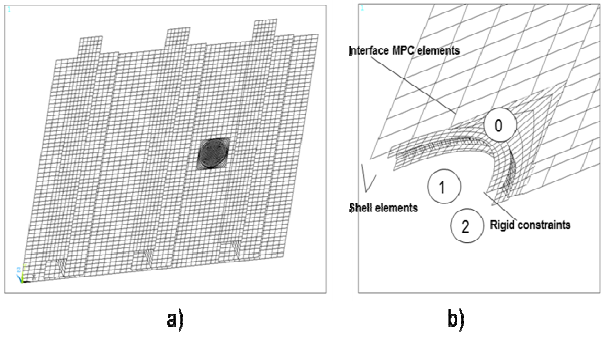


Fig. 5 a) FEM model of the stiffened panel with bay delamination; b) Delamination model details

3 Numerical Application

3.1 Test Case description

The proposed linear approach has been applied to the test case taken from [21]. A stiffened composite panel with a bay delamination has been considered. The main outputs of the linear approach (delamination buckling loads, growth initiation loads and ERR distribution along the delamination front) will be compared to the results from the ANSYS non-linear analysis based on Modified Virtual Crack Closure Technique presented in [21] and to the results of ABAQUS non-linear analysis, based on the VCCT, for validation purposes.

The boundary conditions are shown in Fig. 6, while the geometrical features of the test-case are given in Fig. 7.

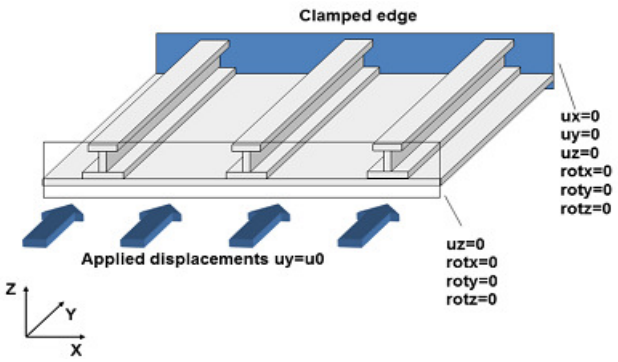
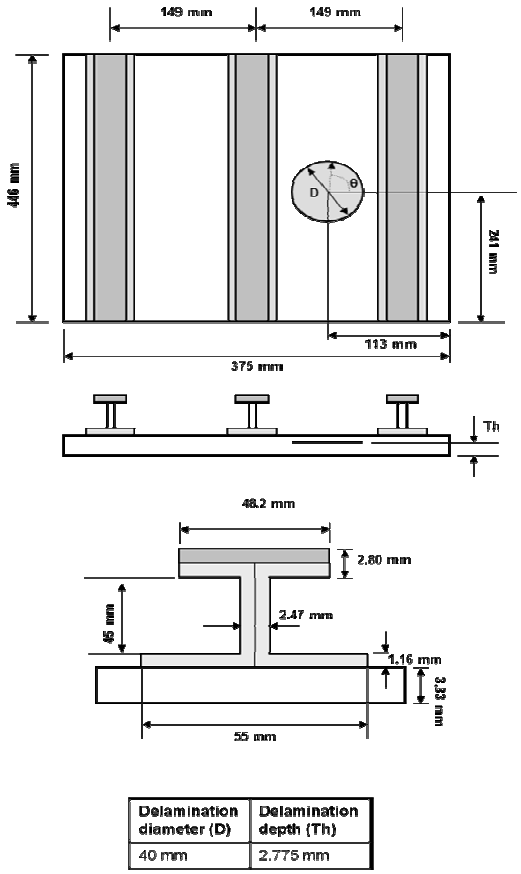


Fig. 6 Applied load and boundary conditions

The stacking sequence of the panel's skin is $[+45°/-45°/0°/90°]_{3S}$, while the stringers consist of three laminates, each with the following lay-up [+45°/-45°/0°3/90°/0°3/-45°/+45°], as shown in Fig.7. In the front view of this sketch, the angle θ is also reported. This angle is introduced to univocally identify a location along the delamination front.



Delamination diameter (D)	Delamination depth (Th)
40 mm	2.775 mm

Fig. 7 Geometrical description of the delaminated stiffened panel configurations

The material properties of the adopted composites (HTA-6376C) are summarized in Table 1.

Property	HTA-6376C
Longitudinal Young's modulus, E_{11} (GPa) - Tension	140.
Transverse Young's modulus, E_{22} (GPa) - Tension	10.5
In-plane shear modulus, G_{12} , G_{13} (GPa)	5.2
In-plane shear modulus, G_{23} (GPa)	3.48
Poisson's ratio, n_{12} n_{13}	0.3
Poisson's ratio, n_{23}	0.51
Critical ERR-Mode I, G_{IC} [40mm] (Jm^{-2})	260
Critical ERR-Mode II, G_{IIC} [60mm] (Jm^{-2})	950
Critical ERR-Mode III, G_{IIIC} (Jm^{-2})	1200
Ply thickness (skin), t1 (mm)	0.13875
Ply thickness (web), t2 (mm)	0.112
Ply thickness (cap), t3 (mm)	0.127
Ply thickness (foot), t4 (mm)	0.105

Table 1 Material properties

3.2 Mesh density – Convergence study

In order to determine the most effective (best compromise between computational cost and accuracy of results) element size across delamination front, different mesh configurations have been considered. The investigated mesh configurations have been characterized by different widths of the elements on the delamination front to determine the optimal mesh size for the FE model.

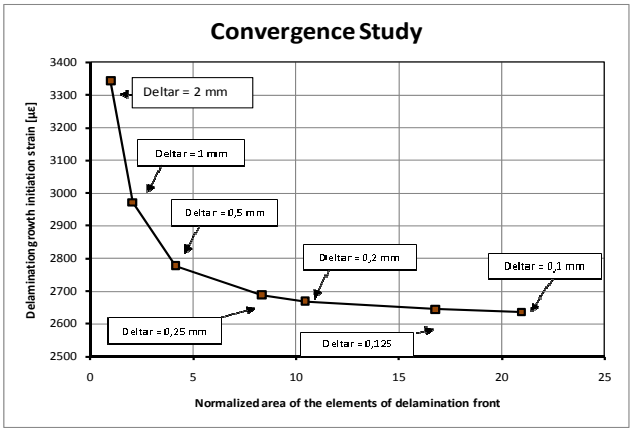


Fig. 8 Delamination growth initiation strain versus the normalized area of the elements along delamination front

For convenience in Fig. 8, each of the analyzed configuration has been identified by the normalized area of the elements along the delamination front. The areas have been normalized with respect to the area of the first configuration (element width *deltar* = 2 *mm*).

As shown in Fig. 8, the configuration characterized by *deltar* = 0.25 *mm* provides a very accurate result using, at the same time, a relatively coarse mesh along θ and keeping the aspect ratio of the shell elements along the delamination front within the warning limits.

Following this sensitivity study, the configuration with *deltar* = 0.25 *mm* has been therefore chosen for the analyses and the comparison test-cases.

3.3 Numerical results and comparisons with literature results

In the considered test-case, a delamination with a diameter of 40 *mm*, is located in the bay, between the fourth and the fifth ply. In Fig. 9, the applied strains as a functions of the applied compressive loads are plotted as evaluated by the linear and non-linear model; the numerical results are also compared to the strain-gauge data produced by an experimental model.

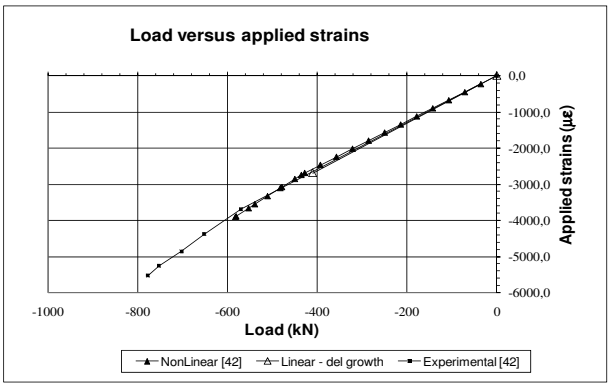


Fig. 9 Load versus applied strains – comparison between linear, nonlinear and experimental results

In Fig. 9, the excellent agreement between the linear and non-linear results, in terms of global longitudinal stiffness, confirms the linearity of the global compressive behavior of the stiffened panel, up to the delamination growth initiation load.

In Table 2, the delamination buckling and growth initiation loads and applied strains, computed according to the linear model, are compared with the non-linear and experimental results.

According to Table 2, the linear model can correctly predict the delamination buckling load and applied strains. The percentage difference between linear and nonlinear delamination buckling load is about 10%.

It should be also noted that the experimental delamination buckling values are affected by the presence of the Teflon film inserts to simulate the artificial delamination, as discussed in [18].

For these specimens, the experimental delamination buckling data cannot be considered reliable for the validation of the model.

	ε^{cr} (delaminationl buckling strain) (με)	F^{cr} (delamination buckling load) (N)	ε^{del} (strain at growth initiation) (με)	F^{del} (growth initiation load) (N)
Simplified linear model (quadratic distribution)	**1496**	**228336**	**2688**	**410235**
Non-linearModel	1345	214127	2747	435354
Experimenta results [32]	1250	193052	2713	419000
Diff % linear-nonlinear results	+10%	+6.2 %	+2.14 %	-6,1 %
Diff % nonlinear-experimental results	+7 %	+9.8 %	+1.2 %	+3.7 %

Table 2 Comparison among major results

As a matter of facts, the very strong capability and effectiveness of the presented linear methodology, coming out from previous numerical results, can be related to the growth initiation stage. The good agreement between the non-linear and the experimental growth initiation strain (1.2% percentage difference and below the experimental scatter of 10%), confirms that the non-linear model can be taken as a validation benchmark for the linear approach when determining the delamination growth initiation.

This good agreement between linear and non-linear model, which can be considered more than acceptable in a preliminary design phase, represents a first proof of the effectiveness of the proposed linear approach. In Fig. 10, the nearly identical numerical and experimental out-of-plane displacements distributions within the delaminated area, at delamination buckling and growth initiation, are shown for the test-case under consideration. This is an evident proof that the out of plane displacements distribution does not undergo changes from the delamination buckling to the growth initiation event.

The predicted location of maximum ERR (-9° inclination with respect to the global *X* axis) is exactly the same for both linear and non-linear model.

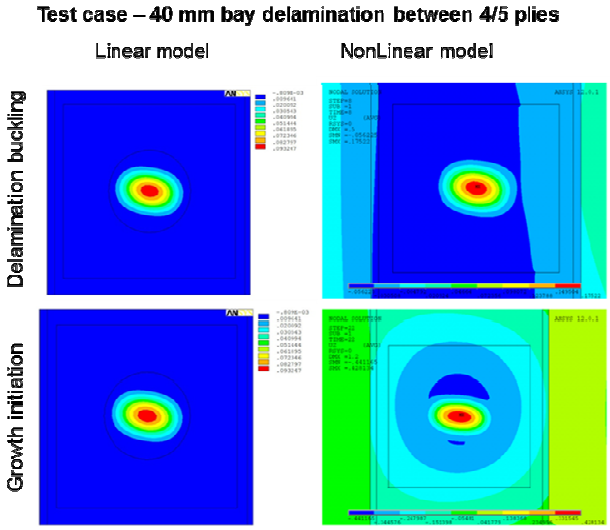


Fig. 10 Applied contour plots of the out-of-plane displacements distributions at delamination buckling and at growth initiation - comparisons between linear model and nonlinear model

In Fig. 11, the distribution along the delamination front of the Mode I contribution of the ERR is shown.

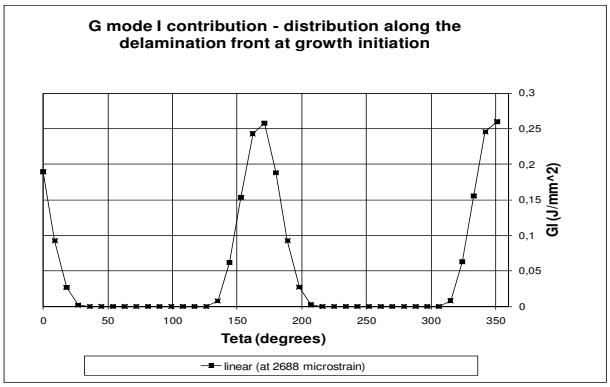


Fig. 11 Distribution along the delamination front of the Mode I ERR contribution at growth initiation

3.4 ABAQUS VCCT approach comparison

In order to further verify the developed linear approach, the stiffened panel described and analyzed in [21] has been modeled and analyzed with the ABAQUS FEM code. Then, the obtained results are compared to the previous ones.

The delaminated region has been modeled by two parts, each one composed by brick volumes, The two parts, representing the sub-laminates, have been connected by node-to-surface interaction with the option "Virtual Crack Closure Technique".

model	Delamination buckling load strain ($\mu\varepsilon$)	Delamination growth initiation strain ($\mu\varepsilon$)	Global buckling strain ($\mu\varepsilon$)
ABAQUS linearized	1256	-	2690
ABAQUS non linear	1483	2325	2679
ANSYS non-linear [42]	1345	2747	-
Linear approach	1496	2668	-
% error linear approach with respect to ABAQUS non linear	0.8%	14.7%	-

Table 3 Summary of the relevant applied strain for the compared numerical models

Table 3 shows that the agreement between the linear approach and the ABAQUS non-linear approach is still good and acceptable for a preliminary design stage. However ABAQUS seems to underestimate the results of ANSYS [21]. In Fig. 12, the load strain curves are plotted for each of the analyzed models.

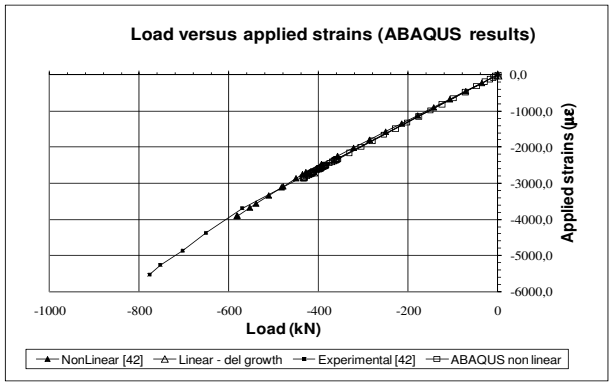


Fig. 12 Load versus applied strains – Comparison

The agreement in terms of structural stiffness among the ANSYS [21], ABAQUS, and linear models and the experimental data is excellent. However, the underestimation of the skin buckling event with ABAQUS is appreciable.

In Fig. 13 the out-of-plane displacements evaluated at the centre of the sub-laminates, as a function of the applied strains, is plotted for each of the non-linear model.

From Fig. 13 it is possible to note that both ANSYS [21] and ABAQUS non-linear models are able to follow the experimental behavior and give a reasonable prediction of the delamination growth initiation applied strains.

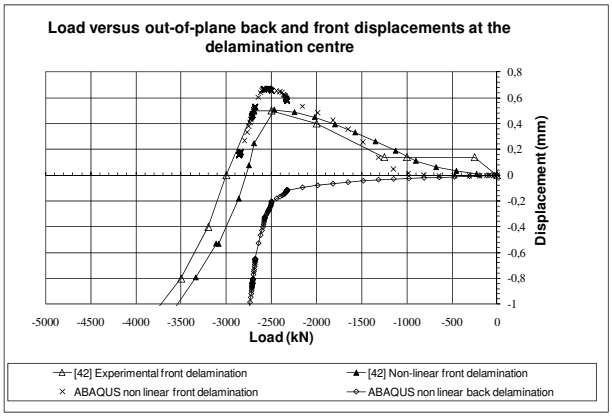


Fig. 13 Out-of-plane displacements versus applied strains – Comparison

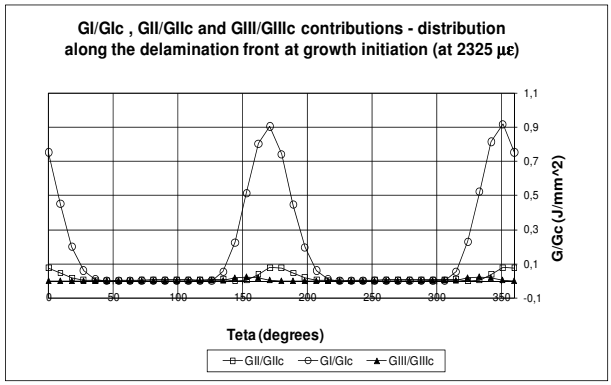


Fig. 14 ERR normalized distributions at growth initiation (2325 *με*)

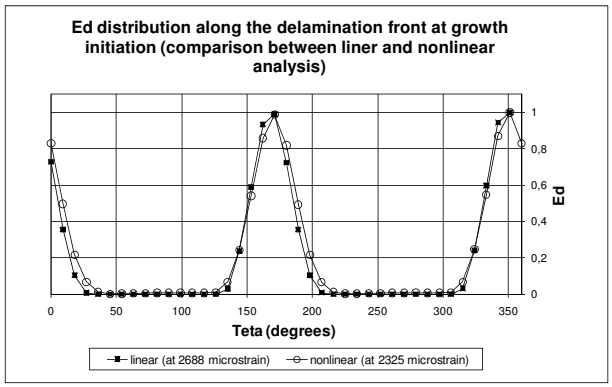


Fig. 15 Ed distributions at growth initiation

However, during the delamination growth, both models seem to produce a load displacement curve which slightly deviates from the experimental trend. Nevertheless, in the final phase of the delamination growth, ABAQUS is able to follow exactly the experimental trend while ANSYS [21] model slightly overestimates the out-of-plane displacements.

The delamination growth initiation event with the ABAQUS non-linear model is located at 2325 με. The predominant fracture mode, as expected, is the Mode I; hence the delamination growth initiation is a consequence of the delamination buckling event.

The distribution of the ERR at growth initiation is very useful to understand the location of growth and the weight of each fracture mode in the growth initiation. In Fig. 14 the distribution of the ERR normalized along the delamination front for the three fracture modes at the growth initiation for the ABAQUS non-linear analysis is shown.

The comparison between the ABAQUS non-linear model and the developed linear model in terms of E_d is shown in Fig. 15.

It is relevant to observe that for the non-linear model the following relation hold $E_d = G_I/G_{Ic} + G_{II}/G_{IIc} + G_{III}/G_{IIIc}$ while for the linear models it is $E_d = G_I/G_{Ic}$.

It is possible to appreciate that both the distributions and the angles of delamination growth initiation (-9°) are in full agreement for the ABAQUS non-linear model and the developed linear model.

4 Conclusion

In the present paper, a novel methodology, able to predict the delamination growth initiation in stiffened composite panel with a bay delamination has been proposed. This methodology, based on the linearized buckling evaluation and on the application of the Energy balance principle, has been implemented in the ANSYS FEM code and proved to be able to reasonably predict, by means of a limited number of linear analyses, the location along the delamination front and the load at which a delamination may propagate.

Comparisons with the literature non-linear and experimental results for a stiffened composite panel with a bay embedded delamination have demonstrated that the proposed approach is affordable, as expected, when the delamination depth falls within a range of relatively small values (≤15% of the total skin thickness). Such a situation is quite representative of manufacturing defects and damages induced by low velocity impact events.

In order to further validate the developed tool, the linear results have been compared with ABAQUS non-liner results and an acceptable agreement has been found again. The location of the delamination growth initiation and the energy release rate distribution at growth initiation are again in excellent agreement. The optimal compromise between the very limited computational effort and the accuracy of the results, suggests possible applications to the preliminary design and optimization of complex damage tolerant composite structures.

Even if the proposed linear approach works properly for thin bay delaminations, some improvements could be made in future work in order to limit the assumptions and to extend the applicability of the model to stiffened panels with thicker delaminations.

References

[1] Ramkumar R. L., "Performance of a Quantitative Study of Instability-related Delamination Growth", NASA CR-166046, 1983.

[2] Chai H., Knauss W. G. and Babcock C. D, "Observation of Damage Growth in Compressively Loaded Laminates", Journal of Experimental Mechanics, Vol. 23, No. 3, 1983, pp. 329-337.

[3] Yokozeki T., Aoki Y., Ogasawara T., "Experimental characterization of strength and damage resistance properties of thin-ply carbon fiber/toughened epoxy laminates", Composite Structures, Vol. 82, No. 3, 2008, pp. 382-389.

[4] Gordnian K., Hadavinia H., Mason P.J., Madenci E.b., "Determination of fracture energy and tensile cohesive strength in Mode I delamination of angle-ply laminated composites", Composite Structures, Vol. 82, No. 4, 2008, pp. 577-586.

[5] Whitcomb J. D., "Approximate Analysis of Postbuckled Through-the-width Delaminations", Composite Technology Review, Vol. 4, No. 3, 1982, pp. 71-77.

[6] Shivakumar K. N. and Whitcomb J. D., "Buckling of a Sub-laminate in a Quasi-Isotropic Composite Laminate", Int. Journal of Composite Materials, Vol. 19, No. 1, 1985; pp. 2-18.

[7] Whitcomb J. D. and Shivakumar K. N., "Strain-Energy Release Rate Analysis of a Laminate with a Postbuckled delamination. Numerical Methods in Fracture Mechenics", NASA TM-89091, 1987.

[8] Whitcomb J. D., "Analysis of a Laminate with a Postbuckled Embedded Delamination, Including Contact Effects", Journal of Composite Materials, Vol. 26, No. 10, 1992, pp. 1523-1535.

[9] Perugini P., Riccio A. and Scaramuzzino F., "Influence of Delamination Growth and Contact Phenomena on the Compressive Behaviour of Composite Panels", Int. Journal of Composite Materials, Vol. 33, No. 15, 1999, pp. 1433-1456.

[10] Riccio A., Perugini P. and Scaramuzzino F., "Modelling Compression Behaviour of Delaminated Composite Panels", Computers & Structures, Vol. 78, 2000, pp. 73-81.

[11] Gaudenzi P., Perugini P. and Riccio A., "Post-buckling behaviour of Composite Panels in the presence of Unstable Delaminations", Composite Structures, Vol. 51, No. 3, 2001, pp. 301-309.

[12] Riccio A., Perugini P. and Scaramuzzino F., "Embedded Delamination Growth In Composite Panels Under Compressive Load", Composites part B: Eng., Vol. 32, No. 3, 2001, pp. 209-218.

[13] De Borst R. and Remmers J.J.C., "Computational modelling of Delamination", Composites Science and Technology, Vol. 66, 2006, pp. 713-722.

[14] ABAQUS MANUAL (revision 6.5-1): Theory.

[15] Greenhalgh E., Singh S., Hughes D., Roberts D., "Impact damage resistance and tolerance of stringer stiffened composite structures", Plastics, Rubber and Composites Processing and Applications, Vol. 28, No. 5, 1999, pp. 228-251..

[16] Greenhalgh E., Singh S., Nilsson K.-F., "Mechanisms and modeling of delamination growth and failure of carbon-fiber reinforced skin-stringer panels", ASTM Special Technical Publication, 2000, 1383:49-71.

[17] Suemasu H., Kurihara K., Arai K., Majima O., Ishikawa T. "Compressive property degradation of composite stiffened panel due to debonding and delaminations", Advanced Composite Materials: The Official Journal of the Japan Society of Composite Materials, Vol. 15, No. 2, 2006, pp. 139-151.

[18] Faggiani A., Falzon B.G., "Optimization strategy for minimizing damage in postbuckling stiffened panels", AIAA Journal, Vol. 45, No. 10, 2007, pp.2520-2528.

[19] Riccio A., Gigliotti M., "A Novel Numerical Delamination Growth Approach For The Preliminary Design of Damage Tolerant Composite Structures", Journal of Composite Materials, Vol. 41, No. 16, 2007, pp.1939-1960..

[20] ANSYS MANUAL (revision 5), Volume III: Elements DN-R300:50-3.

[21] Pietropaoli E., Riccio A., "On the robustness of finite element procedures based on Virtual Crack Closure Technique and fail release approach for delamination growth phenomena. Definition and assessment of a novel methodology", Composites Science and Technology, Vol. 70, No. 8, 2010, pp. 1288-1300.

CEAS 2011

UAS 1

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Innovative Unmanned Aircrafts: role and constraints for GMES applications

E. Cestino , G. Romeo

Aerospace Engineering Department, Politecnico di Torino,

Corso Duca degli Abruzzi 24 – 10129 - Torino, Italy

Keywords: *GMES, Unmanned Aircraft, solar power, flight tests*

Abstract

Within the TANGO project, a three years project funded by the EU commission for the development of new telecom services for GMES (Global Monitoring for Environment & Security) operations, a multicriteria comparison between different UAVs configurations has been performed. Configurations have been assessed regarding a lot of representative scenarios (Risk & Crisis, Maritime, Security, Land cover, Humanitarian Aid). Almost all these scenarios can be performed with success by UAVs. Two innovative configurations has been studied in details: 1) A Solar powered HALE configuration that obtained a very good score for mission similar to "Border Patrol Mission" in which the aircraft would fly from one sea region to the next and track targets or loiter in a region for a period of time. A very long endurance as well long range are the two main mission parameters. 2) A Small Solar powered UAV shows interesting results relative to "Boat Identification" missions which require the deployment of a platform in closer proximity to suspicious vessels. Finally the Small Electric Solar Aircraft (SESA) was positively introduced in a TANGO demonstration.

1 Introduction

The GMES (Global Monitoring for Environment & Security) services are currently being developed to support public policy makers' needs in the domain of environment and security, and rely on a comprehensive Earth observing system, using space borne and in situ techniques.

Although the use of Unmanned Aircraft Systems (UAS) was for many years restricted to military applications, recent technological developments and increased utilization of Unmanned Aircraft Systems (UAS) have widened their application from military operations only to also civil and commercial operations.

Nevertheless, the civil market for the use of UAVs will remain difficult to develop until operational regulations for the use of UAVs in non segregated airspace are developed and approved by ATM authorities.

Potential users have interests to deploy UAS in non segregated airspace. Recent technological and operational improvements give reason to believe that UAS safety and performance capabilities are maturing. UAS, together with satellite telecommunications will be a key component of the future GMES architecture, enabling to expand GMES services and to

enhance their performances where unmet needs can be identified:

• Improvement of the service area through the dissemination of the GMES applications wherever it is needed; reach communities with no other solutions –for instance in the ocean or following a natural disaster-; deployment of Ad hoc networking for crisis management,

• Improvement of the reactivity and freshness of the data through faster scene and in-situ data collection

• Speed up the transfer of data expected as these prototypes services become operational and allow higher volumes to be processed

In this framework, the European Commission has selected the TANGO Project "Telecommunications Advanced Networks for GMES Operations" to develop, integrate, demonstrate and promote new telecom services dedicated to GMES requirements.

TANGO implements a bottom up approach to identify the requirements for telecommunication services that are not met for the delivery of GMES services and try to find solutions throughout the integration of existing systems and advanced solutions in the framework of security and crisis management, fisheries management, maritime surveillance and humanitarian aid.

The present paper presents results and recommendations of an high level analysis with comparison between various platforms in performing different TANGO missions through the development of a multicriteria tool in which inputs from performances characteristics of reference aircrafts are compared to requirements from mission typology.

The last phase of the TANGO Project demonstrated, trough a real experimental phase, the ability to integrate land monitoring existing systems with monitoring aircraft systems based on a small UAS platform, with high range and endurance capabilities, powered by alternative energy sources based on solar cells, batteries and fuel cells in order to develop a more effective monitoring system with the introduction of a local air component with low environmental impact.

2 Description of the procedure

The implementation of a multicriteria analysis for the identification of suitable applications for different classes of UAVs within TANGO thematic (Risk & Crisis, Maritime, Security, Land cover, Humanitarian Aid) has been subdivided in four parts:

- Definition of applications requirements.
- Assessment of the Capabilities & Performance of various UAVs
- Analysis of capabilities versus user needs through a multi-criteria analysis.
- Identification of suitable TANGO applications

2.1 Definition Of Applications Requirements

Several scenarios (e.g. oceanic Observations, fire fighting, oil spills, boarder surveillance etc…), has been identified in term of payload typology (Infrared, EO, SAR, Lidar etc.), characteristics of various payloads (mass, power consumption, volume, data link etc.) and requirements of the mission (all weather capabilities, day/night operation, endurance, altitude, flight speed etc.).

Data about missions requirements has been collected through data sheets aimed to the identification and description of specific applications in term of payload main characteristics, operational needs, advantages and benefits compared to existing implementations.

The analysis presented is an high level analysis and so a detailed description of payloads for each scenario it is not possible and, at this level, it has been identified only the type of sensor required by the scenario and a first estimation of weight, and some operational needs in terms of speed, endurance etc. based on previous experience and available open literature documents. For each scenario a weight index is included in order to describe main missions performance requirements. A high weighting index means the criterion is very important for the mission, a low weighting

index means the criterion is not very important for the mission. For example, in the case of scenario "Oceanic Observations" (Fig. 1) the performance criterion "endurance" has a weighting factor of 6, because a high endurance is important for the mission while for example the landing and take off length are less important for the selected scenario. These weights indices are used in the multicriteria method to give the final ranking. They are very important, so the user has to be very careful when he makes the weighting Fig 1.

performances	weight indices	> o <	required	optimum
vertical take off	0			
HOGE	0			
all weather capabilities	3			
manuvrability	4			
range	5			
endurance	6	>	24[hr]	
range of controllability	3	>	1000[km]	
maximum velocity	5	>	35[m/s]	
climb rate	3			
ceiling	0	>	2500[m]	
payload weight	4	>	30[kg]	
minimum velocity	5	<	60[m/s]	
landing length	2			
take off length	2			
noise on flight	5			
minimum altitude	0	<	10000[m]	
loiter velocity	5			50[m/s]

Fig. 1 Example of Mission Requirements for Oceanic Observation Missions

The weight scale used to characterize missions is reported in Table 1. It should be noted that in case the mission includes some minimum or maximum required performances (>o<) the weight indices measure the importance to overcome the basic requirement, or if the mission include some optimal requested value, weight indices measure the importance to be closer to the optimal value.

Table 1

Mark	Meaning
10	Capability is extremely important for the mission
6	Capability is one of the more important parameters for the mission.
2	Capability is not very important for the mission and a low performance is sufficient to perform the mission.
0	Capability is not critical for mission

2.2 Assessment Of The Capabilities & Performance Of Various UAVs

On the technical side, the UAV system can be characterized by an air system and a ground system, interconnected by data links, control links, and integrated into a technical and decision-making environment. Civil UAVs are also categorized as a function of weight and operational altitude and are usually grouped into the following four main categories as defined below

- MICRO/MINI: less than 7kg in weight with flight altitude below 400ft.
- MINI/SMALL: weight range of about 7- 20 kg with flight altitude below 400 ft, range below 500m, Visual Flight Rules (VFR).
- MALE: (Medium Altitude with/without Long Endurance) with flight altitude above 400 ft, range above 500m, VFR.
- HALE: High Altitude Long Endurance, weight greater than 150 kg with flight altitude above 30000 ft.

During the work presented, the assessment of the Capabilities & Performance of various UAV systems has been made starting from some reference configurations mainly derived from Heliplat and CAPECON projects **[1-3]** and used here only as a performance baseline of the different UAVs categories.

HALE (High Altitude Long Endurance UAVs):

Four types of HALE UAVs has been considered as reference configurations:

- *The solar powered HALE UAV (Politecnico di Torino Prof. Romeo design)* **[1,2]:** This configuration has been designed for very high endurance, high altitude and low weight payload capabilities (140-150Kg – Available power: 1500 W). This system has a reduced number of climb and descend events, which is important when considering interference with the aviation traffic. The solar energy propulsion system reduces or even eliminates the need for the conventional fuel propulsion concept. The solar HALE configuration due to the high efficiency but low thrust has a takeoff length of about 804 m and a landing length of 420 m (without applying brake force). A rate of climb of about 2.4 m/s at maximum power condition was estimated in the preliminary analysis. Due to the regenerative electric propulsion system the solar configuration has a very long endurance (about 6 month) and low noise emission. The loiter, minimum, and maximum velocity at the operative altitude of 17 km are respectively of about 25 m/s, 18 m/s and 36 m/s.
- *The OBW 02 UAV (ONERA Design* **[4]***)* has been designed to fulfill performance mainly defined by a nominal flight profile which the main parameters are the endurance in operation (24h), the corresponding operational ceiling (60 000 ft), minimum time to climb to the maximum ceiling of airliners, typically close to 45 000 ft (30 minutes). This last constraint has been associated with conventional ones dealing with take off and landing distances (close to 540 m for the Balanced Field Length and 620 m for landing) in order to be compatible with the use of conventional runways.
- *PW-114 HALE BLENDED WING (WARSAW UNIVERSITY Design* **[4]***)* configuration has been designed by Warsaw University of Technology (WUT) for high endurance (>24 h), high altitude (20 km) and high payload capabilities (500-700 kg). The aircraft has high climb rate which lets to reach the operational altitude in short time (about 30min). Good take-off and landing characteristics (take-off length of about 430 m and landing of about 550 m for 50 ft obstacle) are presented by this configuration. Cruise velocity (IAS) is equal to 400 km/h for best endurance (at 15 km altitude) and 500 km/h for longest range (at 15 km altitude). Maximum cruise velocity is equal to 690 km/h.
- MODULAR HALE CONFIGURATION (IAI Design **[5]**): The Modular HALE is an aft-fuselage mounted, twin turbofan configuration. The wing is a high AR wing with a reference area of 38m^2. The wing span is 30m, length 12.67m and height 4.09m that means a hangar request similar to the Airbus 320 size. It has two interchangeable non-structural payload bays, one up and one down. A T tail configuration has been selected in order to avoid tail-engines interaction. Take off and Landing distance are respectively 639m and 1094 m respectively. Endurance as well operation radius are different for the four versions (A,B,C,D) that consider different payloads and fuel weights. Aircraft is equipped with two Williams/Rolls-Royce FJ44 located on the fuselage (noise<74dB). The Designer goal is to be able to deploy the UAV in less than 24 hours.

MALE (Medium Altitude Long Endurance UAVs:

Reference configuration is:

- *PW-104 PISTON MALE CONFIGURATION (WARSAW UNIVERSITY Design* **[6]***)* this configuration has been designed for long endurance (>24 h), medium altitude (4 km) and medium payload capabilities (217 kg). The climb rate (about 7m/s at sea level) lets to reach the operational altitude in short time (less than 15 min). It presents good take-off and landing characteristics (take-off length of about 415 m and landing length about of 580 m, considering 50 ft obstacle). Cruise velocity (IAS) is equal to 130 km/h for best endurance (at 4 km altitude) and 152 km/h for longest range (at 4 km altitude).

Maximum cruise velocity is equal to 280 km/h. The aircraft has good manoeuvrability Compact configuration with two tail beams facilitates installation of different payloads in payload compartments. Thus it is easily possible to change the UAV mission.

Rotary UAVs

- AGUSTA-EUROCOPTER Design **[7]**:

Two type of rotary configuration has been considered: 1)Single Main Rotor (Developed by Agusta) 2) Co-axial Rotor (Developed by Eurocopter). First configuration is a conventional helicopter configuration with known and proven technology and

therefore relatively low cost both in development and maintenance. The second Configuration is more complex and challenging for small rotorcraft, but has performance advantages over the conventional single main rotor configuration. The Single Main Rotor configuration is more efficient for UAVs with a MTOW below about 270 kg and the Co-axial Rotor configuration is more efficient for UAVs with a MTOW above 270 kg.

Small UAV

- Politecnico di Torino (Prof.Romeo Design **[8]**)

In order to demonstrate some critical technologies and applications, the flying model (SESA E-Plane Model) has been built at Politecnico di Torino. The flying model SESA was built, within the EC funded project CAPECON. The starting model was modified by replacing its combustion propulsion system with an electric one including a single brushless motor and a NiMh battery system. The structure is realized using glass-fibre reinforced plastic and carbon-fibre composite materials for wing box. Payload capabilities are in the order of 5-6 kg. Within the EC funded project TANGO, the NiMh batteries have been substituted by rechargeable LiPo batteries, mainly utilized during the take-off phase. A new wing with span of 7m was manufactured and 2 square meters of thin high efficiency (21%) mono-crystalline silicon arrays have been bonded over the wing skin; during the level flight the needed power is being achieved from the solar cells system covering the wing. All the structures were designed according to EASA-VLA to withstand a limit load of 3.8.

In order to evaluate and compare the aircrafts selected, the performances criteria shown in Table 2 have been chosen **[9,10]**. Some of these criteria are exclusive like “Vertical takeoff”, “HOGE” and the “All weather capability”. The value for these characteristics can be 1 if the aircraft is able to perform them or 0 if not. Other characteristics are dimensional characteristics that give a high score as higher they are (Range, Endurance, Range of controllability, Maximum velocity, Climb rate, Ceiling and Payload weight). Finally others criteria are dimensional characteristics that give a high score as lower they are (Maneuverability, Minimum velocity, Landing length: Take off length, Noise on flight, Minimum altitude, Loiter velocity).

In Fig 1 are reported the collected requirements for the “Oceanic Observation mission”.

Table 2

Capability	Description
Vertical Take Off	capability to perform vertical take off and landing typical of helicopter.
HOGE	capability to perform hovering out of ground effect typical of helicopters.
All weather capabilities	ability to flight in every adverse weather conditions.
Range	maximum flight range to perform the mission expressed in meters
Range of controllability:	distance of the UAV from the ground control station
Maximum velocity:	maximum velocity that can be reached in level flight expressed in m/s
Climb rate	the altitude expressed in meters that can be achieved in one second, it express the ability to reach the desired altitude quickly
Ceiling	maximum flight altitude in meters
Payload weight	maximum admitted weight of the operative payload expressed in kg.
Manoeuvrability	ability to perform short range manoeuvres, the value is the turn radius expressed in meters.
Minimum velocity	minimum flight velocity in level flight before stalling in m/s.
Landing length	distance necessary to land

	expressed in meters.
Take off length	distance necessary to take off expressed in meters.
Noise on flight	noise produced during the flight operations caused by engines and rotors expressed in dB.
Minimum altitude	Minimum flight altitude correlated to the airspace categorization.
Loiter velocity	cruise speed on level flight expressed in m/s. a suitable aircraft should have a loiter velocity as close as possible to the optimal value required by the specific mission.

2.3 Analysis of capabilities versus user needs through a multi-criteria analysis.

The algorithm implemented is based on inputs from performances characteristics of the aircrafts (J_i) and requirements from mission typology, it analyze each aircraft considering some weight indices given to the missions (W_i) and shows a final score for each aircraft and for each mission excluding those UAVs that are not able to complete the specified mission. The comparison algorithm takes in consideration the gap between the lowest and the highest value of single requirements and considers how near is the aircraft from the best value, then it applies the weight indices required by specified mission and give the global score (Fig 2).

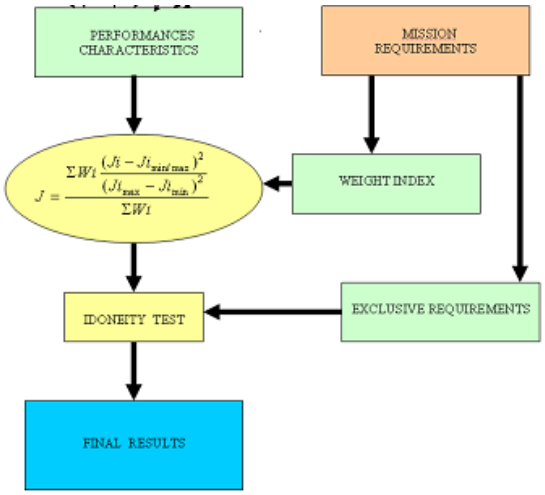


Fig. 2 Multicriteria analysis

3 Results For Selected Missions

Several scenarios in each TANGO thematic has been considered for the analysis.
Below are shown results of some scenarios related to few thematics of the five main scenarios considered during the project TANGO (Fig 3).

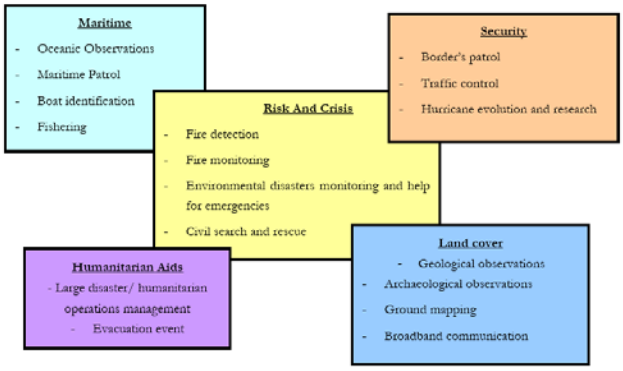


Fig. 3 Thematic and Scenarios considered in the analysis

Maritime Thematic

- Oceanic Observations

The interests of these types of missions mostly involve superficial determination of the water temperature and masses movements to different temperature to evaluate the consequences on the continents climate; Monitoring and check of some icebergs routes, along the oceans, which come off the poles; Definition and identification of the routes along which some protected species move (whales, turtles, etc); Identification and check of some fishing zones to exploit and to assign them; Movements, formations and intensity study of the waves which form themselves in the oceans and their impact along the coasts.
In this case the mission main objective is to monitor phenomena with slow variations on time so a long endurance and range as well the possibility to fly at very low speed are very important criteria. Capacity to reach the operational scenario in a very short time (especially for fishering and maritime traffic control) is another key issue that require a high value of maximum speed and manoeuvrability.

The HALE UAVs seems to be the best candidates for this type of missions (Fig 4).

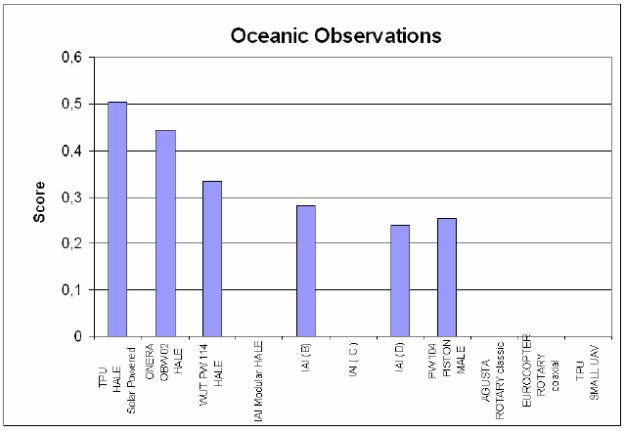


Fig. 4 Example of Results for Oceanic Observations

- Maritime Patrol

Marine surveillance and tracking missions have the main purpose of detection and localization of sea moving targets (ship imaging and classification, man in life raft, oil tankers…); sea pollution monitoring: detection of oil spills, position, thickness and extent of the oil surface.
The aircraft would fly from one sea region to the next and track targets (different ships categories) or loiter at speeds of about 50m/s in a region for a period of time. A long endurance not less than 8hr as well long range mission not less than 1000km are the two main mission parameters. The vehicle management system should allow for direct control or flight path re-direction from a ground station so a high manoeuvrability is requested for the mission. Payloads will consist of various sensors including detection and tracking of ships. Sensor data from the UAV must be available real-time. OTH network communications for command and control of the UAV are required. For these reasons mission require a high payload weight capacity not less than 90kg and better if it is higher.
Also in this case the use of very long endurance HALE UAVs seems the best solution. Results are shown in Fig 5.

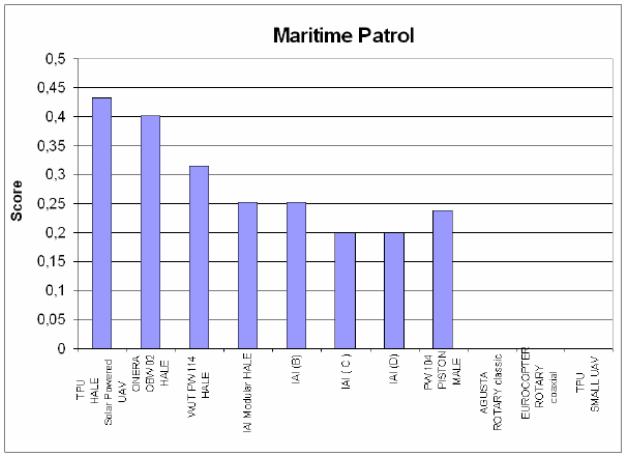


Fig. 5 Example of Results for Maritime Patrol

- Boat Identifications

This mission is an additional support for the previous mission and may include the deployment of a platform in closer proximity to suspicious vessels, the HOGE capability could give a contribution to the success of the mission but is not an excluding requirement. The platform would be released low altitude less than 5000m and in some case a very low fly altitudes is requested so the capability to fly at very low altitude is weighted with an high value. The mission require low turn radius that means high maneuverability in order to fly around the vessel under suspicion. The analysis results shown high rates for Rotary and for small UAVs (Fig 6).

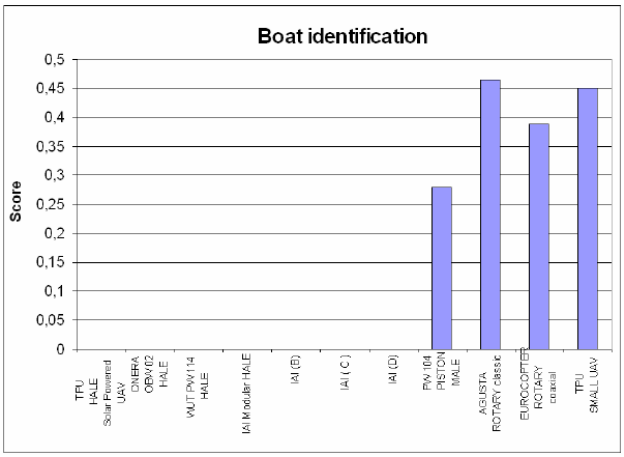


Fig. 6 Example of Results for Boat Identification

Risk & Crisis thematic

- Fire Detection

The objective is a continuative 24 hours vigilance of the territory, to identify fires and to help the turning off. The platforms must have an endurance capability of 24 to 72 hours, Flights emanating and terminating at field locations would require aircraft with high crosswind and gust limits that means a high maneuverability (weight=5). The platform required for this mission may need to have a high range capability greater than 300km and transit from one fire location to the next around the country. Launch and recovery would occur at either field locations or dedicated fire fighting airfield locations (take off and landing capability=4).

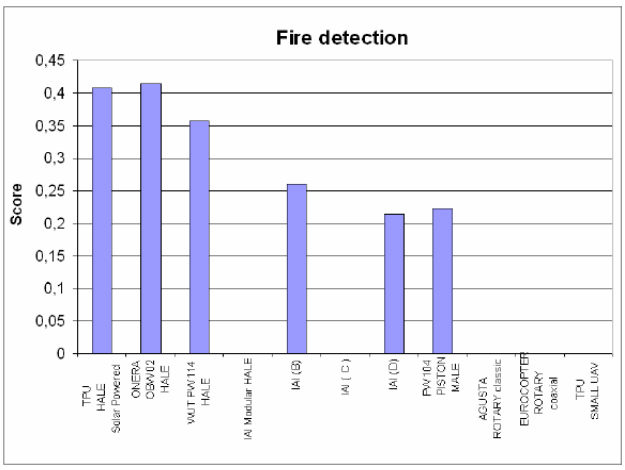


Fig. 7 Example of Results for Fire Detection

Both HALE and MALE categories seem to accomplish the mission with good results (Fig 7).

Security thematic

- Borders Patrol

The UAV flies along the border conducting surveillance, and detecting cross-border activity and informs personnel at the Border Patrol Regional Command Centre. An area supervisor dispatches resources and re-tasks the UAV(s), as appropriate. Once on station, the UAV serves as a communications node and identifies border violators. Desired attributes for the platform include multi-day endurance (>24h weight=6), and the ability to carry multiple payloads (>40kg weight=5). High altitude flight is desirable (weight=7) as it increases platform's covertness. The ability of the sensor package to "steer" the platform to a target of interest is highly desirable (maneuverability=5).

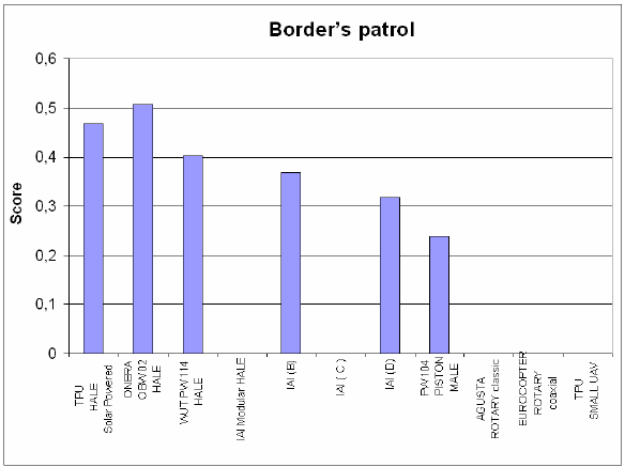


Fig. 8 Example of Results for Fire Detection

The results show that the Solar powered and blended HALE configurations have the best result for this mission (Fig 8)

Land Cover thematic

- Archaeological observations

The main interest of this kind of mission is to check for archaeological areas already existing and evaluation of their conservation state.
The possibility to have a simple take off and landing as well as the hovering over the site of interest are very important parameters for this kind of mission (HOGE=5 Vertical take off=5 landing and take off length=7). In order to increase the quality of the detection the possibility to fly at very low speed is also an important parameter (minimum speed=7).
The analysis shows that, despite all the platforms may accomplish this mission, the small UAVs and Rotary UAVs, in particular, seem to be the best candidates (Fig 9).

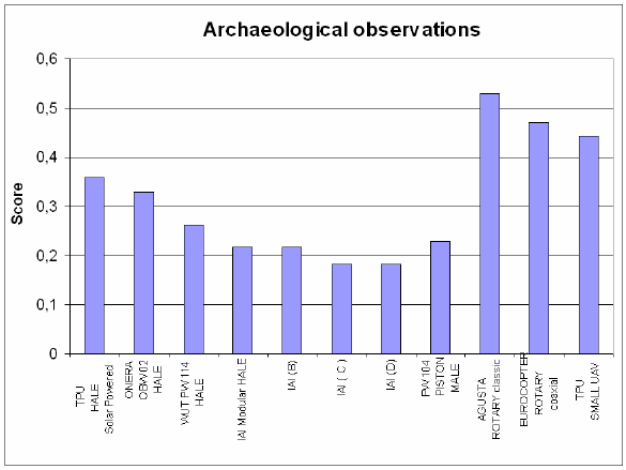


Fig. 9 Example of Results for Archaeological observations

Humanitarian Aids thematic

- Evacuation Event

The goal is to provide images and voice communication during evacuation events. For emergency evacuation, the communications system must be rapidly reconfigurable. For example, mobile ground stations may have to be delivered to a disaster area because existing base station may lose power or connectivity. The ground stations and any UAVs deployed in the area must be able to quickly establish communications and not controlled by a global authority but by a local entity. A UAV that not need a complicated ground control station and able to have a high maneuverability (maneuverability=9), capable to fly at a low speed (minimum velocity = 8) with a good rate of climb (rate of climb =6) and for at least 5 hr could give a great contribution to the operation.

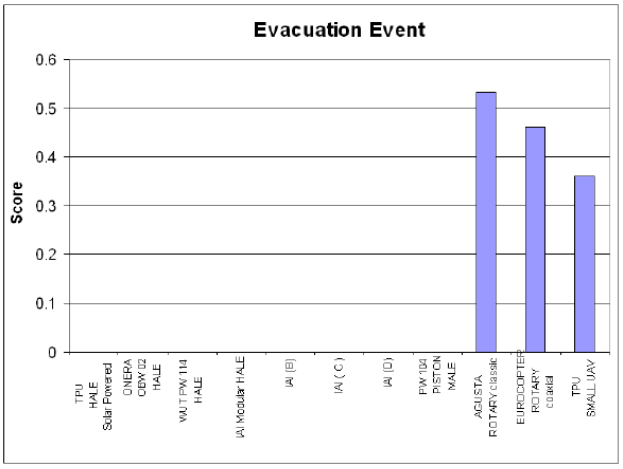


Fig. 10 Example of Results for Evacuation Event

Because several configurations are able to perform almost all the applications, interpretation of these results is very dependent on the preferences of the user. So, it has been decided to keep as much information as possible when giving the primary results, instead of degrading information making interpretations. Results of the preliminary analysis are only a first support and should be interpreted directly by the user, to allow him to make his own trade-offs. Some important additional criteria should be considered during the selection of a UAVs.

A key technology which supports all of the future missions is, for example, the ability to transmit data over the horizon (OTH).

Three classes of Satcom data relay payload have been defined during TANGO Project **[9,10]** and described in terms of mass, volume and power requirements. Selecting Platform for a specific scenario it is important to take into account which are the data link possibilities:

CLASS1: INMARSAT 4L BAND terminal: implementation possible for all reference UAVs excluding Small Solar UAV.

CLASS2: Ku or Ka Satcom terminal: implementation possible for all reference UAVs excluding Small Solar UAV. The Solar HALE UAV is at the limit of its capability due to high power request from amplifier.

CLASS3: Optical Terminal: implementation possible for all reference UAVs excluding Small Solar UAV. The small size of the Small Solar UAV preclude the use of high bandwidth geostationary satellite systems. The only systems capable of being installed are those which have small antennas, and thus it is the Low Earth Orbit (LEO) systems (Iridium/Globalstar). An OTH scenario is possible but with link at 2200–9600 baud.

Another important factor in order to choose the proper UAV for a specific mission is the cost. Preliminary results indicated the following costs for the different UAV classes: combustion engine HALE are between 1000 to 3000 €/FH, Solar configuration presents an asymptotic value of about 1000 €/FH , MALE are in the range of 300-900 €/FH and Rotary configuration about 500-1400 €/FH with the coaxial configuration slightly more expensive.

Small solar UAV is the less expensive platform with a cost range of about 100-200 €/FH.

Finally the choice of a suitable UAV is strongly dependant on the development of a UAS SRM (Safety Risk Management) framework, supporting regulators and applicants through provision of detailed guidelines for each SRM step to be conducted, including 1) system description, 2) hazard identification, 3) risk analysis, 4) risk assessment, 5) risk treatment.

The purpose is that all potential risks of the newly proposed UAS operations are controlled so that the existing safety level does not decrease as compared to the current aviation system with manned aircraft only.

4 The use of the SESA UAV in the TANGO demonstration.

Within the project TANGO a flight demonstration by the scaled Solar Powered UAV (SESA) was carried out to monitor boats and ships prior to boarding in port **[10,11]**. Demonstrations would integrate satellite telecommunication solutions with on-going GMES developments in the framework of fisheries management. Inclusion of UAVs in the global relay infrastructure enables quasi real time and continuous access to dedicated zones for monitoring or surveillance. SESA can be used, in cooperation with Iridium satellite system, as a reconnaissance aircraft sending real time images and videos to the ground control station. The flight test was carried out starting the flight from S. Giorgio airfield (5km south Lido di Tarquinia, about 10 km north of Civitavecchia. In Fig 11 shows the mission architecture. During the demonstration the SESA UAV was equipped with a) video camera heading forward, for see & avoidance purpose. b) photo/video camera heading downward; the video images are continuously wireless sent to the ground control station c) An autopilot system was installed on board for an autonomously flight up to 40km of distance by a highly integrated data acquisition, processing and control system. d) Satellite based Communication System was installed onboard for the Iridium satellite communication network aiming to a BLOS flight.

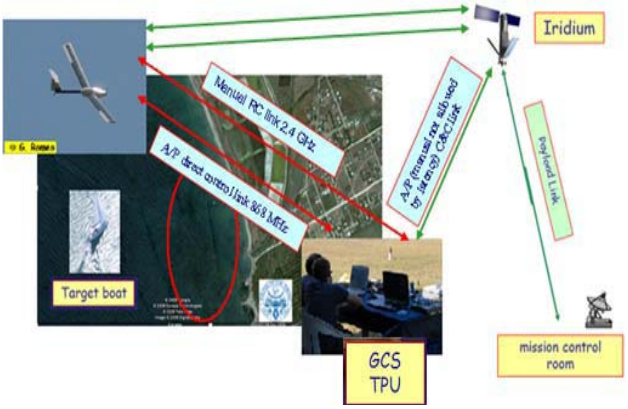


Fig. 11 Mission Architecture

The experimental tests validated main technologies for high altitude very long endurance flight.

5 Conclusion

Paper shown that today's technology allows to define and produce different types of UAVs, for a high number of potential GMES applications.
A multi-criteria analysis, which compares the mission requirements with the performance of different UAV configurations has been developed as a decision support for future end-users. The limits of the proposed algorithm mainly are the difficulty of taking into account additional criteria such as costs and level of risk introduced by a specific UAS configuration.

The ability to use, on the maritime surveillance scenario, a real configuration of Small UAV was demonstrated by means of an experimental campaign.

ACKNOWLEDGEMENTS

The authors acknowledge the important contribution of European Commission by the funding programmes:

- HELINET: "Network of the Stratospheric Platforms for Traffic Monitoring, Environmental Surveillance and Broadband Services". EC 5° FP–2000-03; WP Leader

of the HeliPlat Design and Manufacturing: Prof. G. Romeo

- CAPECON: "Civil UAV Applications & Economic Effectivity of Potential Configuration Solutions". EC 5° FP. - WP Leader of 3 HALEs Design: Prof. G. Romeo.
- ENFICA-FC: "ENvironmentally Friendly Inter City Aircraft powered by Fuel Cells. EC-6FP-2005-30779-AERO-1, Ott. 2006-09. Coordinator: Prof. G. Romeo
- TANGO: Telecommunications Advanced Networks for Gmes Operation (SIP5-CT-2006-030970-2006-09).Coor. Eads Astrium.

The authors acknowledge Eng. M. Pacino, Eng M.Francesetti, Dr. G.Correa and Dr. F.Borello for their important contribution during all the experimental part of the work.

References

[1] Romeo G., Frulla G., Cestino E., Corsino G., "HELIPLAT: Design, Aerodynamic and Structural Analysis of Very-Long Endurance Solar Powered Stratospheric UAV", *Journal of Aircraft*, Vol.41, No.6, 2004, pp. 1505-1520.

[2] Romeo G., Frulla G., Cestino E., "Design of high-altitude long-endurance solar-powered unmanned air vehicle for multi-payload and operations", *Part G Journal of Aerospace Engineering*, Vol.221, No.18, 2007, pp. 199-216.

[3] Tsach S., Peled A., Penn D., Touitou D.H., "The CAPECON Program Civil Applications And Economical Effectivity Of Potential Uav Configurations", *AIAA 3rd Unmanned Unlimited Technical Conference, Workshop and Exhibit*, Chicago-Illinois, 2004.

[4] Goraj Z., Frydrychewicz A., Hermetz J., Le Tallec C., "HALE UAV platform optimised for a specialized 20-km altitude patrol mission", *24th ICAS Congress*, Yokohama-Japan, 2004.

[5] Tsach S., Chemla J., Penn D., "UAV Systems development in IAI – Past, Present & Future", *2nd AIAA "Unmanned Unlimited" Systems, Technologies, and Operations*, San Diego, California, 2003.

[6] Goraj Z., "PW-125 MALE UAV design project developed in Warsaw University of Technology", *AIAA 3rd Unmanned Unlimited Technical Conference, Workshop and Exhibit*, Chicago, Illinois, 2003.

[7] Mouritsen S., "Rotary Wing Requirements Technical Report", CAPECON Report ID3/3, 2003.

[8] Romeo G., Pacino M., Borello F., "First Flight of Scaled Electric Solar Powered UAV for Mediterranean Sea Border Surveillance & Forest Fire Monitoring", *Aerotecnica Missili e Spazio* Vol.88, No.1, pp.8-19,2009.

[9] Cestino E., Pacino M., "Report on UAV & associated role/constratints for GMES application Part 1", TANGO-2008-D3.2/2-Part1 Technical Report (approved by Romeo G.& Defever S.).

[10] Cestino E., Pacino M., "Report on UAV & associated role/constratints for GMES application Part 2", TANGO-2008-D3.2/2-Part2 Technical Report (approved by Romeo G.& Defever S.).

[11] Indregard M., Lefort T., Gallagher R., Santos P., Borges P, Aznar V.,Romeo G., Cestino E., "Scenario definition for fisheries management", TANGO-2008-D5.2.1-1 Technical Report.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

SAMONIT - Polish Research Program For Development of an Advanced Aerial Surveillance System

Zdobyslaw Goraj[*], **Miroslaw Rodzewicz**[*], **Wojciech Grendysa**[*], **Marek Jonas**[*], **Cezary Górniak**[*]
[*]*Warsaw University of Technology, Poland*

Jaroslaw Hajduk[#]
[#]*Air Force Institute of Technology, Poland*

Keywords: *UAV, design, flight testing*

Abstract

This paper is focused on design, optimisation and testing of a mini UAS called SAMONIT. The main goal of the activity within this project is to design an UAV system for long endurance border surveillance and monitoring having high level requirements in term of reliability and safety and being affordable for potential users in terms of cost. The design process is treated as an interdisciplinary approach, and includes a selection of thick laminar wing, aerodynamic optimisation of swept wing, stability analysis, weight balance, structural and flutter analysis, many on-board redundant systems, reliability and maintainability analysis, safety improvement, cost and performance optimisation. A baseline configuration (straight wing with V-tailplane placed behind the main wing) will be used as a reference for more advanced layouts, especially swept tailless configuration, considered as the goal configuration in two versions – with pulling and pushing propellers. Comparisons between these 3 configurations will be used to show expected advantages and possible drawbacks. Full scale free flight airplane testing is in progress, both in manual and automated modes.

1 Introduction

The main goal of the activity within this project is to design an UAV system for long endurance border surveillance and monitoring [1], having high level requirements in term of reliability and safety (expected by other users of national airspace and people on ground) and being affordable for potential users in terms of cost. One such system (2 airplanes + sensors + Ground Control Station) must cost no more than 200 000 euros [2]. The design process is treated with an interdisciplinary approach, and includes a selection of a thick laminar wing [3], aerodynamic optimisation of swept wing, stability analysis, weight balance, structural and flutter analysis, many on-board integrated redundant systems, reliability and maintainability analysis, safety improvement, cost and performance optimisation. Assumed endurance is 24 hours, the main aircraft parts and systems weight is as follows: composite structure - 15 kg, power unit - 6 kg, communication + navigation + flight control systems – 6 kg, emergency parachute – 2 kg, payload –14 kg, fuel for 24 hours flight – 24 kg. Wing area is equal to 1.4 m^2, wing aspect ratio is 11.4. Stall speed at SLF is 24 m/s, typical landing speed is 20 m/s. Typical lift coefficient

during patrol is equal to 0.46, aerodynamic efficiency is 15, power coefficient (required power over the available power) during typical patrol is equal to 0.44. The airplane is powered by two KOMATSU gasoline engines, model G800BPU, 8 bhp each, air cooled two stroke cycle opposed cylinder type. This paper is focused mainly on the design aspects, a selection and refinement of its external layout – choice between straight wing with V_tail configuration and tailless, swept wing configuration and is based on the author's experience gained during former, UAS related design activity, [5-12]. Design details, technology of manufacturing processes encompassing both negative moulds and positive aircraft components, and progress in production of prototype is shown and discussed.

2 Design layout

Tailless, swept wing configuration for a long endurance surveillance mission was adopted, mainly due to its aerodynamic characteristics and utility properties [13-15]. Such a layout has improved aerodynamic efficiency due to a lower wetted area, also possessing less parts that could be exposed to damage, especially during a difficult landing. Self-stable, 15% wing section was intentionally designed [16] to increase negative pitching moment by inclination of trailing edge up.

Chord of the wing is constant along the span. Wing tip is twisted down with respect to the wing central part on -3° to move the stall from elevon area into the wing central part. The so-called end plates are mounted at the wing tips and they play the role of vertical stabilizers. These plates have rudders, deflected outside of the wing only (the reason that rudders are not deflected inside the wing is to avoid the interference between rudders and elevons).

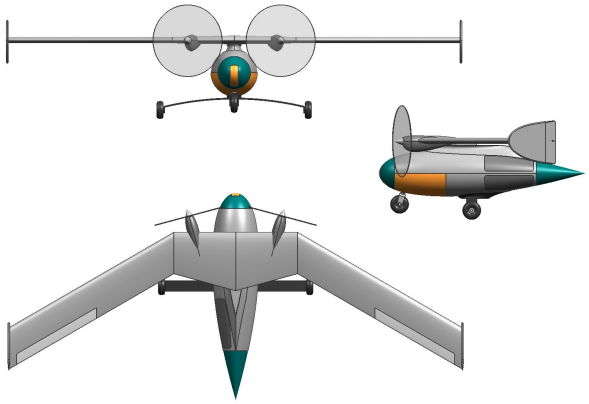


Fig. 1 ***Swept wing tailless configuration – version with wing tip plates [13-15]***

Fuselage plays the role of container, and holds different on-board systems – navigation, communication, surveillance and fuel. It was assumed that initially FLIR system will be used. In second stage of aircraft development also a SAR system can be hosted into the fuselage, so the fuselage (container) dimensions are dictated by one of such systems available on the marked (SANDIA mini SAR,[17]).

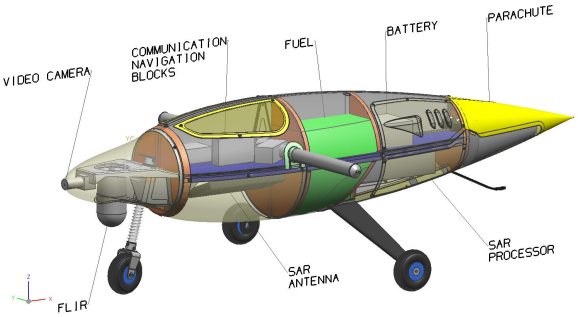


Fig. 2 ***Fuselage - container for sensors, electronics and fuel tank***

On-board SAR system [17] is optional – either can be selected and mounted, or not, and therefore its centre of gravity (CG) is located in order not to influence the whole aircraft's CG location. A SAR antenna is relatively large and it is the reason why the fuselage cross-section is also large. Fuselage overall shape was designed and optimised to minimise its drag coefficient. It was found that a streamline (droplet) shape will fulfil this requirement well. The large fuselage cross-section, in a natural way, offers a

possibility to design a high capacity fuel tank, which can be located ideally at the aircraft centre of gravity (CG). The rear part of the fuselage is used as a container for a recovery parachute. Such a recovery system offers the possibility to land in a difficult environment (where vertical falling is unavoidable), and also increases the safety factor in standard operation, when in an emergency. This two-stage parachute system consists of smaller breaking parachute (a piloting chute) and bigger recovery parachute (to be opened after a deceleration phase). This COTS system is routinely used in para-gliding and produced in Poland by AirPol-Legionowo ltd.

Tailless configuration offers many advantages over the orthodox configuration, however it is still aerodynamically difficult and can reveal some unexpected phenomena (trim issues, stability etc.). It is the reason why it has been decided to develop a second configuration with a straight wing and V-tail control surface placed behind of the main wing, called further baseline configuration. It was assumed that both configurations (tailless and baseline) are unified as much as possibly (they have the same wing aspect ratio, wing loading, power-to-weight ratio, on-board systems, fuselage etc.). The wing section of the baseline configuration, the Wortmann profile (FX66-17A-II-182/20), was selected because in this case, the pitching moment can be controlled by elevator's deflection and there is no need to modify the pitching moment through the change of profile camber in the TE region. The baseline configuration will be used as a reference for different comparisons between both configurations and for evaluation of expected advantages and possible drawbacks. Attachment of the wing on a pylon above the fuselage will facilitate an easily change of this wing (from swept to straight and vice versa) using the same fuselage, landing gear system etc.

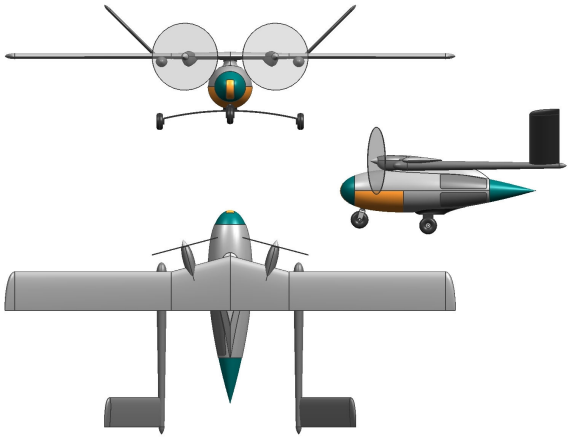


Fig. 3 Fuselage - container for sensors, electronics and fuel tank

Additionally, to reduce the unavoidable risk connected with first flights of unconventional configuration (tailless one), it was decided to design, manufacture and test in flight, a scaled (1:2) tailless configuration model. These flight tests have shown that the configuration initially proposed for the tailless layout must be changed. It appeared that during the take-off run a normal reaction force acting on the front landing gear was too high, mainly due to nose-down pitching moment generated by highly located engines. Elevons were not able to produce the counterbalancing, nose-up pitching moment and the accelerating model turned over on the runway. After the necessary redesign process the new layout of tailless configuration are shown at Fig.4-6 and baseline configuration at Fig.7.

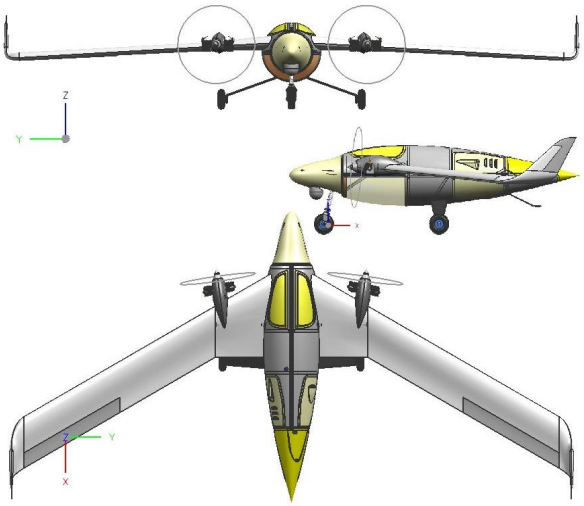


Fig. 4 Swept wing tailless configuration – current version with the wing moved down and wing tips optimised for lateral dynamic stability

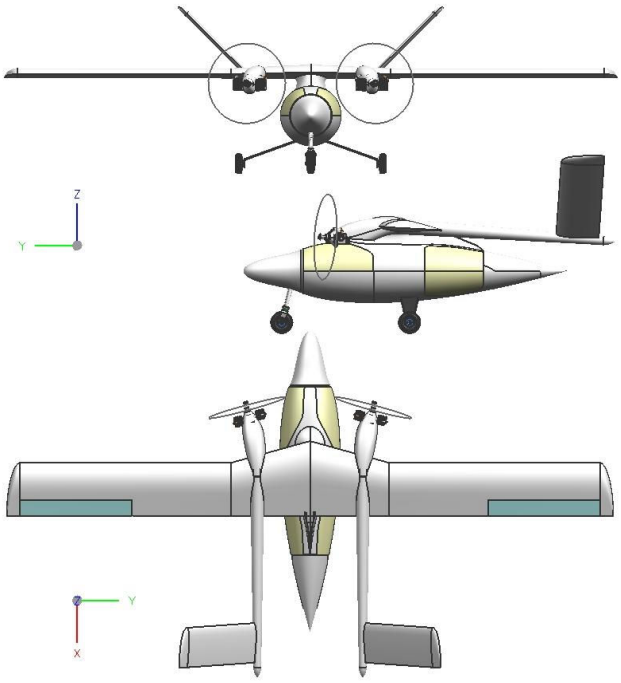


Fig. 5 Baseline configuration – redesign to be unified with last, modified version of the swept wing configuration



Fig. 6 Swept wing configuration with pushing propellers –flying pre-prototype no 3

Pre-prototypes no 1 (Fig.5) and 2 (Fig.4) are designed to be operated from typical concrete runways. Pre-prototype no 3 (Fig.6) could be operated both from traditional runways and from catapult and can be captured with a net, Fig.7. Pushing propellers are protected against damage when aircraft is decelerated by a net, Fig.7.

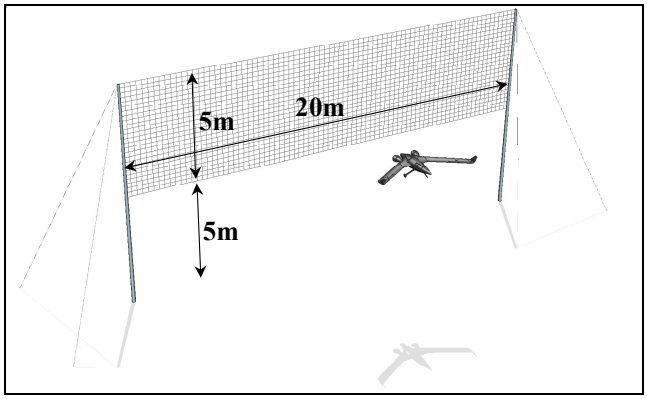


***Fig. 7* Aircraft being recovered in the net**

3 Selection of power unit

The two-engined configuration was selected to increase the safety level (according to statistics gathered in IAI, one of the most frequent critical failures of UAVs was engine dysfunction [ ] and redundancy can increase the so-called MTBL (Mean Time Between Losses). Engines are placed along the wing's centreplane, close to the aircraft's vertical plane of symmetry. Moreover, they are skewed at 6° outward with respect to vertical plane of symmetry to decrease yawing engine torque when one engine is not working. In order to decrease the level of vibration (important for stable usage of sensors) the two cylinder engines (KOMATSU) were selected. Another desired feature of these engines, especially required for long endurance mission, is their high mean life. The design philosophy and production technology, of the main components of this type of engines, are based on experience gained from the family of engines produced for pumps widely applied in civil engineering. Therefore they are likely to able to work for several dozens of hours with no failure. With respect to earlier requirements of power access, the available power was increased to 8 hbp and therefore it will make it possible to increase MTOW in the future, if necessary. A further reduction of vibration transferred from engine into the aircraft structure is possible through the flexible friction dampers of an engine suspension system. The special dumpers developed by KOMATSU were also used for

noise reduction. The engine spark device (originally supplied with the engine by KOMATSU) together with an additionally designed and manufactured coil was used as a current generator.

4 Components of aircraft structure

The structure is designed from carbon fibre. The main parts of fuselage consists of horizontal and vertical plates, attached to a set of frames and with a few easily disassembled covers. Plates are designed for a sandwich construction. The wing spar in the fuselage area is made from a steel, with a circular cross-section (ϕ 50 * 48), being hardened for high strength. The integral fuel tank, in the fuselage central part, has 49 litre capacity.

Wing (and tailplane) are built from the same materials as the fuselage. Straight wing consists of a central part and two outer wings. Two tailbeams are attached to the wing central part and ended with V-tails. Central part of the wing is attached to the fuselage through a pylon and therefore aerodynamic interference is reduced. Wing has a single spar only, Fig.9-11.

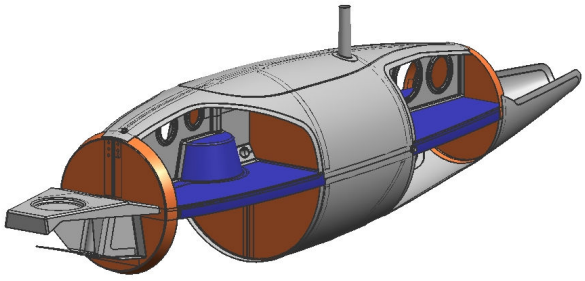


Fig.8 Design structure of sensor container - lower and upper covers disassembled for inspection

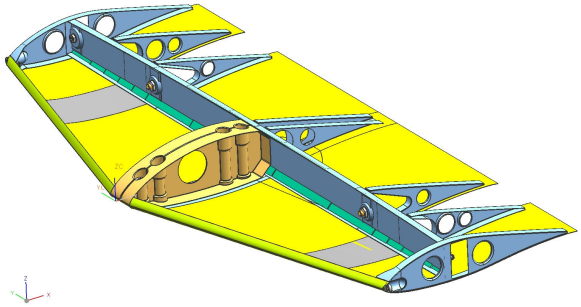


Fig.9 Design details of the straight wing central part

Swept wing is divided into left and right parts, Fig.10. Both parts are slid over the pipe spar standing out from the fuselage. Fast joining can be performed screwing down two screws, one per each wing, Fig.10-11.

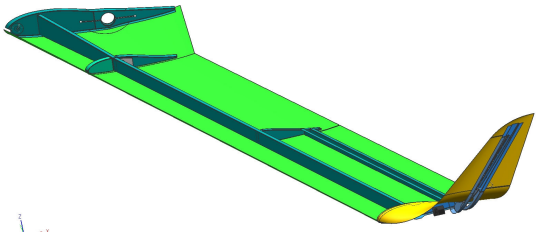


Fig.10 Design details of the swept wing

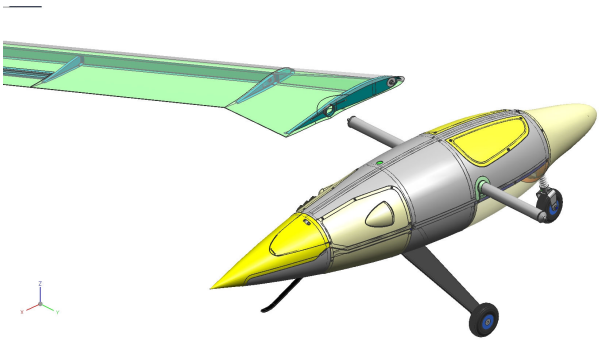


Fig.11 Method of join between swept wing and fuselage

The aircraft with straight wing has the V-tails, mounted on tailbeams. The method of attachment between tailbeams and the wing central parts, and between tailbeams and the engine's mount, is shown at Fig.12.

Main landing gear was designed as a composite spring beam, Fig.13, and the front landing gear is of telescopic type, external spiral spring and internal pneumatic damper.

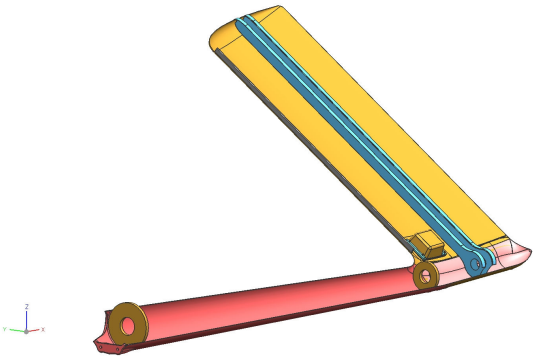


Fig.12 Baseline configuration - design details of right V-tail



Fig.13 Main landing gear

5 Parachute recovery system

This system is important to decrease the probability of a ground collision – the case in which either innocent people could be injured or/and infrastructure could be damaged. The system is based on the well-known parachute produced by AirPol ltd (KSKY-36,7), widely used by para-gliding pilots (standard load equal to 70 kg), fully ertified. To avoid damage to the aircraft's structure due to a rapid deceleration, the opening of main parachute is preceded by opening of a breaking, much smaller parachute (area of 1 m^2), which is shown at Fig. 14.



Fig.14 Successive phases of parachute deployment

6 Summary of design features

The main design features can be summarised as follows:

- Tailless layout is the target configuration. Baseline layout is used for comparison and evaluation only;
- Aircraft design is easy for transportation on the trolley behind a passenger car and easy for quick assembling / disassembling;

- Tailless configuration with selfstable wing section and geometry offering wide range of angles of attack is resistant to stall;
- Two engines of high reliability offer increased MTBF level with high power access;
- Parachute recovery system increases safety level for people and property on the ground;
- Main landing gear made from carbon fibre, endowed with efficient breaking system, ensures short, well dampened landing. Front landing gear consisting of telescopic type external spiral spring and internal pneumatic damper, ensures good damping characteristics and protects against potential turnover on the runway;
- Wing tip plates used on tailless configuration increase wing aspect ratio, provide lateral stability and good controllability in sideslip;
- Centre of gravity (CG) does not travel with respect to aircraft during flight mission;
- The swept wing configuration was modified to fulfil stability requirements. Corrections to be done by the Automatic Flight Control System (AFCS) in terms of dynamic stability will be marginal.

7 Free-flight experiments

Taxi tests, landing runs and free-flight experiments were initiated in Dec.2009 and are still in progress. Several free-flights were performed, either in manual or automated modes. All 3 airplanes were in the air longer than 3 hours and authors have in their disposal a very wide material including logs, videos from the ground and from the 3 cameras attached to the airplane. In Fig.15 there is a trajectory seen over the terrain and in Fig.16 there is a trajectory and turning points seen on the control monitor in GCS.



Fig.15 Trajectory supervised by autopilot (turning points and flight altitudes were changed on-line when in progress)

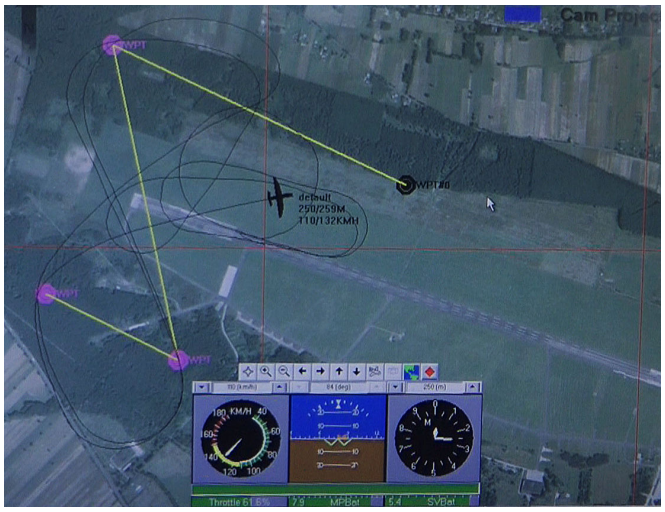


Fig.16 Trajectory and turning points seen on the control monitor

8 Conclusion

SAMONIT project, focused on the design, optimisation, manufacturing and testing of a surveillance long endurance mission aircraft, named as PW-141, is well advanced. Three versions ((1) baseline with straight wing and empennage placed after the main wing; (2) swept, tailless, tractor configuration and; (3) swept, tailless, pushing propellers configuration) are fitted with fully autonomous flight control system. Before the free flight experiments all these pre-prototypes were aerodynamically tested in ϕ5 WT. The baseline,

straight wing configuration was extensively tested in the ground vibration experiments and then numericaly checked with respect to flutter. The main goal of the whole programme is to optimise the platform and its systems to obtain the economically competitive aircraft, safe, affordable and efficient in its main long endurance, surveillance mission.

References

[1] European Civil Unmanned Air Vehicle Roadmap, www.uavnet.com.

[2] CAPECON Project No GRD1-2001-40162 ("Civil UAV Applications & Economic Effectivity of Potential CONfiguration Solutions"), Technical documents 2002-2004. V FR of European Union.

[3] Steinbuch M., Marcus B., Shepshelovich M., "Development of UAV Wings – Subsonic Designs". AIAA 2003, 41th Aerospace Sciences Meeting & Exhibit. Jan., 2003, Reno, NV.

[4] Allouche M., "Civil UAV Safety Issues – Airworthiness and Operational Certification aspects". UAVNET, Stockholm, Oct.2001, www.uavnet.com

[5] Goraj Z., "Civilian Unmanned Aerial Vehicles – Overview of European Effort and Challenges for the Future". Aviation Journal, Vilnius 2003, Aviation, Vol.VII, No 2, pp.1-18, 2003.

[6] Goraj Z., Frydrychewicz A., „Development Approach of the PW_103 - An Increased Reliability MALE UAV - Under the CAPECON Project Within the V FR of EU", Proc. Of 24th ICAS Congress, Yokohama 2004, Paper 443.

[7] Gadomski J., Hernik B., Goraj Z., "Analysis And Opimisation of a Male UAV Loaded Structure". Aircraft Engineering and Aerospace Technology, Vol.78, No. 2, 2006, pp.120-131.

[8] Goraj Z., Frydrychewicz A., de'Tallec C., Hermetz J., "HALE UAV platform optimised for a specialized 20-km altitude patrol mission", Proc. Of 24th ICAS Congress, Yokohama 2004, Paper 285.

[9] Goraj Z., Frydrychewicz A., Świtkiewicz R., Hernik B., Gadomski J., Goetzendorf T., Figat M., Suchodolski S., Chajec W., "High Altitude Long Endurance Unmanned Aerial Vehicle of a New Generation – a Design Challenge for a Low Cost, Reliable And High Performance Aircraft", Bulletin of Polish Academy of Sciences, Vol.52, No.3/2004, pp.173-194.

[10] Goraj Z., "PW-125 MALE UAV design project developed in Warsaw University of Technology". AIAA 3rd conference "Unmanned Unlimitted", Chicago, Sept. 2004, AIAA paper 6326.

[11] Goraj Z., "Design Challenges associated with development of a new generation UAV". Aircraft Engineering and Aerospace Technology, Vol.77, No. 7, 2005, pp.361-368.

[12] Figat M., Goetzendorf-Grabowski T., Goraj Z., "Aerodynamic Calculation of Unmanned Aircraft". Aircraft Engineering and Aerospace Technology, Vol.77, No. 6, 2005, pp.467-474.

[13] Goraj Z., "An Integrated System for Border Surveillance and Monitoring Based on Very Light MALE UAV". Proc. Of 25th ICAS Congress, Hamburg, Sept 2006, Paper 417.

[14] Goraj Z., "UAV platforms designed in WUT for border surveillance". AIAA Infotech @ Aerospace 2007 Conference and Exhibit. San Francisco - Rohnert Park, 7-10 May, 2007.

[15] Goraj Z., "Mini UAV Designed for Surveillance Long Endurance Mission". The 4th Internat. Symposium on Innovative Aerial / Space Flyer System. Tokyo, 14 Jan. 2008, Plenary, Invited Lecture. Proceedings of the 4th Int. Symposium, pp.11-20.

[16] Galinski C., "Wing section designed for PW-141". SAMONIT internal report (unpublished). Warsaw, May 2007.

[17] www.sandia.gov/radar.sar.html

[18] Goraj Z. et al., "Mini UAV design and optimization for long endurance mission". Proc. Of 26th ICAS Congress, Anchorage, Alaska, Sept 2008, Paper 437.

[19] Goraj Z., Chajec W., "Aeroelastic analysis of remotely controlled research vehicles with numerous control surfaces". International Journal of Structural Integrity, Vol. 2 Iss: 2, April 2011.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Dynamics-Aided Inertial Navigation for a UAV during GPS outage

Lv Pin, Liu Jianye, Lai Jizhou and Zhang Ling
Navigation Research Center, Nanjing University of Aeronautics and Astronautics, China

Keywords: *Dynamics Aided, INS, GPS outage, UAV*

Abstract

Dynamics-aided INS utilizes available control information and dynamic model to obtain UAV's motion information, and fuses theses information with INS using a filter. The integration of dynamics model with INS leads to the compensation of inertial sensors' errors, causing the improvements of navigation accuracy, especially in GPS-denied environments. Our purpose is to explore the navigation accuracy improvement brought by the UAV dynamics model aiding technique in short periods. This paper explores the feasibility of using dynamics model aided INS in an UAV which is flying in an area where GPS signals are intermittent. Bedsides, we analyze the sensitivity of navigation accuracy on the error of aerodynamic coefficients, and some simulations are designed.

1 General Introduction

Since INS is a diverging system, its error spreads over time. So INS is generally used combined with another navigation system whose error is relatively constant over time, and Global Position System (GPS) is a good choice. When GPS measurements are available, the combined navigation accuracy is usually satisfactory. However, when GPS signals are absent for some reason (e.g., environment factors or jamming), the INS will be used for navigation as a stand-alone system. In this case, position errors of UAV increase rapidly with time since INS is a dead-reckoning navigation system [2]. Especially for small UAVs, accurate INSs are bulky and power-hungry, so low-accuracy micro-electromechanical systems (MEMS) INSs are always used, and GPS outage will lead to a huge decrease of position accuracy.

A common way to circumvent this situation is to fuse other sensors with the INS. Vision navigation and celestial navigation are applicable schemes for certain types of UAVs, but both of them need corresponding sensors which require space, power and extra cost. An alternative scheme is to use information from the dynamic model of UAV, which is called dynamics-aided INS. Dynamics-aided INS utilizes available control information and dynamic model to obtain UAV's motion information, and fuses theses information with INS using a filter. The integration of dynamics model with INS leads to the compensation of inertial sensors' errors, causing the improvements of navigation accuracy, especially in GPS-denied environments. This aiding scheme can be adopted in nearly any UAV whose dynamics model is known, requiring no additional sensors and with low cost, so it has good application value.

This concept was first proposed by M. Koifman [3], it was demonstrated that INS aided by aircraft dynamic model could get better positioning accuracy than pure INS. Later Mitch Bryson discussed two different methods

(velocity and attitude aiding or acceleration and rotation rate aiding) for aiding INS using aircraft dynamic model information [4], the characteristic of two methods were shown. Dynamics-aided INS is widely used in vehicles, since constraints of the vehicle motion can also be introduced to the system (e.g., a vehicle's sideways velocity is zero) which aircrafts don't have. Recently the contribution of dynamics-aided method to INS in vehicles was tested by Simon J.Julier [5] and Itzik Klein [6]. Dynamics-aided INS was applied to underwater vehicles by Øyvind Hegrenæs [7~9].

Our study is carried out in the following scenario: an UAV is traveling in an area where GPS signals are intermittent (e.g., urban canyon), equipped with an INS and a GPS receiver. During the UAV's flight, GPS signals are unavailable at some time. Our purpose is to explore the navigation accuracy improvement brought by the UAV dynamics model aiding technique in short periods.

The rest of the paper is organized as follows: section 2 describes the inertial components error and the INS error equations; section 3 describes the UAV aircraft dynamic model used to aid INS; section 4 describes the filter structure. Section 5 presents results of simulation that demonstrate the effect of dynamics-aided INS. Section 6 presents the conclusions.

2 INS Error Model

The accuracy of INS is greatly affected by the error of inertial components, which includes gyros and accelerometers. Gyro error is supposed to consist of bias, white noise and first order markov noise: bias is constant error; white noise represents noise of short time characteristics; first order markov noise represents noise of long time characteristics. So the raw measurement of a gyro is given by Eq. (1).

$$\omega_m = \omega_n + \varepsilon_b + \varepsilon_w + \varepsilon_m \qquad (1)$$

ω_m is the measurement value of gyro, ω_n is the nominal value of gyro, ε_b is bias , ε_w is white noise, ε_m is markov noise. The raw measurement of a accelerometer is given by Eq. (2).

$$f_m = f_n + \delta_b + \delta_m \qquad (2)$$

f_m is the measurement value of accelerometer, f_n is the nominal value of gyro, δ_b is bias , δ_m is markov noise.

In this paper the following coordinate frames.are used: Inertial frame (i-frame), Earth Centered Earth Fixed (e-frame) frame, North-East-Down (NED) frame (n-frame) and Body frame (b-frame). ω_{ib}^b is gyro's output, which means angular velocity vector of b-frame) with respect to i-frame resolved in the former. f_{ib}^b is accelerometer's output, which means specific force vector of b-frame) with respect to i-frame resolved in the former.

From ω_{ib}^b and f_{ib}^b , attitude and position can be calculated through INS algorithm. Figure.1 shows the basic principle of INS.

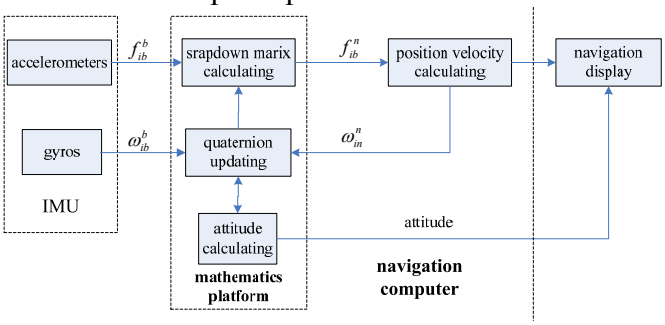


Fig.1 The principle of RSINS

In order to compensate INS error, it is necessary to develop INS error model. Error model as Eq. (3) is adopted in this paper. l is the latitude, L is the longitude, h is the height above the earth surface. $\delta l, \delta L, \delta h$ represents their errors. $v = [v_E, v_N, v_U]$ is the vehicle velocity, $\delta v = [\delta v_E, \delta v_N, \delta v_U]$ represents their errors. $\delta\alpha, \delta\beta, \delta\gamma$ represents platform angle error. Ω is earth rate. R is earth radius. $\delta\omega = [\delta\omega_x, \delta\omega_y, \delta\omega_z]$ is the gyro error, $\delta f = [\delta f_x, \delta f_y, \delta f_z]$ is the accelerometer error. C_{11}, C_{22}, C_{33} represents scale factor error of three gyros, $C_{12}, C_{13}, C_{22}, C_{23}, C_{31}, C_{32}$ represents misalignment angle of three gyros. D_{11}, D_{22}, D_{33} represents scale factor error of three

accelerometers, $D_{12}, D_{13}, D_{22}, D_{23}, D_{31}, D_{32}$ represents misalignment angle of three accelerometers.

$$\begin{bmatrix} \delta\dot{\alpha} \\ \delta\dot{\beta} \\ \delta\dot{\gamma} \\ \delta\dot{v}_N \\ \delta\dot{v}_E \\ \delta\dot{v}_D \\ \delta\dot{L} \\ \delta\dot{l} \\ \delta\dot{h} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \delta\alpha \\ \delta\beta \\ \delta\gamma \\ \delta v_N \\ \delta v_E \\ \delta v_D \\ \delta L \\ \delta l \\ \delta h \end{bmatrix} + \begin{bmatrix} -C_{11} & -C_{12} & -C_{13} & 0 & 0 & 0 \\ -C_{21} & -C_{22} & -C_{23} & 0 & 0 & 0 \\ -C_{31} & -C_{32} & -C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & D_{11} & D_{12} & D_{13} \\ 0 & 0 & 0 & D_{21} & D_{22} & D_{23} \\ 0 & 0 & 0 & D_{31} & D_{32} & D_{33} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta\omega_x \\ \delta\omega_y \\ \delta\omega_z \\ \delta f_x \\ \delta f_y \\ \delta z_z \end{bmatrix} \quad (3)$$

A detailed description of $A_{11} \sim A_{33}$ the parameters in Eq. (3) is given in an appendix.

3 UAV Dynamic Model

The force applied to UAV make UAV do various maneuvers in flight. UAV dynamic model consists of dynamic equations and kinematic equations. UAV's dynamic equations represent the relationship between forces, moments and its flight status (thrust, control surface angle etc.). UAV's kinematic equations represent the relationship between forces, moments and its kinetic parameters (accelerometer and angular acceleration). So UAV's kinetic parameters, which can be used to aid INS, can be got through dynamic model.

The forces and moments acting on UAVs are mainly due to thrust provided by the aircraft's motor and aerodynamic forces caused by control surface, which are given by Eq. (4)~(9).

$$X_a = T + \frac{1}{2}\rho V^2 S_w C_{X_a} \quad (4)$$

$$Y_a = \frac{1}{2}\rho V^2 S_w C_{Y_a} \quad (5)$$

$$Z_a = \frac{1}{2}\rho V^2 S_w C_{Z_a} \quad (6)$$

$$L = \frac{1}{2}\rho V^2 S_w b C_l \quad (7)$$

$$M = \frac{1}{2}\rho V^2 S_w c_a C_m \quad (8)$$

$$N = \frac{1}{2}\rho V^2 S_w b C_n \quad (9)$$

X_a, Y_a, Z_a are components of the total force resolved in velocity frame. L, M, N are components of the total moment resolved in velocity frame. S_w is the UAV's platform wing area, b is the UAV's platform wing span, c_a is the UAV's platform wing chord, ρ is the atmospheric density of surrounding airflow, V is UAV's velocity, T is UAV's thrust force. C_{X_a}, C_{Y_a}, C_{Z_a}, L, M, N are coefficients of UAV's forces and moments, and they are functions of UAV's velocity and attitude. Theses coefficients can be evaluated through Eq. (10)~(15).

$$C_{X_a} = C_{D0} + AC_{Z_a}{}^2 \quad (10)$$

$$C_{Y_a} = C_{Y\beta}\beta + C_{Y\delta_r}\delta_r + C_{YP}\left(\frac{b}{2v}\right)p + C_{Yr}\left(\frac{rb}{2v}\right) \quad (11)$$

$$C_{Z_a} = C_{L0} + C_{L\alpha}\alpha + C_{L\delta_e}(M)\delta_e \quad (12)$$

$$C_l = C_{l\beta}\beta + C_{l\delta_a}\delta_a + C_{l\delta_r}\delta_r + C_{lp}\left(\frac{pb}{2v}\right) + C_{lr}\left(\frac{rb}{2v}\right) \quad (13)$$

$$C_m = C_{m0} + C_{m\alpha}\alpha + C_{m\delta_e}\delta_e + C_{mq}\left(\frac{qc_A}{2v}\right) + C_{m\dot{\alpha}}\left(\frac{\dot{\alpha}c_A}{2v}\right) + C_{m\dot{\delta}_e}\frac{\dot{\delta}_e c_A}{2v} \quad (14)$$

$$C_n = C_{n\beta}\beta + C_{n\delta_a}\delta_\alpha + C_{n\delta_r}\delta_r + C_{np}(\frac{pb}{2v}) + C_{nr}(\frac{rb}{2v}) \quad (15)$$

Where α, β are the angle of attack and angle of sideslip, the parameters $C_{(\cdot)_{(\cdot)}}$ are UAV aerodynamic coefficients.

After forces and moments are solved, UAV's velocity, Euler angles and rotation rates can be got through kinematic equations as Eq. (16)~(21).

$$\dot{V}_{bx} = \omega_{bz}V_{by} - \omega_{by}V_{bz} + g_x + (C_{X_a}/m) \quad (16)$$

$$\dot{V}_{by} = \omega_{bx}V_{bz} - \omega_{bz}V_{bx} + g_y + (C_{X_y}/m) \quad (17)$$

$$\dot{V}_{bz} = \omega_{by}V_{bx} - \omega_{bx}V_{by} + g_z + (C_{X_z}/m) \quad (18)$$

$$\dot{\omega}_{bx} = C_3\omega_{by}\omega_{bx} + C_4\omega_{by}\omega_{bz} + C_1C_l + C_2C_n \quad (19)$$

$$\dot{\omega}_{by} = C_7\omega_{bz}\omega_{bx} - C_6(\omega_{bx}^{\ 2} - \omega_{bz}^{\ 2}) + C_5C_m \quad (20)$$

$$\dot{\omega}_{bz} = C_9\omega_{by}\omega_{bx} - C_3\omega_{by}\omega_{bz} + C_2C_l + C_2C_n \quad (21)$$

Where $V_b = [V_{bx}, V_{by}, V_{bz}]$ is UAV's velocity in b-frame, $\omega_b = [\omega_{bx}, \omega_{by}, \omega_{bz}]$ is UAV's rotation rates in b-frame, g_x, g_y, g_z are components of gravitational acceleration. C_{0-9} are given by Eq. (22)~(31).

$$C_0 = I_{xx}I_{zz} - I_{xx}^{\ 2} \quad (22)$$

$$C_1 = I_{zz}/C_0 \quad (23)$$

$$C_2 = I_{xz}/C_0 \quad (24)$$

$$C_3 = C_2(I_{xx} - I_{yy} + I_{zz}) \quad (25)$$

$$C_4 = C_1(I_{yy} - I_{zz}) - C_2I_{xz} \quad (26)$$

$$C_5 = 1/I_{yy} \quad (27)$$

$$C_6 = C_5I_{xx} \quad (28)$$

$$C_7 = C_5(I_{zz} - I_{xx}) \quad (29)$$

$$C_8 = I_{xx}/C_0 \quad (30)$$

$$C_9 = C_8(I_{xx} - I_{yy}) + C_2I_{xz} \quad (31)$$

$I_{xx}, I_{yy}, I_{zz}, I_{xz}$ are moment of inertia of UAV, which are related to UAV's structural and called structural coefficients.

UAV's dynamics model coefficients include structural coefficients and aerodynamic coefficients. There is a certain degree of uncertainty with dynamics model coefficients during UAV real flight, and dynamics model couldn't provide accurate information to INS. Structural coefficients are concerned with the mass distribution and shape of UAV, which are relatively stable. Aerodynamic coefficients are usually evaluated for an UAV using a combination of wind tunnel testing and empirical analysis, which are concerned with flight environment (e.g., velocity, attitude, wind speed). Usually real flight environment is not completely consistent with the testing environment, so there exists error in aerodynamic coefficients. In this paper, the sensitivity analysis of navigation accuracy on the error of aerodynamic coefficients will be studied through simulation, which is an inevitable problem when dynamics-aided INS is applied to AUVs.

4 Filter Configuration

The UAV dynamics model and the INS outputs are fused using a filter. In INS/GPS integrated navigation, a Kalman filter (KF) is used. KF is one of the best estimation techniques for integrating signals from short-term high performance systems, like INS, with reference systems exhibiting long-term stability, like GPS. However, KF only works well under appropriately predefined linear error models. Since UAV's dynamics model has strong nonlinear characteristics, an extended Kalman filter (EKF) is adopted which is suitable for nonlinear model.

The state equation can be established as Eq. (32) referenced to section 2 and section 3.

$$\begin{bmatrix}\dot{x}_D \\ \dot{x}_I\end{bmatrix} = \begin{bmatrix}A_D & 0 \\ 0 & A_I\end{bmatrix}\begin{bmatrix}x_D \\ x_I\end{bmatrix} + \begin{bmatrix}w_1 \\ w_2\end{bmatrix} \quad (32)$$

x_D is the error vector of dynamics-based navigation, consisting of 9 states as Eq. (33).

$$x_D = [\delta l_D, \delta L_D, \delta h_D, \delta V_{ND}, \delta V_{ED}, \delta V_{DD}, \delta\Phi_{RD}, \delta\Phi_{PD}, \delta\Phi_{HD}] \quad (33)$$

$\delta R_D = [\delta l_D, \delta L_D, \delta h_D]$, $\delta V_{ND} = [\delta V_{ND}, \delta V_{ED}, \delta V_{DD}]$ and $\delta\Phi_D = [\delta\Phi_{RD}, \delta\Phi_{PD}, \delta\Phi_{HD}]$ are position error, velocity error and attitude error of dynamics-based navigation separately.

The matrix A_D is the error model of dynamics-based navigation, which is expressed by Eq. (33). U is the control input of UAV.

$$A_D = \frac{\partial f(X,U)}{\partial X} \quad (34)$$

x_I is INS error vector, which contains 13 states as Eq. (35).

$$x_I = [\delta l_I, \delta L_I, \delta V_{NI}, \delta V_{EI}, \delta\Phi_{RI}, \delta\Phi_{PI}, \delta\Phi_{HI}, \nabla_x, \nabla_y, \nabla_z, \varepsilon_x, \varepsilon_y, \varepsilon_z] \quad (35)$$

$\delta R_I = [\delta l_I, \delta L_I, \delta h_I], \delta V_{NI} = [\delta V_{NI}, \delta V_{EI}, \delta V_{DI}]$ and $\delta\Phi_I = [\delta\Phi_{RI}, \delta\Phi_{PI}, \delta\Phi_{HI}]$ are position error, velocity error and attitude error of INS separately. $\nabla = [\nabla_x, \nabla_y, \nabla_z]$ represents gyro's error, $\varepsilon = [\varepsilon_x, \varepsilon_y, \varepsilon_z]$ represents accelerometer's error.

A_I is the dynamics matrix of INS, which can be got through Eq. (3). w_1 is a white noise vector of dimension 15, and w_2 is a white noise vector of dimension 9.

The measurement equation can be established as Eq. (36).

$$Z = H\begin{bmatrix} x_D \\ x_I \end{bmatrix} + w_3 \quad (35)$$

The measurement vector Z is the difference of position, velocity and attitude between INS and dynamics-based navigation. w_3 is a white noise vector of dimension 7, and a detailed description of H is given in an appendix.

The filter configuration can be expressed as Fig. 2.

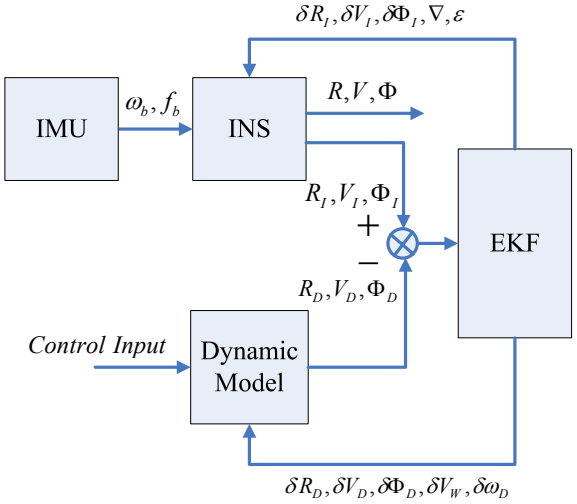


Fig.2 Block diagram of the filter configuration

5 Simulation Results

Our study is carried out in the following scenario: an UAV is traveling in an area where GPS signals are intermittent (e.g., urban canyon), equipped with an INS and a GPS receiver. During the UAV's flight, GPS signals are unavailable at some time. We adopt one typical flight path (Fig. 3), which includes level flight, take off and landing, and slalom-like flight, to demonstrate the usefulness of the UAV dynamics model aiding technique. During the flight, GPS signals are set to be lost respectively lasting 1min, 5mins and 10mins. When GPS is outage, dynamics model aided INS works and its navigation results will be compared with a pure INS's to demonstrate the accuracy improvement.

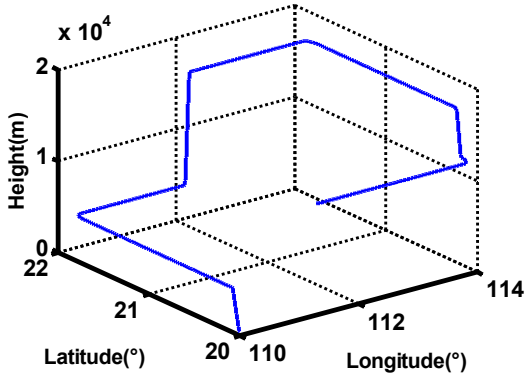


Fig.3 Simulation Trace

The simulation conditions are set as follows:

1) As to three gyros, bias、first-order Markov process and random constant is each $0.1\text{deg}/h$, scale factor symmetrical error is

1ppm, and installation error is $4''$;

2) As to three accelerators, bias and first-order Markov process is each 5×10^{-5} g, scale factor error is 50ppm, and installation error is $4''$;
3) Simulation time is 1h, step length is 0.02s, and the initial position is $[110^\circ, 20^\circ, 500m]$.
4) GPS is outage at the 10th minute, from the 20th to 25th minute, from 40th to 50th minute.

Figure 4 shows the position error.

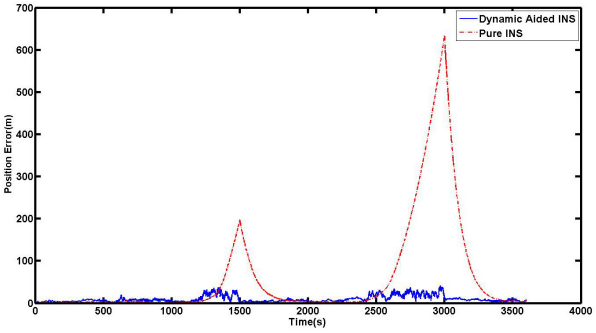


Fig.4 Position Error of Dynamic Aided INS and Pure INS

Through Fig.4 it can be concluded that position error of pure INS is far larger than that of dynamic model aided INS during GPS outage. The advantage of dynamic model aided INS becomes more and more evident with the time of GPS outage increasing.

In order to analyze the sensitivity of navigation accuracy on the error of aerodynamic coefficients, simulation is designed as follows: different degrees of aerodynamic coefficients error (5%, 10% and 20% respectively) are set, and corresponding navigation results will be analyzed to confirm the sensitivity of navigation accuracy on the error of aerodynamic coefficients.

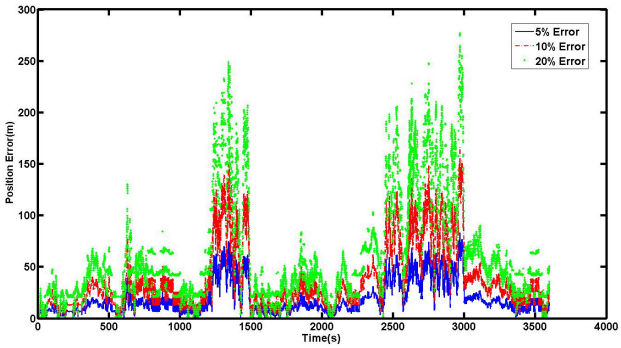


Fig.5 Position Error of Different Degrees of Aerodynamic Coefficients Error

It can be concluded that the position error become larger as the growth of the error of aerodynamic coefficients.

6 Conclusion

The UAV dynamics model aiding technique with the stand-alone INS can efficiently reduce the navigation errors during GPS outages. The navigation accuracy improvement brought by the UAV dynamics model aiding technique in short periods is explored. Simulation is carried out and the advantage of dynamics model aiding technique is verified. Besides, the sensitivity of navigation accuracy on the error of aerodynamic coefficients is analyzed through simulation, which has good reference value to the engineering application of this technique. Furthermore, the technique can also be used when GPS is available, which increases the redundancy of data input to the filter, enabling the navigation system more reliable.

References

[1] Ishibashi S., Tsukioka S. and Yoshida H., "Accuracy Improvement of an Inertial Navigation System Brought about by the Rotational Motion", *OCEANS*, England, 2007, pp. 1-5.

[2] Zhang L., Liu J., Lai J., "Rotating fiber optic gyro strapdown inertial navigation system with three rotating axes", *Transactions of Nanjing University of Aeronautics & Astronautics*, Vol. 25, No.4, 2009, pp. 289-294.

[3] M. Koifman, I. Y. Bar-Itzhack., "Inertial Navigation System Aided by Aircraft Dynamics", *IEEE Transactions on Control Systems Technology*, Vol. 7, No. 4, 1999, pp. 487-493.

[4] Mitch Bryson, Salah Sukkarieh, "Vehicle Model Aided Inertial Navigation for a UAV using Low-Cost Sensors", *Australasian Conference on Robotics and Automation*, Canberra, Australia, 2004.

[5] Simon J.Julier, Hugh F.Durrant-Whyte., "On the Role of Process Models in Autonomous Land Vehicle Navigation Systems", *IEEE Transactions on Robotics and Automation*, Vol.19, No.1, 2003, pp. 1-14.

[6] Klein, Itzik, Filin, Sagi, Toledo, Tomer, "Pseudo-Measurements as Aiding to INS during GPS Outages", *NAVIGATION*, Vol. 57, No. 1, 2010, pp. 25-34.

[7] Øyvind Hegrenæs, Oddvar Hallingstad, "Towards Model-Aided Navigation of Underwater Vehicles", *Modeling, Identification and Control*,

Vol. 28, No.4, 2007, pp. 113-123.

[8] Ø. Hegrenæs, O. Hallingstad, and B. Jalving, "A framework for obtaining steady-state maneuvering characteristics of underwater vehicles using sea-trial data", *Proceedings of the 15th IEEE Mediterranean Conference on Control and Automation*, Athens, Greece, 2007, pp. 1–6.

[9] Øyvind Hegrenæs, Einar Berglund. "Model-Aided Inertial Navigation for Underwater Vehicles", *IEEE International Conference on Robotics and Automation*, Pasadena, USA, 2008, pp. 1069–1076.

Appendix

The following matrixes are parameters in Eq. (3)

$$A_{11}=\begin{bmatrix} 0 & -\left(\Omega\sin L+\frac{V_E}{R}\tan L\right) & \frac{V_N}{R} \\ \left(\Omega\sin L+\frac{V_E}{R}\tan L\right) & 0 & \Omega\cos L+\frac{V_E}{R} \\ -\frac{V_N}{R} & -\Omega\cos L-\frac{V_E}{R} & 0 \end{bmatrix} \quad (A.2)$$

$$A_{12}=\begin{bmatrix} 0 & \frac{1}{R} & 0 \\ -\frac{1}{R} & 0 & 0 \\ 0 & -\frac{\tan L}{R} & 0 \end{bmatrix} \quad (A.2)$$

$$A_{13}=\begin{bmatrix} -\Omega\sin L & 0 & \frac{V_E}{R^2} \\ 0 & 0 & \frac{V_N}{R^2} \\ -(\Omega\cos L+\frac{V_E}{R\cos^2 L}) & 0 & \frac{V_E\tan L}{R^2} \end{bmatrix} \quad (A.3)$$

$$A_{21}=\begin{bmatrix} 0 & -f_D & f_E \\ f_D & 0 & -f_N \\ -f_E & f_N & 0 \end{bmatrix} \quad (A.4)$$

$$A_{22}=\begin{bmatrix} V_D/R & -2\left(\Omega\sin L+\frac{V_E}{R}\tan L\right) & v_N/R \\ 2\Omega\sin L+\frac{V_E}{R}\tan L & \frac{V_N\tan L+V_D}{R} & 2\Omega\cos L+\frac{V_E}{R} \\ -\frac{2V_N}{R} & -2\left(\Omega\cos L+\frac{V_E}{R}\right) & 0 \end{bmatrix} \quad (A.5)$$

$$A_{23}=\begin{bmatrix} -V_E(2\Omega\cos L+\frac{V_E}{R\cos^2 L}) & 0 & \frac{V_E{}^2\tan L-V_NV_D}{R^2} \\ 2\Omega(V_N\cos L-V_D\sin L)+\frac{V_NV_E}{R\cos^2 L} & 0 & -\frac{V_E(V_N\tan L+V_D)}{R^2} \\ 2\Omega V_E\sin L & 0 & \frac{V_N{}^2+V_E{}^2}{R^2} \end{bmatrix} \quad (A.6)$$

$$A_{31}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (A.7)$$

$$A_{32}=\begin{bmatrix} \frac{1}{R} & 0 & 0 \\ 0 & 1/R\cos L & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (A.8)$$

$$A_{33}=\begin{bmatrix} 0 & 0 & -\frac{V_N}{R^2} \\ \frac{V_E\tan L}{R\cos L} & 0 & -\frac{V_E}{R^2\cos L} \\ 0 & 0 & 0 \end{bmatrix} \quad (A.9)$$

The following are coefficients of H in Eq. (36)

$$\begin{aligned} &H_{1,1}=H_{2,2}=H_{3,4}=H_{4,5}=H_{5,10}=H_{6,11}=H_{7,12}=1 \\ &H_{3,18}=H_{4,19}=H_{6,11}=H_{7,22}=-1 \\ &H_{1,17}=\frac{-1}{(R+h)\cos\lambda},H_{2,16}=\frac{-1}{(R+h)\cos\lambda} \\ &H_{5,20}=\frac{-\cos\Phi_{ZI}}{\cos\Phi_{YI}},H_{5,21}=\frac{-\sin\Phi_{ZI}}{\cos\Phi_{YI}},H_{6,22}=\sin\Phi_{ZI} \\ &H_{6,21}=-\cos\Phi_{ZI},H_{7,20}=-\tan\Phi_{YI}\cos\Phi_{ZI} \\ &H_{7,21}=-\tan\Phi_{YI}\sin\Phi_{ZI} \end{aligned} \quad (A.10)$$

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Mathematical Model of Unmanned Aircraft In Ground Effect for Optimal Collision Avoidance

C. Grillo
Università degli Studi di Palermo, Palermo, Italy

Keywords: *Ground Effect, UAS, Non Linear Mode,*

Abstract

The fundamental aim of the present paper is to model the aerodynamic characteristics of an Unmanned Aerial Vehicle(UAV) flying in ground effect. The second aim is to determine the relationship between the optimal avoidance maneuver and the control to execute it. In fact, this relationship is basilar to the development of a guidance scheme capable of approximating the optimal trajectory in real time. In *the first part of this work ,a non-linear model is built in order to model the aerodynamic characteristics of an UAV flying In Ground Effect. To use a single model in the whole range of flying altitude, aerodynamic coefficients are modeled by means of hyperbolic equations. In particular, these ones depend on the ground distance. Such a model is employed to describe aerodynamic forces and moments in the equations of motion which are used in order to obtain optimal trajectories for solving the collision avoidance problem.*

1 General Introduction

As it is well known, UAVs are today employed for several civil and military missions, such as territorial surveillance (forest fire detection, volcanoes monitoring, etc.), law enforcement, disaster assistance, frontier surveillance, agricultural surveying, power-line monitoring, archaeological sites surveillance and many others [[1]], [[2]], [[3]]. In particular, they can be usefully employed in applications which require flight in ground effect (IGE) above water surface [[4]]. This is the case, for instance, of antisubmarine missions [[5]] and, in the civil field, of search and rescue operations or servicing and maintenance of oil rigs. as a consequence, they need to operate at very low altitude, under the influence of extreme ground effect [1], [2], [3]. Safe IGE flight requires high performance flight control systems, particularly when dealing with relatively small aircraft (and this is often the case for UAVs) and/or flight above rough sea.

As it is known, when an aircraft flies near the ground, the lift increases, the induced drag decreases, the neutral point shifts, the pitching moment at zero lift varies and it is necessary to take into account the bank angle(ϕ) derivatives[6].

. The stability characteristics of such a vehicle are strongly correlated to the ground distance, so the primary goal of researchers in this field is the evaluation of aerodynamic coefficients in presence of ground effect. Many Authors have carried out studies in this field[7],[8],[9], [10].In spite of all these studies, there are no mathematical model which affords to analyze the whole flight range of an aircraft. To study

take-off and or landing manoeuvres, for example, the aerodynamic coefficients are modelled by using semi-empirical equations Often, medium values of these ones are used . Nevertheless, stability characteristics are not determined in the whole range of the h/b *ratio* (altitude-span ratio) from in to out ground effect.. Therefore, up to now, two different mathematical models are used to study the dynamic behaviour of aircraft out or in ground effect, [7],[9], [10],[11],[12].

In the present paper, a general mathematical model is built to obtain non – linear analytical equations for aerodynamic coefficients both Out of Ground Effect and In Ground Effect conditions. Such a mathematical model is used to obtain optimal trajectories for solving the collision avoidance problem for a particular UAV in Tandem – Canard architecture [5], [13].

2 Mathematical Model for Aerodynamic coefficients

The studied UAV, as previous stated, has a tandem – canard arrangement; it is particularly suitable to very low altitude missions (for example forest fire detection, volcanoes monitoring and/or battle field surveillance). The fixed tail surface (S_f)is identical to the main wing(S_b) .

All the aerodynamic coefficients have been evaluated in function of classical variables and including also the flight altitude. Obviously, because of the strong nonlinearities due to the ground effect presence, the superposition principle is not applied.

2.1 Longitudinal Coefficients

The influence of ground effect on aerodynamic coefficient has been evaluated as an angle of attack (α) variation, as a downwash(ε) variation and as an aspect ratio(λ) variation due to flight altitude [6], [14]. So, the equations of Lift, Drag and Pitching Moment are:

$$L = \frac{1}{2}\rho V_w^2 S\left(\frac{1}{2}c_{Lf} + \frac{1}{2}c_{Lb}\cdot\cos\varepsilon\right) \qquad (1)$$

$$D = \frac{1}{2}\rho V_w^2 S\left(\begin{array}{l} c_{D0} + \Delta c_{D0}\left(\delta_{el}\right) + \\ + \frac{1}{2}\frac{c_{Lf}^2}{\pi\lambda_f} + \frac{1}{2}\frac{c_{Lb}^2}{\pi\lambda_b}\cos\varepsilon \end{array}\right) \qquad (2)$$

$$\begin{aligned} M = {} & \frac{1}{2}\rho V_w^2 Sc\left(-\frac{1}{2}\frac{h_{nb}}{c}\left(c_{Lb}\cos\left(\alpha_g - \varepsilon\right) + \right.\right. \\ & + \left(\frac{1}{2}c_{D0} + \frac{c_{Lb}^2}{\pi\lambda_b}\right)\sin\left(\alpha_g - \varepsilon\right)\Big) + \\ & + \frac{1}{2}\frac{h_{nf}}{c}\left(c_{Lf}\cos\left(\alpha_g\right) + \right. \\ & \left.\left.\left(\frac{1}{2}c_{D0} + \frac{c_{Lf}^2}{\pi\lambda_f}\right)\sin\left(\alpha_g\right)\right) - c_{m0}\right)\Big) \end{aligned} \qquad (3)$$

Obviously, c_{Lf} and c_{Lb} (respectively forward and backward wing lift coefficient) depend on angle of attack of the single wing:

$$\alpha_f = i_t + \arctan\left(\frac{w - q\cdot h_{nf}}{u}\right) + \Delta\alpha_{IGE} \qquad (4)$$

$$\begin{aligned} \alpha_b = {} & \arctan\left(\frac{w + q\cdot h_{nf}}{u}\right) - \\ & + \left(\varepsilon - \left(\left.\frac{\partial\varepsilon}{\partial\alpha}\right|_{OGE}\left(1 - \Delta\varepsilon_{IGE}\right)\cdot 4q\frac{h_{nf} + h_{nb}}{V_w}\right)\right) + \Delta\alpha_{IGE} \end{aligned} \qquad (5)$$

$\Delta\alpha_{IGE}$ is the angle of attack variation due to ground distance.

In a previous research [15], it has been found that aerodynamic coefficients can be expressed by hyperbolic equations. These ones afford to evaluate aerodynamic coefficients both in Out Ground Effect (OGE) and In Ground Effect (IGE) conditions.

$$\Delta\alpha_{IGE} = \left(\frac{3.3574 + 0.0045 \cdot \left(\frac{h}{b}\right)^{-1.748}}{3.3574} - 1 \right) \cdot \left(1 + 0.001317 \left(\frac{h}{b}\right)^{-1.75}\right) \cdot \alpha_{OGE} \quad (6)$$

$$\Delta\varepsilon_g = \varepsilon \frac{b_{eff}^2 + 4(H_h - H_w)^2}{b_{eff}^2 + 4(H_h + H_w)^2} \quad (7)$$

where b^2_{eff} is the wing span in IGE and it has been evaluated by using calassical methodologies[16], [17].

Moreover:

$$H_w = \frac{h}{b} b - h_{nb} \sin\alpha_g$$
$$H_h = \frac{h}{b} b + h_{nf} \sin\alpha_g \quad (8)$$

In order to evaluate the aspect ratio variation due to altitude, the variation of the induced angle of attack caused by the altitude has been evaluated for each wing:

$$-\frac{1}{\pi\lambda_e} = \frac{c_{L\alpha_{IGE}}}{c_{L\alpha_{OGE}}} \cdot \frac{1}{c_{L\alpha_{airfoil}}} \left(1 - \frac{c_{L\alpha_{airfoil}}}{\pi\lambda_{OGE}}\right) - \frac{1}{c_{L\alpha_{airfoil}}} \quad (9)$$

2.2 Lateral Coefficients

The variation of either Side Force, or Rolling and Yawing Moment due to Ground Effect have been obtained taking into account :

- bank angle(ϕ) effects
- angle of attack (α) variation on aerodynamic surfaces
- aspect ratio(λ) variation due to flight altitude

Side force has been evaluated by modeling the variation of the angle of attack (α) of the vertical tail [16] :

$$\Delta\alpha_g = -18.24(C_L)\frac{\sigma_g}{\lambda} + \frac{r_g T_g (C_L)^2}{57.3 C_{L\alpha}} - r_g B_g \quad (10)$$

with:

$$\sigma_g = e^{-2.48\left(\frac{h}{b/2}\right)^{0.768}} \quad (11)$$

$$r_y - \sqrt{1 + \left(\frac{h}{b/2}\right)^2} - \frac{h}{b/2} \quad (12)$$

$$T_g = \frac{1}{8\pi} \frac{\frac{h}{c}}{\left(\frac{h}{c}\right)^2 + \frac{1}{64}} \quad (13)$$

To evaluate B_g the methodology of [16] has been used.

Rolling and Yawing Moments have been calculated, by inserting Eqs (6), (7), (9) into Eqs (1) (2),(4), (5), and by modeling the ϕ effects.

In fact, near the ground, left and right wing have different distance from the solid surface, so ϕ generates an angle of attack variation .

In particular, for a positive bank angle, the right(h_r) and left (h_l) wing distances from the ground are:

$$h_r = h + y \sin\phi$$
$$h_l = h - y \sin\phi \quad (14)$$

(h is the distance from the ground of the center of gravity).

By inserting Eq.(14) into Eqs (6), (7), (9), it is possible to model lift and drag to evaluate rolling and yawing moments.

3 Collision Avoidance problem

As it is well known, the dynamic equations of a rigid UAV, flying over a flat Heart , in body axes are:

$$\dot{u} = rv - qw - g\sin\theta + \frac{F_{X,B}}{m}$$
$$\dot{v} = -ru + pw + g\sin\phi\cos\theta + \frac{F_{Y,B}}{m}$$
$$\dot{w} = qu - pv + g\cos\phi\cos\theta + \frac{F_{Z,B}}{m}$$

$$\dot{p} = (c_1 r + c_2 p)q + c_3 \bar{L} + c_4 N$$

$$\dot{q} = c_5 pr - c_6 (p^2 - r^2) + c_7 M$$

$$\dot{r} = (c_8 p - c_2 r)q + c_4 \bar{L} + c_9 N$$

$$\dot{\phi} = p + \tan\theta (q \sin\phi + r \cos\phi)$$

$$\dot{\theta} = q \cos\phi - r \sin\phi \qquad (15)$$

$$\dot{\psi} = \frac{q \sin\phi + r \cos\phi}{\cos\theta}$$

$$\dot{x}_E = u \cos\theta \cos\psi + v(-\cos\phi \sin\psi + \sin\phi \sin\theta \cos\psi) + \\ + w(\sin\phi \sin\psi + \cos\phi \sin\theta \cos\psi)$$

$$\dot{y}_E = u \cos\theta \sin\psi + v(\cos\phi \cos\psi + \sin\phi \sin\theta \sin\psi) + \\ + w(-\sin\phi \cos\psi + \cos\phi \sin\theta \sin\psi)$$

$$\dot{h}_E = -u \sin\theta + v \sin\phi \cos\theta + w \cos\phi \cos\theta$$

The state vector for the system of Equations (15) is :

$$\underline{X}^T = [u \quad v \quad w \quad \phi \quad \theta \quad \psi \quad p \quad q \quad r \quad x_E \quad y_E \quad h_E]$$

with the distance from ground (h) instead of the altitude z.

By using the above described mathematical model of Aerodynamic forces and moments it is possible to solve the collision avoidance problem.

The best strategy is to maximize wrt the controls the timewise minimum distance between the aircraft and the obstacle. At the maximin point of the encounter, the distance between the UAV and the obstacle has a minimum wrt the time, which occurs when the relative position vector is orthogonal to the relative velocity vector. In this way, we obtain an inner boundary condition to be satisfied at the maximin point separating the two main branches of the maneuver: the avoidance branch and the recovery branch. As a consequence, a one-subarc Chebyshev problem wich is not solvable in a direct way can be transformed into a two-subarc Bolza problem solvable via the multiple-subarc sequential gradient-restoration algorithm (SGRA)[18],[19].

4 Conclusion

A non-linear model has been built in order to model the aerodynamic characteristics of an UAV flying In Ground Effect. The obtained model can be used in the whole range of flying altitude,because of the aerodynamic coefficients depend, in a direct way, on the ground distance. Such a model has been employed to describe aerodynamic forces and moments in the equations of motion. To cope with the Collision Avoidance problem it has been proposed to transform the Chebyshev problem of optimal control into a Bolza problem, which can be solved , numerically, applying the multiple-subarc sequential gradient-restoration algorithm

References

[1] Jane's,"Unmanned Aerial Vehicles and Targets", Alexandria, VA, USA, Jane's Information Group Inc., 2009

[2] Van Blyenburg P., " Overview of the Current UAV Situation in the World", Euro UVS New, 1999

[3] AA.VV, "Development and Operation of UAV's for Military and Civil Application", *NATO RTO-EN-9,* 2000

[4] Nebylov A., Sharan S., " Concepts & Principles for Creating an Autonomous and Intelligent WIG Vehicle for Coastal Patrolling and Search & Rescue Operations", *Proceedings of the 9 International Conference on Fast Sea Transportation*, 2007, pp. 530-536

[5] Grillo C., Caccamo C., Gatto C., Pizzolo A., "Dinamiche in Effetto Suolo di un UAV non Convenzionale per la Sorveglianza Antisom", *Atti del XVIII Congresso Nazionale AIDAA*, Volterra (PI), Italia, 2005

[6] Rozhdestvensky K. V, " Aerodynamics of a lifting system in extreme ground effect", Springer, 2000

[7] Kumar P.E., "On the Stability of the "Ground Effect Wing" Vehicle'' Ph.D. Dissertation, The University of Southampton, UK, 1969

[8] Hall I.A.," 'An Investigation into the Flight Dynamics of Wing in Ground Effect Aircraft Operating in Aerodynamic Flight'', MSc Dissertation, Cranfield University, UK, 1994

[9] Staufenbiel R.W., Schlichting U.J. ," Stability of

Airplanes in Ground Effect'', Journal of Aircraft, Vol.25 No.4,1988
[10] Delaye H.,"An Investigation into the Longitudinal Stability of Wing in Ground Effect Vehicles". MSc Dissertation ,Cranfield University, UK, 1997
[11] Chun H.H., Chang C.H. "Longitudinal Stability and Dynamic Motions of a Small Passenger WIG Craft'', Ocean Engineering,2001
[12] Chun H.H., Chung K.H. .''Performance Analysis of Wing-In-Ground-Effect Craft'' *Proceedings of International Conference on Wing-In-Ground-Effect Craft*, The Royal Institution of Naval Architects, London,1998
[13] Grillo, C., Caccamo, C., Gatto, C., Pizzolo, A. "Trajectory tracking for an ultralight WIG", *Proceedings of FAST2007*, Shanghai, 2007
[14] Asselin M.,"An Introduction to Aircraft Performance", *AIAA* Educational Series, 1997.
[15] Grillo C., Gatto C." Dynamic stability of wing in ground effect vehicles: a general model" *Proceedings of FAST2005*, St. Petersburg, 2005
[16] ESDU, Engineering Data. *IHS*, 1972
[17] Roskam J.,"Airplane Design Part VI. Preliminary Calculation of Aerodynamic, Thrust and Power Characteristics", The University of Kansas, Lawrence, Kansas, 1990.
[18] Miele A., Wang T., "Optimal Trajectories and Guidance Schemes for Ship Collision Avoidance", Journal of Optimization Theory and Applications, Vol. 129, No. 1, 2006
[19] Miele A., Wang T., "Maximin Approach to the Ship Collision Avoidance Problem via Multiple-Subarc Sequential Gradient-Restoration Algorithm", *Journal of Optimization Theory and Applications*, Vol. 124, No. 1, pp. 29-53, 2005

.

CEAS 2011

Space and Sensors

Design and testing of a miniaturized Fourier transform infrared spectrometer for Mars observation

Bortolino Saggin, Diego Scaccabarozzi, Marco Tarabini
Politecnico di Milano, Department of Mechanics, Lecco's Campus, via M. D'Oggiono 18/A, 23900 Lecco, Italy

Giancarlo Bellucci
INAF-IFSI, via fosso del cavaliere 100, Roma 00133, Italy.

Keywords: *FTIR spectrometer, infrared, calibration, qualification, testing*

Abstract

The design and testing of a miniaturized Fourier transform spectrometer devoted to Mars observation are presented. MIMA (Mars Infrared Mapper) is an FTS based on a michelson interferometer, working in the near-middle infrared range and devoted to Mars observation. The instrument design was guided by strict mass and size limitations, low power availability, all demanding constraints. These requirements combined with the mechanical and thermal environments predicted during the mission, led to a challenging thermo-mechanical design with a constant search for unusual solutions and tradeoff among lightness, resistance and instrument performances. Structural and thermal models were developed and validated through an intensive testing activity. Despite the tested model included many commercial components, in particular for the proximity electronics, the path for the development of a flight quality version was fully identified and the successful testing of the DM (Development Model) proved the soundness of the instrument concept.

1 Introduction

MIMA is a Fourier spectrometer developed to be a part of the Pasteur Instrument Payload for the ExoMars Rover Mission. Mapping of the mineralogical composition of the Martian surface and analysis of its atmosphere, were the scientific goals of the instrument. Moreover, identification of interesting areas for detailed studies using the contact instruments onboard the rover, devoted to the search of biological activity were additional goals. The instrument was conceived with a double pendulum scheme for the interferometer, and two channels to cover the 2 to 25 μm spectral range with a resolution of 5 cm^{-1}. Unfortunately, the last mission redesign bound to mass and cost reductions of the ExoMars program, led to MIMA descoping because its mineralogical and atmospheric studies were considered of secondary importance with respect to the main exo-biologic mission's goals. Despite this unfortunate ending, MIMA development was carried-out up to the qualification testing level. This advanced development stage will hopefully be exploited for the design of similar instruments for future planetary missions. Strong constraints guided the instrument design:

being MIMA mounted on the rover's pointing platform, its overall mass had a ceiling of 1 kg. This was for instance less than half the mass of the Mini-TES experiment (Miniature Thermal Emission Spectrometer), a similar FTS mounted onboard the rovers of the NASA-MER mission. Moreover, instrument size was restrained to a volume of 140 mm x 140 mm x 120 mm and the allowed power consumption to 7 W. A design load of 1000 m/s^2 was considered because the type of loading caused by landing on Mars was uncertain at the time the design started. In addition to this harsh mechanical environment, the non-operational temperature range for MIMA was from -90 to 120 °C. The upper value derived from the wished possibility of dry heat sterilization while the lower depended on the requirement of using no power for thermal control during the non operative phase. These conditions led to an unusual design. Unconventional solutions were developed for the main instrument components. Purposely designed kinematic mountings were devised to mount the optical components warranting high mounting stiffness and low thermally induced stress [1]. An innovative calibration strategy was developed with a calibration cover system designed to perform different tasks, i.e. protecting the instrument entrance window by external contamination or dust deposition, perform as a black body for the LW channel calibration or as a diffuser to reflect the light of the calibration source for the SW channel [2]. Eventually, an instrument mockup was manufactured, calibrated and underwent most of the qualification tests foreseen by the mission [3]. Thermo-vacuum tests were performed, instrument performances were verified while decreasing the optical bench temperature down to -70°C and after cycling to -90°C, values identified as design limits for operative and non operative conditions.

The paper summarizes the full development activity and is organized as follows: MIMA optical layout is firstly described then thermo-mechanical design is presented and testing activity is concludes the work.

2 Optical design

A sketch of MIMA optical layout is shown in figure 1. The incoming radiation passes through the entrance window, is divided by the beamsplitter optics and travels within the instrument following two interferometer arms. After reflection at the cubic corner mirrors (CCM), the radiation goes back to the beamsplitter, generates the interference and is finally focused on the detectors by a parabolic mirror.

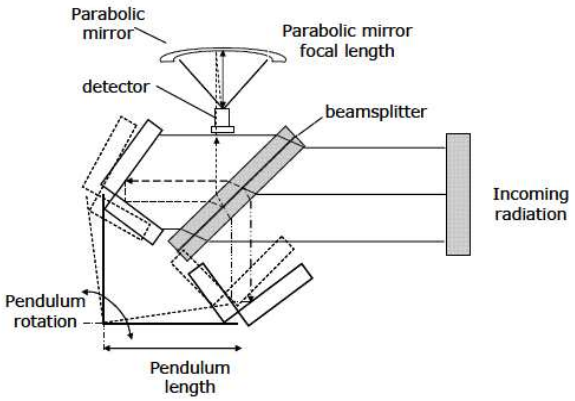


Figure 1 MIMA optical layout.

Such configuration was adopted since more compact than the classical Michelson interferometer setup where optical path difference (OPD) is varied by moving linearly one mirror. The latter configuration moreover requires an accurate linear guide for the moving mirror, capable of warranting during the motion a mirror parallelism within a fraction of micrometer in order not to compromise the system efficiency. This requirement would have been more critical in our case owing to the strict mass budget. Moreover, the heritage on of the double pendulum scheme was available from the previous development of the Planetary Fourier Spectrometer, for the European Mars Express mission [4]. The required 2.0-25 μm wavelengths range could be covered by one optical system with good efficiency but could not be matched by a single uncooled detector with the required signal to noise ratio. Two detectors mounted side by side, a fotoconductive PbSe for the Short Wavelength (SW) range from 2 to 5μm and a Pyroelectric for the 5 to 25 μm allowed to solve the problem.

Being the two sensing areas placed one next to the other (see Figure 2) and in the focus of the same optical system, two nearly contiguous fields of view are achieved and no dichroic mirror is implemented. Such solution was mainly guided by the tradeoff between instrument efficiency and mass. [4]. Geometrical features of the final optical layout are summarized in table 1. It is worth noticing that the mass allocated for the optics was 200g, 20% of the overall instrument budget.

Optical part	Features
Parabolic Mirror	clear aperture with a diameter of 24 mm; focal distance of 17.5 mm;
Cube Corners	clear aperture useful diameter: 37 mm; overall size total diameter: 38 mm or lower angular accuracy 1 arc sec;
Beamsplitter/ Compensator disk	size: ellipsoidal, axes 58mm and 37mm, thickness 5 mm (matching to 0.05 mm within the same pair); coating R / T = 50 / 50 % ± 20 %
PbSe / Pyroelectric detectors	size: 1x1 mm² / 0.85x1.15 mm² SW/LW FOV ≈1.6°

Table 1 MIMA optical features.

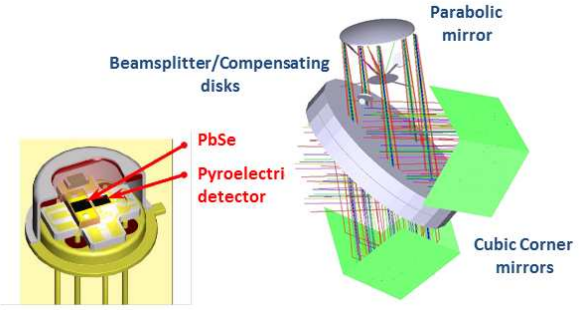


Figure 2 (left) sketch of the MIMA detector assembly, (right) tridimensional model of the optical layout used for ray tracing and optical efficiency analyses.

3 Thermo-mechanical design

Numerical models were developed to predict instrument behavior considering both thermal and mechanical environments. Modal and stress analyses were performed to achieve a lower natural frequency higher than 140 Hz to survive both takeoff and landing loads at Mars. Moreover, an instrument thermal model was built to determine the temperature distributions during the different mission phases; moreover, its outputs have been used for the thermo-elastic analysis bound to assess and reduce the thermal stresses and the distortions of the optical system during the operative phases on the surface of Mars that affect the instrument efficiency.

3.1 FE mechanical model

Figure 3 shows a screenshot of MIMA solid model.

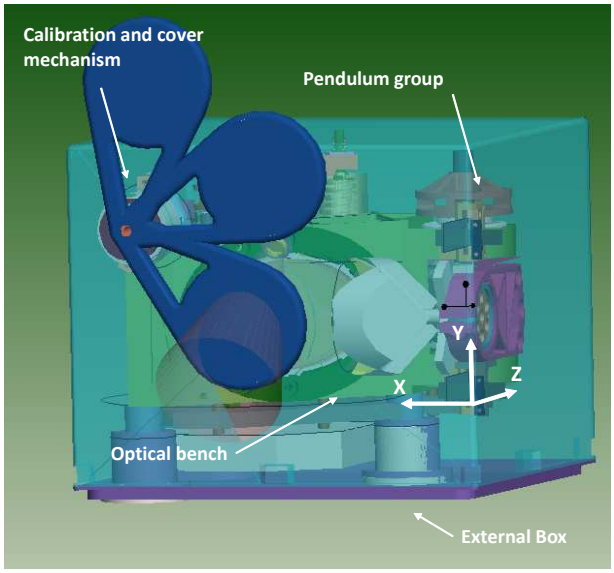


Figure 3 MIMA solid model.

The system can be divided in four main subassemblies, i.e. Optical Bench (OB), Double Pendulum (DP), Beamsplitter Assembly (BA) and External Box group (EB). OB is an aluminum alloy structure, like most of the interferometer components. This is the main structural element that was optimized with a separated set of analyses mainly driven by stiffness, with a requirement of the first natural frequency above 140 Hz. The overall instrument is packaged inside the external box. This assembly has been designed to achieve a sealed enclosure with a single venting path through an EPA filter so that less demanding Planetary

Protection procedure can be implemented for the instrument internal components. Moreover, the EB has a thermal function because it protects the internal part from the severe thermal conditions of mars atmosphere. To speed-up the structural optimization a simplified model, including many concentrated inertial elements was implemented. More details about the FE model and a complete description of all the analyses can be found in [6]. The connection of the instrument to the rover was a critical issue since two opposite goals were sought: high stiffness was desired to achieve resonances above the sine excitation frequency range i.e. within 10-100 Hz range. On the contrary, low resonance frequencies with high damping ratios were sought to reduce the transmission of the random vibration whose spectrum was spread between 20 and 2000 Hz. To match these opposite requirements passive non-linear tunable dampers were designed. Their dynamic properties were individually tuned with a specific testing reproducing the excitation amplitudes expected at the various frequencies. Outputs of the dampers tuning tests were also stiffness and damping characteristics to be used in FEM dynamic analyses.

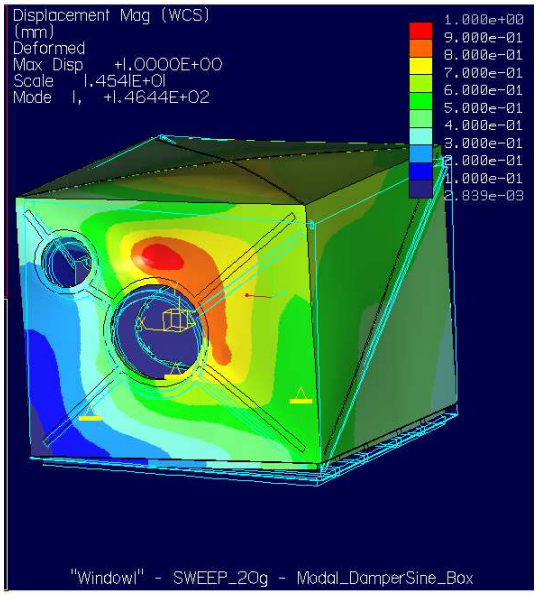


Figure 4 First vibration mode, MIMA model,146 Hz. Displacements are shown in color scale from red to blue.

Figure 4 shows the first natural frequency for damper's parameters corresponding to the low level sweep sine condition while table 2 provides model results also for the random environment.

Modes	SINE [Hz]	RANDOM [Hz]
I	146	102
II	154	131
III	170	160

Table 2 Vibration modes with equivalent dampers parameters for sine sweep and random excitations.

Static analyses were made to compute stresses arising from expected loading during the landing on Mars. Being the direction of the impact with mars surface initially unknown, 1000 ms^{-2} was applied as quasi-static acceleration in all directions of the reference system of Figure 3. Z direction proved to be the worst loading condition for both kind of excitations. Results are reported in table 3.

Dampers parameters	Material	VM Stress [MPa]	FOS	Rs [MPa]	MS
SWEEP SINE	Aluminum	60	2	380	2.2
	KBr	0.6	2.5	5	2.3
	Epox.-Aram. Composite	110	2.5	400	0.5
RANDOM	Aluminum	28	2	380	47
	KBr	0.25	2.5	5	7.9
	Epox.-Aram. Composite	50	2.5	400	2.2

Table 3 Von Mises stresses computed on the worst loading case using connections for sweep sine and random conditions.

Material properties were derived from MIL-HDBK-5J database. For each material the margin of safety (MS) was computed according to ECSS standards as follows:

$$MS = \frac{\sigma_{ADM}}{FOS * \sigma_{MAX}} - 1$$

σ_{ADM} is the material yield stress, σ_{MAX} is the maximum Von Mises (VM) stress and FOS is the factor of safety derived from ECSS standards [7]. Analyses proved that MIMA design is fully compatible with the mechanical environment. Moreover, the implemented

dampers allow to strongly reducing the stresses due to the random excitation, critical for the optics in particular for the KBr (potassium bromide), weak material used for the beam splitter.

3.2 Thermo-elastic analyses and instrument performances.

A FE thermal model was built in order to predict temperature distribution, thermal stresses and optical misalignments within the expected environment. Thermoelastic deformations are one of the most critical causes of instrument performances loss and, in the worst case, they may lead to failures affecting instrument integrity. A thermal model was developed considering four main components: moving equipment, electronics and instrument cover. Different Mars scenarios were simulated seeking for the worst condition. These were identified in the cold operative scenario characterized by a cyclical change of Mars temperature within -100÷-20 °C. The temperature predicted over the instrument is reported in figure 5.

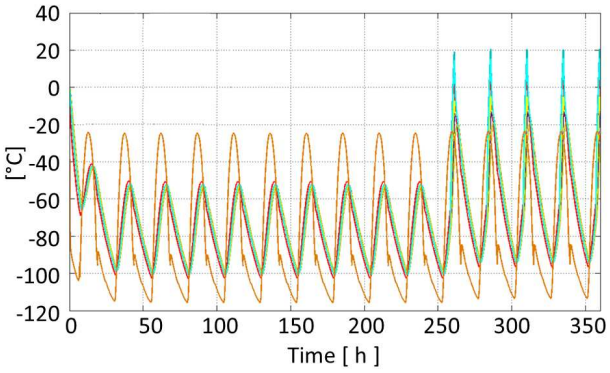


Figure 5 Temperatures of the main instrument parts. Blue line represents the electronic board, the green ones the interferometer internal parts while the orange line for the cover group.

The simulation represents ten days with MIMA in off state where the temperature of the interferometer never goes above –50°C, followed by a measurement session with temperature varying cyclically within the range -100÷20 °C. Temperatures fields from thermal analysis were imported and interpolated in the FE structural model. A worst case condition for the beamsplitter assembly is shown in figure 6.

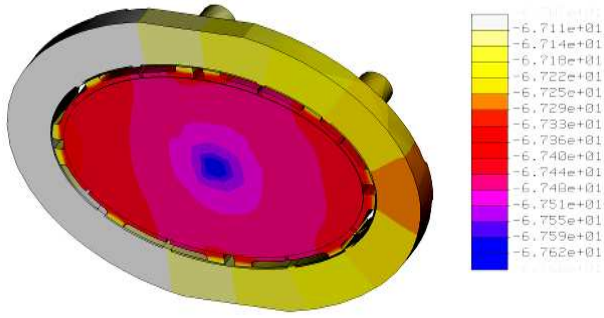


Figure 6 Temperature distribution over the beamsplitter assembly with OB and DP respectively at -66.7 °C and -72.4°C.

Thermal stress analysis showed on the optics a maximum VM stress close to KBr yield stress. Beside this critical result, the analysis was a worst case because the model did not account for any stress release within the glue therefore an experimental validation of the optics mounting through thermal cycling was carried-out; the successful test eventually led to the validation of the optics mounting. This activity was anyway one of the most critical issue that eventually led to the refinement of a technique for the mounting of optical components in wide temperature ranges applicable in a more general case and already experimentally qualified[1]. Optical misalignments were computed from the thermo-elastic analysis. The need of a kinematic mounting of the interferometer on the supporting structure of the rover was evidenced due to the different temperatures and CTE of the two structures. The identified solution was based on dampers with different stiffness along two directions in the plane of the mounting.

4 Testing activity

MIMA design was accomplished leveraging on an intensive testing activity aimed to verify the conceived solutions and assess performances and resistance of many components in the vide temperature range [7]. In the following paragraphs, the thermal and mechanical tests performed on the instrument DM are described.

4.1 Optical stability testing

Performance test were carried-out at different operating temperatures to prove that the misalignment induced by thermal distortions was below the acceptable value as predicted by the thermo-elastic analysis. The criterion for the alignment verification was based on the evaluation of the instrument efficiency in the low wavelength range that is the most sensitive region to misalignmens. The responsivity (DN /Radiance) and NESR at room temperature and at -70°C are shown in figure 7 and 8. It can be seen that the responsivity increases at lower wavenumers, (thanks to cut-off shifting of the colder detector) and is almost stable in the rest of the range, demonstrating that the instrument alignment does not degrade significantly.

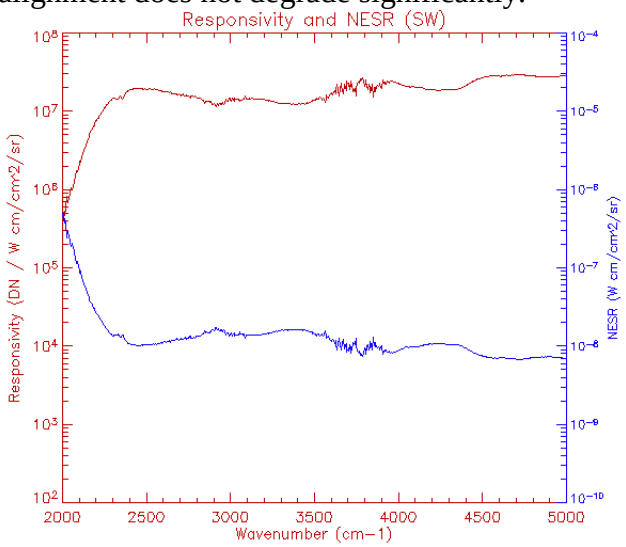


Figure 7 Responsivity and NESR at room temperature

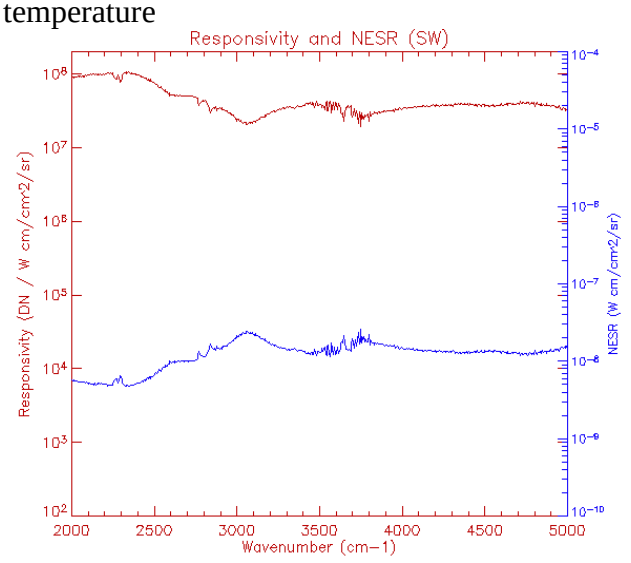


Figure 8 Responsivity and NESR at -70 °C

4.2 Mechanical Qualification testing

The mockup was tested at the excitation levels defined for the mission [8]. The test facility was based on the LDS V830 SPA-16 electro-dynamic shaker, with LMS SCADAS III controller for sine, random and shock excitation. Acceleration signals were acquired by means of three accelerometers placed in the bottom part of the mockup. In figure 9 pick-ups position and measurement axes are shown. The test procedure, in compliance to the ECSS standards, is summarized in the following steps:

- "*resonance search":* it is a sweep sine at 5 ms^{-2} within the range 20-2000 Hz. This is repeated after each power test in order to detect any changes occurred in the instrument behavior. This must be regarded as a complementary check respect to the always present visual inspection and functional test procedure;
- *"sine test"*: instrument is excited in the range 10-100 Hz with a constant acceleration amplitude of 200 ms^{-2} between 20÷60 Hz and a sweep rate of 2 oct/min;
- *"random test"*: mockup is vibrated for 2.5 minutes with a random profile within 20-2000 Hz and an RMS value of 140 ms^{-2};
- *"shock test":* the maximum shock achievable with the shaker is applied to the instrument.

These steps were performed for all three directions of figure 9. Before, during and after each test “visual inspection” procedure allowed checking for damages or unexpected behavior.

Acceleration levels along Z axis are shown in figure 10. The main result was that MIMA DM successfully withstood the full mechanical test program. Concerning the general dynamic behavior, it is noticeable that in the frequency range up to150 Hz no modes with significant amplification were present. This validated the FE models prediction and prevents dangerous dynamic coupling with the low frequency modes of the rover module.

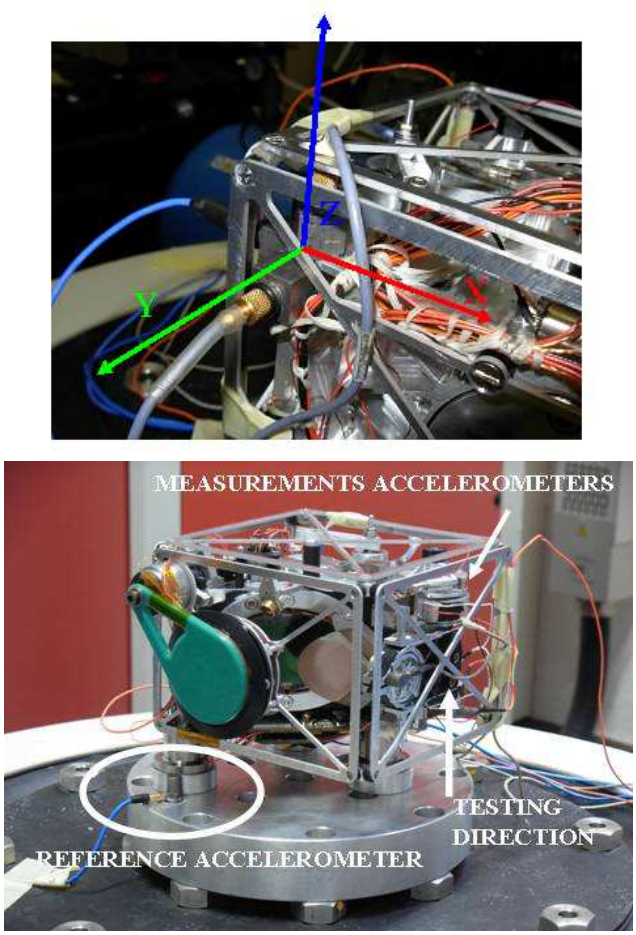


Figure 9 (top) Mounting of the accelerometers used to measure mockup vibration under qualification testing activity. (bottom) Mockup mounted on the testing facility is shown.

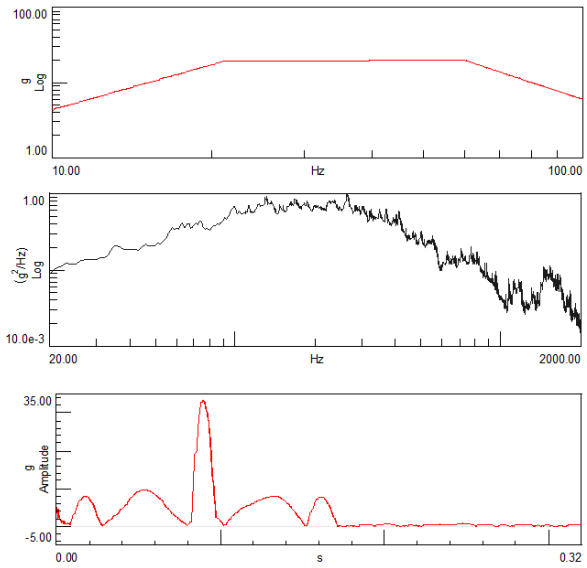


Figure 10. Measured acceleration spectrum for sweep sine, (top), PSD during random (middle) and acceleration time history of the shock (bottom) tests along Z direction.

5 Conclusions

MIMA Fourier spectrometer was born to face harsh mechanical and thermal environments, like those expected for the planned robotic exploration of the Mars surface. Innovative solutions were conceived to fulfill the challenging constraints imposed by the mission and above all the requirement on the mass budget. Design led to a flight quality model fully compliant in terms of expected mechanical resistance, dynamic behavior and radiometric performances. So, the robustness of the proposed concept was proved and the basis for the development of future experiments for planetary exploration was successfully set.

Acknowledgment

The authors gratefully acknowledge the financial support to the MIMA/Exomars instrument provided by ESA and ASI, the European and the Italian Space Agencies.

References

[1] B. Saggin, M. Tarabini, D. Scaccabarozzi., "Infrared optical element mounting technique for wide temperature range", *Applied Optics*, Vol. 24, No. 11, 2011, pp. 1872-1880.

[2] B. Saggin, D. Scaccabarozzi, "Design and Optimization of the Calibration Procedure for a Miniaturized Fourier Transform Spectrometer", *App. Spectr.*, Vol. 65, No 6, 2011, pp. 627-633.

[3] ExoMars Mechanical and Thermal General Design and Interface Requirements - EXM-MS-SSR-AI-0003.

[4] Formisano V., Angrilli F., Fonti S., Saggin B., et al., "The Planetary Fourier Spectrometer (PFS) onboard the European Mars Express mission", *Planetary and Space Science 53* (2005) 963–974

[5] S. Fonti, G. A. Marzo, R. Politi, G. Bellucci, B. Saggin, "MIMA, a miniaturized infrared spectrometer for Mars ground exploration: part II, optical design", *Proceeding of SPIE*, 2007.

[6] D. Scaccabarozzi., "Mechanical design of a FTIR for Mars" Ph.D. Dissertation, Department of Mechanics, Politecnico di Milano, Italy, 2010.

[7] B. Saggin, E. Alberti, L. Comolli, M. Tarabini, G. Bellucci, S. Fonti, "MIMA, a miniaturized infrared spectrometer for Mars ground exploration: part III, thermomechanical design", *Proceeding of SPIE*, 2007.

[8] ExoMars Environment and Test Requirements - EXM-MS-SSR-AI-0004.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Uncertainty experimental evaluation of the Mercury surface reconstruction for the SYMBIO-SIS stereo camera

M. Pertile, A. Aboudan, G. Colombatti, S. Debei
CISAS G. Colombo University of Padova, Italy

Keywords: *Uncertainty, planetary exploration, 3D surface reconstruction, vision systems*

Abstract

The present paper analyzes in detail the measurement uncertainty achievable in 3D reconstruction of a planetary surface by means of stereo or multiple images. The analysis is performed for SIMBIO-SYS experiment onboard BepiColombo ESA space mission. The main uncertainty sources, that propagate to the indirect measurement of 3D points of the surface by means of the stereo camera (STC), are evaluated. A simplified model is used to experimentally evaluate the uncertainty sources in laboratory and allows to define a procedure employable during the future development phases of the mission.

1 Introduction

The measurement uncertainty achievable in 3D reconstruction of a planetary surface by means of stereo or multiple images is deeply analyzed. The analysis is performed for the experiment named SIMBIO-SYS on board BepiColombo ESA space mission. The launch is scheduled for 2014 and the rendezvous with Mercury will take place after a journey of about six years. The three main sub-systems of SIMBIO-SYS are the High spatial Resolution Imaging Channel (HRIC), the Stereo Channel (STC) and the Visible and Infrared Hyperspectral Imager Channel (VIHI). SIMBIO-SYS is able to reconstruct the 3D planetary surface using the colored stereo images of STC, to acquire colored images with high spatial resolution with HRIC and to perform the spectroscopic analysis of the surface by VIHI. The main task of STC is to define the morphological and geological characteristic of the surface. Thus, STC will be an important instrument to measure geophysical parameters and it will provide useful information for the strategy of observations to be performed with the high resolution instrument HRIC.

Due to the importance of scientific tasks and the particular structure (two optical channels aimed at +/- 20° with reference to the Nadir direction), STC will acquire images of the same surface part in different time instants and with different inclination, respectively +20° with the forward channel and–20° with the backward one. The present paper describes in detail the measurement uncertainty achievable in 3D reconstruction of a surface. To maximize the trustworthiness of results during the early phases of SIMBIO-SYS development, experimental tests are performed with a simplified model of STC. The employment of a

simplified but representative model is due to the program schedule; the flight model will be built in 2012. On the other hand, before a detailed design of the flight model is carried out, it is necessary to evaluate the main uncertainty sources associated with input quantities on a simplified prototype.

A complete mathematical model of STC has been developed; it allows to calculate the 3D position of planet surface points, expressed in a predetermined frame of reference. The same mathematical model is employed also to propagate the uncertainty of directly measurable quantities to the 3D position of surface points. The simplified laboratory model has been used to experimentally evaluate the uncertainties associated with directly measurable quantities, so that uncertainty propagation to the indirect measurement is made trustworthy.

The following section 2 describes the mathematical model of STC that is employed for the indirect measurements and for uncertainty propagation, describing also possible variations. Section 3 deals with uncertainty analysis, describing the main uncertainty sources identified and evaluated experimentally, and also the uncertainty propagation methods that are employed. Section 4 summarizes and compares the uncertainties obtained by different methods and models.

2 Measurement algorithm

To perform the indirect measurement of the 3D observed points, some physical parameters and quantities have to be measured. For each camera, its intrinsic parameters (focal length, optical center position, pixel densities along the two sensor axes, parameters of optical distortion) and its extrinsic parameters (position and orientation in a chosen frame of reference) have to be evaluated. Moreover, the 2D position, expressed in pixels, of the projection on the sensor plane of each 3D observed point is also required. The 3D position of a single point can be reconstructed by two or more images that comprise the same point. Thus, if the same point on the surface is acquired once by the forward camera and once by the backward camera, its position in space can be evaluated. The number of images that comprise the same set of visible points depends on the spacecraft velocity and altitude from ground and varies with the position along the orbit. For this reason, to improve the 3D reconstruction, multi-view algorithms able to exploit all available images (and not only two) are employed in this work, [1], [2] . Using multi-view or stereo algorithms, 3D positions of surface points are the output quantities of an indirect measurement, having as input quantities the intrinsic and extrinsic parameters and the 2D positions of points projected on the image sensor.

In the following sub-sections the employed algorithms are described.

2.1 Calculation of 2D points projected on the image plane

The 2D position on the image plane of the surface points is determined by a detector algorithm, that locally analyzes intensity levels of the image and finds the characteristic points (features or keypoints) of the acquired scene. A detector finds points or local regions that are the projection on the image of characteristic points or regions (landmarks) of the viewed scene, while a descriptor yields a representation of each point or region detected in the image. A descriptor allows to find similar regions in different images and to identify points or regions that are the projections on different images of the same point or region in the 3D scene.

Following the advices of [3], the selected detector and descriptor are those described in [4], that defines the approach named Scale Invariant Feature Transform (SIFT). This method tries to identify and describe characteristic points also in presence of scale and orientation variations. For this reason, SIFT method is usually described as invariant to scale and orientation variations.

The detector is based on the difference-of-Gaussian function convolved with the image:

$$D(x,y,\sigma) = \\ = \left[G(x,y,k\sigma) - G(x,y,\sigma)\right] * I(x,y) \qquad (1)$$

Where $I(x,y)$ is the acquired image intensity at the x, y position; $G(x,y,\sigma)$ is the 2D Gaussian with variance σ^2; varying σ makes the considered scale vary; the symbol * denotes the convolution between the difference-of-Gaussian function and the image; k is a multiplicative constant appropriately chosen. The image characteristic points are found out by the maxima or minima values of D function, as defined by Eq. 1, varying position x, y and scale σ. Each detected feature is associated with an orientation, that is computed by the local gradient. More details can be found in [4].

After the characteristic points and regions of the image have been found, the descriptor yields a description of the identified local region. The SIFT descriptor is based on a distribution and represents the region around an identified point using a 3D histogram of positions and orientations. The Euclidean distance is employed to compare the descriptors of two different images and to find out the corresponding features among different images. In detail, each characteristic point identified in one image is associated with the characteristic point of a second image having the minimum Euclidean distance, if this distance is lower than a predetermined threshold.

2.2 Camera model

The camera model employed in this analysis is described in [5], [6]. STC comprises two cameras, each of which has a frame of reference with the z axis aligned with the optical axis.

If X, Y, Z denote the position of a point in 3D space, and x, y are the 2D positions on the image plane, the camera model allows to write:

$$\mathbf{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^T = \lambda \begin{bmatrix} x & y & 1 \end{bmatrix}^T = \lambda \mathbf{x} \qquad (2)$$

where λ is the point depth. Each camera is characterized by a set of intrinsic parameters that define the functional relationship between the coordinates x, y on the image expressed in length units and the coordinates x', y' expressed in pixels. The direct model is:

$$\mathbf{x}'_i = \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot S_x & 0 & x'_{0,i} \\ 0 & f \cdot S_y & y'_{0,i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{K} \cdot \mathbf{x}_i \qquad (3)$$

Where f is the focal length in length units; S_x and S_y are pixel densities respectively along x e y; x'_0, y'_0, are coordinates in pixels of optical center.

To perform the measurement of 3D points is necessary to know also the position and orientation of the cameras when the corresponding images are acquired. These parameters are named extrinsic parameters of the vision system.

2.3 Calculation of 3D points

To evaluate the 3D points three methods are employed and compared: middle point triangulation, least squares approach with more than two images, the Bundle Adjustment method.

2.3.1 Middle point triangulation with two images

When cameras are calibrated, the 3D position of a point can be measured by the algorithm of the middle point, assuming that the coordinates x_1, y_1, x_2, y_2 of the point are known for two images of the same scene. If the two corresponding points on the image planes are ideally projected towards the scene, the two obtained straight lines should intersect each other in the 3D point. In actual facts, the two straight lines are skew and do not intersect due to several uncertainties sources. The middle point algorithm finds the segment of minimum distance between the two skew lines and calculates the 3D point as the middle point of the segment, as described in [7].

2.3.2 Method with more than two images

This method is described in [5], [6] and treats the acquired images as an array of images without grouping them two by two to obtain stereo images. Considering n characteristic points that are viewed in m images (grouped in one array), the depth λ_j associated with the point j seen from the first camera of the array can be calculated by:

$$\lambda_j = -\left(\frac{\sum_{i=2}^{m} \left(\widehat{\mathbf{x}}_i^j \, {}^i\mathbf{P}_{1i} \right)^T \widehat{\mathbf{x}}_i^j \, {}_1^i\mathbf{R} \, \mathbf{x}_1^j}{\sum_{i=2}^{m} \left\| \widehat{\mathbf{x}}_i^j \, {}^i\mathbf{P}_{1i} \right\|^2} \right)^{-1} \qquad (4)$$

where j= 1, … ,n; i=1, …, m; ${}_1^i\mathbf{R}$ is the rotation matrix from frame 1 of image 1 to frame i of image i; ${}^i\mathbf{P}_{1i}$ is the origin of frame 1 with reference to the origin of frame i and also expressed in frame i; $\mathbf{x}_i^j$ denotes the coordinates in length units of point j in the image i; $\widehat{\mathbf{x}_i^j}$ denotes the cross product, i.e. $\widehat{\mathbf{x}_i^j} \, {}^i\mathbf{P}_{1i} = \mathbf{x}_i^j \times {}^i\mathbf{P}_{1i}$. Each 3D point can be calculated by:

$${}^w\mathbf{X}_j = \lambda_j \cdot {}_1^w\mathbf{R} \cdot \mathbf{x}_1^j + {}^w\mathbf{P}_{01} \qquad (5)$$

where j=1, …,n; ${}^w\mathbf{P}_{01}$ is the origin of frame 1 expressed in the world reference frame; ${}_1^w\mathbf{R}$ is the rotation matrix from frame 1 to the world reference frame. In this work, the approach with multiple images is applied using 4 images of the same scene.

2.3.3 Bundle Adjustment with two images

The flow chart of the 3D reconstruction of a scene using the Bundle Adjustment (BA) approach starting from two images is depicted in Fig. 1. As in the two methods already described, the first step is always the identification of a series of corresponding features between two images, which are used in BA algorithm. The features identified in the two images, together with the a priori knowledge of the intrinsic and extrinsic parameters are input quantities of the BA algorithm, which calculates the 3D positions of points and adjusts the parameters (both intrinsic and extrinsic) to improve the matching among all employed images. The parameters obtained by the BA algorithm can be used in subsequent steps to find dense correspondences and to reconstruct a Digital Elevation Map (DEM) of the scene.

In the BA algorithm 3D positions of points and camera parameters are updated to minimize a cost function appropriately defined. Thus, from a mathematical point of view, BA is a problem of parametric non linear evaluation, usually with large dimensions and of sparse type, [8] - [10].

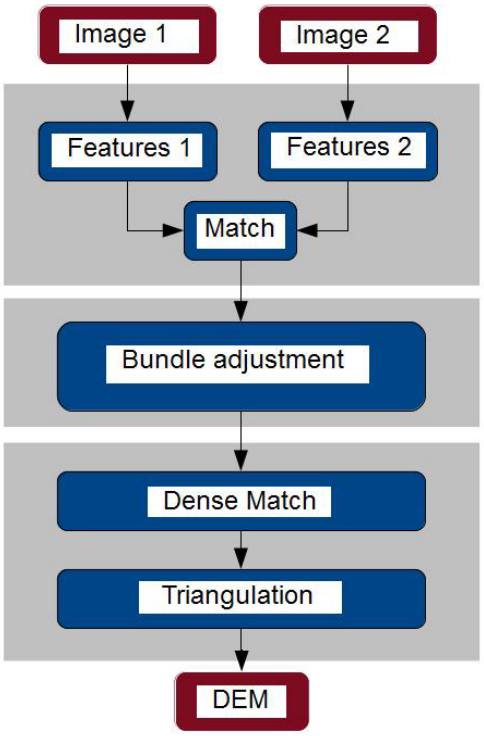


Fig. 1 Bundle Adjustment approach with two images.

The cost function usually is built as the sum of the re-projection errors of the common features identified in images. In some application examples, the cost function may comprise other quantities, e.g. the position of Ground Control Points (GCPs) or position measurements of the cameras performed by the Global Positioning System. In the case of

planetary surfaces as in this paper, the position of GCPs and GPS are not available.

To combine in the same calculation the contribution of measurements of different types that may have different order of magnitude, suitable weights are used. In detail, the weights are selected as functions of the covariance matrices associated with the measurement uncertainties, see [6],[10].

Even if the BA algorithm allows to treat several images, it is applied with only two images in this paper. If $\mathbf{x}_i^j$ denotes the measured 2D positions of the feature j on image i, $\mathbf{X}_j$ is the 3D position of point j in the scene, and $\mathbf{a}_i$ is the vector comprising the intrinsic and extrinsic parameters of the scene i, the cost function can be defined by:

$$f_{ob}(\mathbf{X},\mathbf{a}) = \frac{1}{2}\sum_i \sum_j \left(\Delta \mathbf{x}_i^j\right)^T \mathbf{W}_{i,j} \left(\Delta \mathbf{x}_i^j\right) \qquad (6)$$

Where $\Delta \mathbf{x}_i^j = \mathbf{x}_i^j - \Pi\left(\mathbf{X}_j, \mathbf{a}_i\right)$ is the difference between the positions of the features measured on images and the positions of the corresponding features obtained as re-projection of 3D points with the cameras models (intrinsic and extrinsic parameters). Assuming that $\mathbf{x}_i^j$ measurements are affected by uncertainty of Gaussian type, if the matrices $\mathbf{W}$ are calculated as the inverse covariance matrices of measurements, the cost function is the argument of the joint probability density of all measurements, [11], [12].

The minimum value of the cost function can be evaluated by a suitable minimization algorithm, e.g. Gauss-Newton, Levemberg-Marquardt or similar. The obtained values of $\mathbf{X}_j$ and $\mathbf{a}_i$ are the optimum ones, in the sense that they maximize the probability of points positions $\mathbf{x}_i^j$ in the analyzed images.

3 Uncertainty analysis

A detailed analysis of the uncertainty obtainable in the reconstruction of 3D points is performed using the probabilistic approach described in the Guide to the Expression of Uncertainty in Measurement (GUM) and supplement 1 [13], [14]. Since the analyzed measurement is indirect, the main uncertainty sources for the input quantities are evaluated and then propagated combining them with the mathematical model of STC.

To obtain a trustworthy evaluation of the uncertainties associated with input quantities, experimental tests are performed using a simplified model of STC. The set-up that is used is depicted in Fig. 2; it comprises a simulated rocky scene and a mechanical guide to move a camera and its lens. In this first simplified model the employed camera is of commercial type and is selected with the main characteristics as similar as possible to STC. The inclined position (+/- 20° with reference to the Nadir direction) of the STC cameras is obtained by suitably positioning the camera on the guide. To light the rocky scene, the set-up comprises a lamp, mounted on the guide to experimentally simulate the angular variations of lighting.

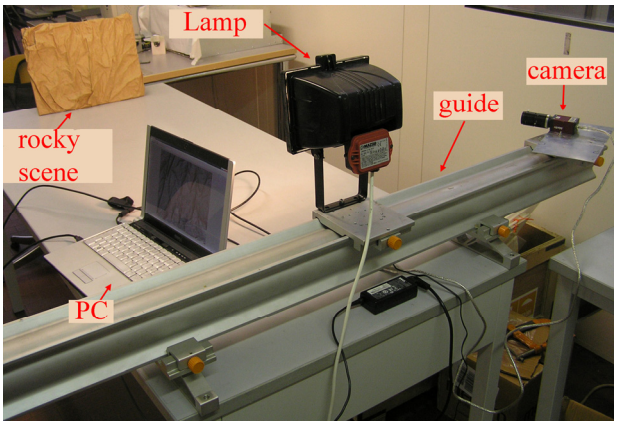


Fig. 2 Experimental set-up.

Particular attention was placed in evaluating the uncertainty associated with the 2D position of image features. More than one uncertainty source is identified and evaluated for the 2D position, e.g. the lighting variation contribution is experimentally analyzed with different lighting positions to simulate operative conditions.

3.1 Main uncertainty sources

The main uncertainty sources are identified for each input quantity: the position of features in images, intrinsic and extrinsic parameters. The experimental analysis with the simplified model of STC allows to evaluate the main uncertainty sources in the initial development phases of the instrument and to define an evaluation procedure applicable also for subsequent models of STC.

3.1.1 2D feature position

Two uncertainty sources are evaluated for the 2D position of features in images: reading uncertainty of the sensor (commonly named "noise" of the sensor) and uncompensated variation of lighting conditions between images.

To evaluate the uncertainty contribution to the 2D feature position due to sensor noise, 400 images of the same simulated scene in the same shooting and lighting conditions are acquired. A simulated rocky scene is employed since sensor noise depends on the scene and lighting conditions. The employed scene and lighting are not uniform; hence, the uncertainty of the acquired intensity levels can not be directly evaluated using a simple frequency histogram of the acquired image, as it is often performed. In this case, for each pixel of the image, the mean value is calculated using all 400 acquired images. These mean values change pixel by pixel due to the non uniform scene. The mean value of each pixel is then subtracted to the acquired intensity value of the same pixel in the acquired images. After this de-trending procedure that removes the scene structure form images, the Probability Density Function (PDF) of the uncertainty associated to the image intensity can be evaluated using a frequency histogram, pixel by pixel and in the whole image. The described procedure requires image noise to be of random type.

The obtained evaluation of image noise is more consistent with operative conditions than other methods, e.g. using a calibrated gray target uniformly lighted. The PDF that is evaluated for the whole image is illustrated in Fig. 3.

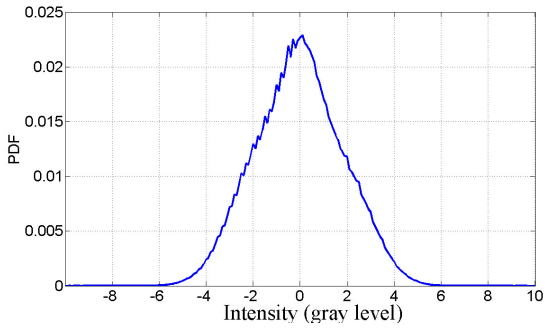


Fig. 3 PDF of the intensity values evaluated with 400 images and associated with image noise.

The PDF of intensity values is used to evaluate the uncertainty of the 2D position of image features. To this aim, a Monte Carlo numerical simulation is performed using a numerical propagation function that comprises the detector and descriptor algorithms and applies them to two images to find out corresponding features. Every Monte Carlo iteration, one of the two images is kept constant, while the other one is obtained from the first one adding pixel by pixel the estimated noise. The outputs of the propagation function are the distances between corresponding features identified in the two images. All distances obtained from all Monte Carlo iterations are gathered in two frequency histograms having on the horizontal axis respectively the distances along x and y axes. In this way, an evaluation of 1D PDFs of feature displacements due to image noise is obtained. The PDF along y axis is illustrated in Fig. 4, while the PDF along x axis is not depicted, since it is substantially equal.

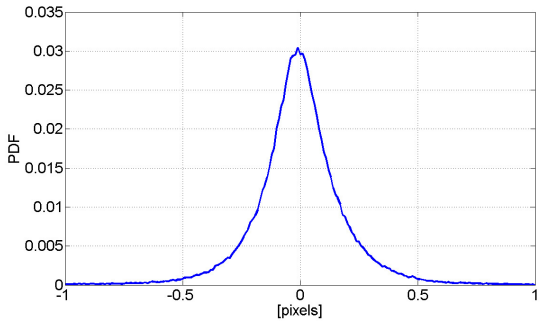


Fig. 4 Evaluated PDF of image feature displacements due to image noise.

The second uncertainty contribution to the

uncertainty of feature positions that is evaluated is the variation of lighting direction between images. During its orbit around Mercury, STC will be able to acquire several images of the same surface area in different time instants. The reasons of this time delay between images are two: the particular configuration with axes inclined of +/- 20° that produces a time delay between the images acquired by the forward channel and those acquired by the backward one; the second reason is that subsequent orbits may yield images partially overlapping that must be fused together during the data elaboration phase. The maximum angular variation of the sun with reference to the surface that may takes place between two images of the same surface area is evaluated from the mission plan. It is assumed that there is not intensity variation of the lighting source between images. The angular variation of the lighting (sun) yields shadow displacements and, hence, a planar translation of features on the image plane between two images of the same scene. To experimentally evaluate this uncertainty contribution a movable lamp aimed to the simulated rocky scene is used. 20 images of the same scene are acquired, moving gradually the lamp to simulate the maximum estimated angular variation. One of the 20 images is employed as reference image, while the others are compared one by one with the reference image to evaluate the movements of the corresponding features. For each couple of images, the displacements of all common features are recorded. Then, all displacements of all couples are collected in two frequency histograms having on the horizontal axis respectively the movements along x and y axes. In this way, an evaluation of 1D PDFs of feature displacements due to the possible angular variation of lighting direction is obtained. The evaluated PDF along x axis is depicted in Fig. 5, while that along y axis in Fig. 6. The two evaluated PDF along x and y axes are slightly different: the PDF along y axis is wider than the other one. This result is due to the fact that the translation direction of the lamp is substantially parallel to y axis of the image sensor. In this analyzed case, it is also clear that the uncertainty contribution due to lighting variation is larger than that due to image noise.

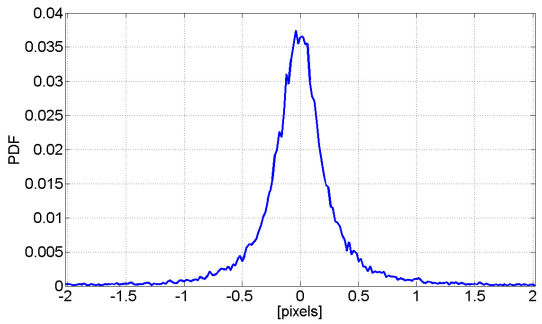


Fig. 5 Evaluated PDF of image feature displacements along x axis due to angular variation of lighting.

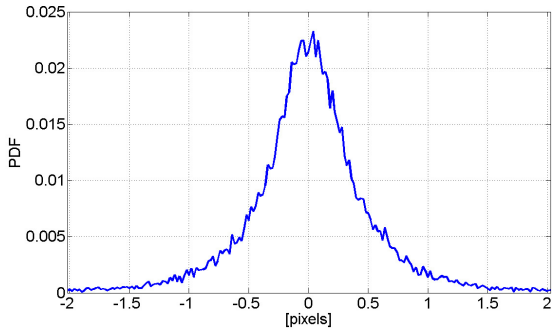


Fig. 6 Evaluated PDF of image feature displacements along y axis due to angular variation of lighting.

3.1.2 Intrinsic and extrinsic parameters

Beside the 2D position of features, the other input quantities that are used in the 3D surface reconstruction are the intrinsic and extrinsic parameters.

The intrinsic parameters can be evaluated by the procedure described in [15], that exploits several images of a planar chessboard having known dimensions. This procedure is applied to the simplified STC model to evaluate the intrinsic parameters. The uncertainty of intrinsic parameters is evaluated by the Monte Carlo method. In this sub-phase, the propagation function is the non linear function described in [15]; the input quantities are the chessboard dimensions and the positions of the chessboard points projected on the image plane; the output quantities are the intrinsic parameters of the

cameras. The uncertainty of input quantities should be known to perform the propagation and evaluate the uncertainty of outputs. In this way the following uncertainty values (with level of confidence 68.3%) are obtained: +/- 3 pixels for fSx and fSy; +/-0.15 pixels for the optical center position. The intrinsic parameters of the STC flight model will be experimentally evaluated on Earth and then, during the flight, they will be adjourned by calibration procedures similar to the one described above for the simplified model of STC (employing images of fixed stars instead of the chessboard).

The positions and orientations of cameras, that define the extrinsic parameters, are calculated for STC from the spacecraft arrangement and from the mission plan. The uncertainty of the extrinsic parameters comprises two contributions:

1. Uncertainty in the measurement of the spacecraft position and attitude;
2. Uncertainty of positioning of the two cameras with reference to the spacecraft frame.

For the linear position, the first contribution, which is calculated from Ephemerides, is much larger than the second one, which is negligible. The obtained values are: +/-10 m with level of confidence 68.3% [16].

The uncertainty of the camera orientation is calculated from the measurement uncertainty of the spacecraft attitude performed by the onboard instruments and from the assembly uncertainty of the cameras on the spacecraft. The two contributions are comparable and, combined together, yields an uncertainty equal to +/- 0.0065° with level of confidence 68.3%.

The uncertainty of intrinsic and extrinsic parameters is correlated with the thermo-mechanical stability of the internal structure of STC in the operative conditions on the Mercury orbit.

3.2 Uncertainty propagation methods

Uncertainty propagation is performed with different approaches: by means of the Monte Carlo simulation method (MCM) for the middle point triangulation with two images, the least squares approach with 4 images, and the BA method; by means of an analytical propagation method only for the BA method. The Monte Carlo simulation is widely used in this work, since the whole propagation function is complex and non linear.

For the BA method, it is possible to demonstrate that the covariance matrix of estimated parameters is, see [12]:

$$\boldsymbol{\Sigma}_{\hat{\mathbf{X}},\hat{\mathbf{a}}} = \left[\mathbf{J}\left(\hat{\mathbf{X}},\hat{\mathbf{a}}\right)^T \mathbf{W}^{-1} \mathbf{J}\left(\hat{\mathbf{X}},\hat{\mathbf{a}}\right) \right]^{-1} \qquad (7)$$

Where **J** is the Jacobian matrix of the cost function, $\hat{\mathbf{X}},\hat{\mathbf{a}}$ denote respectively the values estimated by BA for the 3D positions **X** and for the intrinsic and extrinsic parameters **a**. The uncertainty of the 3D positions X of points can be calculated from this analytical expression.

In detail, the combinations of the methods for 3D point calculation with the methods for uncertainty propagation that are employed are:

1- 3D reconstruction with middle point triangulation from 2 images and uncertainty propagation with MCM.

2- 3D reconstruction with the least squares method from 4 images and uncertainty propagation with MCM.

3- 3D reconstruction with BA from 2 images and uncertainty propagation with the analytical method.

4- 3D reconstruction with BA from 2 images and uncertainty propagation with MCM.

4 Results and discussion

In this paper the analysis is carried out supposing 4 images acquired with the spacecraft at a latitude equal to 13° N and an altitude of 400 km above the ground.

The first two images are acquired by the forward channel oriented concordant with the spacecraft direction, the other two images are acquired by the backward channel. The baseline (distance between the forward and the backward channel during acquisition of a common scene) is supposed to be about 350 km.

Table 1 shows a comparison among the uncertainties evaluated in the four cases

enumerated in section 3.2, with a level of confidence 68.3%. Uncertainties are expressed in meters and the x axis is aligned with the nadir (local vertical) direction. As it is expected, uncertainties are wider along x direction.

	Middle point, 2 images, MCM	4 images, MCM	BA, 2 images, analytical method	BA, 2 images, MCM
x	924	780	518	508
y	562	509	139	185
z	490	476	238	213

Table 1 Evaluated uncertainties of 3D points with level of confidence 68.3%.

For BA, the results obtained by the analytical approach and those by MCM are consistent. Moreover, it is clear that the worst method is the middle point method with two images. The BA approach decreases the obtained uncertainty with reference to the simpler middle point triangulation, and yields uncertainty also better than the multi image approach with 4 images.

5 Conclusion

The measurement uncertainty for 3D reconstruction of a planetary surface by means of stereo or multiple images was analyzed in detail. The analysis was applied to the stereo camera STC of the SIMBIO-SYS experiment onboard BepiColombo ESA space mission. A procedure, employable also during the future development phases of the mission, to evaluate the main uncertainty sources was described and applied using a simplified laboratory model. A possible future work should be the comparison of the indirect measurements obtained by the cameras with the direct measurement of the same laboratory surface obtained by a reference instrument, e.g. a laser scanning instrument.

References

[1] Celenk M., Bachnak R. A., "Multiple stereo vision system for 3-d object reconstruction", *IEEE Int. Conf. on Systems Engineering*, Aug 1990, pp. 555-558.

[2] Esteban C. H., Schmitt F., "Multi-Stereo 3D Object Reconstruction", *Computer Vision Image Understanding*, Dec. 2004, vol. 96, n°.3 Spec.Iss., pp. 367-392.

[3] Mikolajczyk K., Schmid C., "A performance evaluation of local descriptors", *IEEE Trans. on PAMI*, vol. 27, no. 10, 2005, pp. 1615–1630.

[4] Lowe D.G., "Distinctive Image Features from Scale-Invariant Keypoints", *Int. J. of Computer Vision*, 2004, 60(2), pp. 91–110.

[5] Ma Y., Soatto S., Kosecka J., Sastry S. S:, "*An invitation to 3-D Vision*", 2004, Springer.

[6] Hartley R., Zisserman A., "*Multiple view geometry in computer vision*", 2003, Cambridge University Press.

[7] De Cecco M., Pertile M., Baglivo L., Parzianello G., Lunardelli M., Setti F., "Uncertainty analysis for multi-stereo 3d shape estimation", *IEEE AMUEM*, 2009, Bucharest, Romania.

[8] Chen Q., Medioni G., "Efficient iterative solutions to m-view projective reconstruction problem", 1999, *Int. Conf. Computer Vision & Pattern Recognition*, pp. II:55–61.

[9] Granshaw S., "Bundle adjustment methods in engineering photogrammetry", *Photogrammetric Record*, 1980, 10(56), pp. 181–207.

[10] Triggs W., McLauchlan P., Hartley R., Fitzgibbon A., "Bundle adjustment: A modern synthesis", in: W. Triggs, A. Zisserman and R. Szeliski (eds), *Vision Algorithms: Theory and Practice*, 2000, LNCS, Springer Verlag.

[11] Cooper M.A.R., Cross P.A., "Statistical concepts and their application in photogrammetry and surveying", *Photogrammetric Record*, 1988, 12(71), pp. 637–663.

[12] Cooper M.A.R., Cross P.A., "Statistical concepts and their application in photogrammetry and surveying (continued)", *Photogrammetric Record*, 1991, 13(77), pp. 645–678.

[13] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, "*Guide to the Expression of Uncertainty in Measurement*", 1995, Geneva, Switzerland: International Organization for Standardization, ISBN 92-67-10188-9.

[14] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, "*Evaluation of measurement data—Supplement 1 to the Guide to the Expression of Uncertainty in Measurement — Propagation of distributions using a Monte Carlo method*".

[15] Zhang Ji, Boutin M., Aliaga D. G., "Robust Bundle Adjustment for Structure from Motion", *Int. Conf. on Image Processing ICIP*, October 8-11 2006, , pp. 2185-2188, Atlanta, Georgia, USA.

[16] Experiment Interface Document for SIMBIO-SYS part B. EID-B BC-EST-RS-2535, rev. 4.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Uncertainty evaluation and calibration of a vision-based system for displacement and rotation measurement in planetary exploration space missions

M. Pertile, S. Debei
CISAS G. Colombo, University of Padova, Italy

Keywords: *uncertainty evaluation; vision systems; visual odometry, displacement and rotation measurement; planetary exploration*

Abstract

In planetary exploration space missions, motion measurement of a vehicle on the surface of a planet is a very important task. In this work a visual-odometry solution is analyzed. Particularly, a vision-based instrument for displacement and rotation measurement is described and calibrated using a simulated rocky scene. The most significant uncertainty sources are found out by experimental tests. Different motion types are considered and the evaluated uncertainty are compared.

1 Introduction

In planetary exploration, motion measurement of a vehicle on the surface of a planet should be very accurate in order to track the vehicle also for long paths. The odometric evaluation of vehicle position and attitude performed measuring the rotation of wheels has wide uncertainty due to slippage of wheels on a natural, often sandy or slippery, surface. Moreover, on extraterrestrial planets GPS-like positioning systems are not yet available and inertial navigation sensors exhibit unacceptable drifts. Thus, the need of a reliable and accurate motion instrument is particularly relevant. In this work a vision-based instrument for displacement and rotation is described and calibrated using a simulated rocky scene. Displacement and rotation is measured through the images taken by a vision system. The use of stereo systems for visual-odometry is well known, e.g. see [1] -[6]. Stereo-processing allows estimation of the three dimensional (3D) location of landmarks observed by a stereo-camera. If the same landmarks are acquired and detected by a stereo-system that moves from an initial position to a final one, the two 3D point clouds allow to evaluate the position and orientation of the stereo-system in the final position with reference to the initial one. [5] describes a method for visual odometry based on a stereo camera, providing experimental results gathered during development and flight phases of the NASA's twin Mars Exploration Rovers Spirit and Opportunity landed on the surface of Mars in January 2004. A recent work, i.e.[7], emphasizes the importance of a detailed and correct uncertainty evaluation, and describes a method for visual-odometry that allows to reduce the final measurement uncertainty. However, [7] does not take explicitly into account the uncertainty contribution of the feature detector that allows to find the projection of landmarks in the image plane. In the cited works, a detailed calibration

of the vision system as a measurement instrument is not present, while in our analysis it is performed according to [8].

Once the intrinsic and extrinsic parameters of a stereo system are carefully determined, one of the main uncertainty sources is the position on the image plane of the features detected and matched in corresponding images. Detectors find out regions that are projections of landmarks and can be used as features, while descriptors provide representations of the detected regions. A suitable description allows to search similar regions between images and to perform their matching, which is required to measure the 3D position of features and then the position of the stereo system. Thus, several detectors and descriptors were considered in the preliminary phases of this work, see [9] – [16]. Particularly, [13] compares several different detectors invariant to scale and affine transformations and finds out that the best results are obtained by the Hessian-Affine detector and the Maximally Stable Extremal Regions (MSER) approach. The last one performs well on images containing homogeneous regions with distinctive boundaries. [16] makes a comparison among different descriptors and concludes that the Scale Invariant Feature Transform SIFT and the Gradient Location and Orientation Histogram GLOH descriptors are the best. For the reasons described above, in the present work, the selected detector is the Hessian-Affine (MSER was discarded since the simulated rocky scenes do not contain homogeneous regions with distinctive boundaries), while SIFT descriptor is chosen.

2 Measurement algorithm

In this section, the procedure employed to perform the displacement measurement of the stereo system is described. The goal is to calculate the displacement and rotation of a calibrated stereo system using the images acquired in an initial position and in a second one. Thus, the input quantities to be measured are a displacement and/or a rotation of the vision system and the output of the indirect measurement is a numerical evaluation of the displacement and rotation vectors. Rotation is described by a sequence of Euler angles around axes X, Y, Z.

The procedure begins with the detection of image features (keypoints) and the calculation of their image coordinates for both cameras in two different positions. Then, a triangulation phase allows to compute the 3D coordinates of the detected features and, finally, the displacement vector is calculated. In the following paragraphs, there is also a brief description of the employed camera model.

2.1 2D points calculation

As said before, the Hessian-Affine detector was chosen for feature extraction. In the Hessian-Affine detector the image position of interest points is detected using the Hessian matrix, the scale-selection is based on the Laplacian, and the shape of the elliptical region is determined with the second moment matrix of the intensity gradient. Particularly, each feature is located at the local maximum of the determinant of the Hessian matrix:

$$H(\mathbf{x},\sigma_D)=\begin{bmatrix} I_{xx}(\mathbf{x},\sigma_D) & I_{xy}(\mathbf{x},\sigma_D) \\ I_{xy}(\mathbf{x},\sigma_D) & I_{yy}(\mathbf{x},\sigma_D) \end{bmatrix} \qquad (1)$$

whose components are the second derivatives of the image, expressed in levels of gray. The local image derivatives are computed with Gaussian kernels of scale σ_D (differentiation scale).

Then, for each local structure of the acquired scene, the detector tries to find a characteristic scale, which is useful to detect the same local structure when the relative distance between the scene and the camera is changed. To this purpose, the characteristic scale is selected as the one at which there is maximum similarity between the feature detection operator and the local image structures. In this case the selected operator is the Laplacian of Gaussians LoG:

$$\left|LoG(\mathbf{x},\sigma_n)\right| = \sigma_n^2 \left|I_{xx}(\mathbf{x},\sigma_n) + I_{yy}(\mathbf{x},\sigma_n)\right| \quad (2)$$

When the size σ_n of the LoG kernel matches with the size of a blob-like structure the response attains an extremum. Also in this case I_{xx} and I_{yy} are the second derivatives of the image. See [10] and [16], for details. A second advantage of this algorithm is that the size of the region is selected independently of image resolution for each point.

To obtain features with invariance to affine transformations, the iterative estimation of elliptical affine regions is applied. The affine shape of the point neighborhood is determined by the transformation that projects the affine pattern to the one with equal eigenvalues of the second moment matrix. The second moment matrix can be defined by:

$$\mu(\mathbf{x},\Sigma_I,\Sigma_D) = \det(\Sigma_D)\, g(\Sigma_I) * \begin{bmatrix} I_x^2(\mathbf{x},\Sigma_D) & I_x(\mathbf{x},\Sigma_D) I_y(\mathbf{x},\Sigma_D) \\ I_x(\mathbf{x},\Sigma_D) I_y(\mathbf{x},\Sigma_D) & I_y^2(\mathbf{x},\Sigma_D) \end{bmatrix} \quad (3)$$

where Σ_D is the covariance matrix that defines the Gaussian kernels used to compute the local image derivatives, and Σ_I is the covariance matrix of a Gaussian window g(Σ_I) used to smooth the derivatives.

The detector described above allows to compute the 2D position of the detected features. After that, for each feature in the image of a camera, the corresponding feature in the other camera has to be found out, if any. Moreover, corresponding features have to be found out between images acquired by the same camera but in two different positions. These correspondences are evaluated using the SIFT descriptor, which is a distribution based approach. The descriptor represents the local region around each detected feature by a 3D histogram of gradient locations and orientations, for details see [9]. The Euclidean distance is used to compare the descriptors. Each detected feature in a first image is associated with the feature in the second image that has the minimum Euclidean distance, providing that this distance is below a set threshold.

2.2 Camera model

The employed camera model can be found in [4]. The stereo system comprises two cameras, 1 and 2, each camera has a corresponding frame of reference with its z axis aligned with the optical axis. The two cameras are joined together with their optical axes substantially parallel. Considering the model of each camera, the generic position of a point feature comprised in the field of view of both cameras may be written as:

$${}^i\mathbf{X} = {}^i\begin{bmatrix} X & Y & Z \end{bmatrix}^T = \lambda_i \begin{bmatrix} x_i & y_i & 1 \end{bmatrix}^T = \lambda_i \mathbf{x}_i \quad (4)$$

where i is 1 or 2, depending on which camera is considered; ${}^1\mathbf{X}$ (or ${}^2\mathbf{X}$) is the point position expressed in frame 1 (or 2) associated with camera 1 (or 2); $\mathbf{x}_1$ (or $\mathbf{x}_2$) is the projection of the point ${}^1\mathbf{X}$ (or ${}^2\mathbf{X}$) with an ideal camera aligned like camera 1 (or 2) with a focal length equal to 1 (in length units); $\lambda_i \in \mathbb{R}^+$ is a scalar parameter associated with the depth of the point.

Each camera is characterized by a set of intrinsic parameters which are evaluated during camera calibration, and defines the functional relationship between projection $\mathbf{x}_i$, expressed in length units, and projection $\mathbf{x}'_i$, expressed in pixels (x'_i and y'_i are respectively the column number and row number, from the upper left corner of the sensor); an ideal pinhole camera reveals the following direct model:

$$\mathbf{x}'_i = \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot S_x & 0 & x'_{0,i} \\ 0 & f \cdot S_y & y'_{0,i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{K} \cdot \mathbf{x}_i \quad (5)$$

where S_x is the pixel density along x_i axis; S_y is the pixel density along y_i axis; f is the focal length in length units; $x'_{0,i}$, $y'_{0,i}$ are the distances (respectively in pixel columns and rows) between the upper left corner and the

principal point (intersection of the optical axis with the sensor).

A stereo system is also characterized by its extrinsic parameters. Particularly, to perform triangulation and determine the 3D position of the scene points, the relative position and rotation between the two cameras have to be known. These parameters are named extrinsic and should be experimentally evaluated before any measurement with the stereo system can be performed. Generally, the procedure to evaluate the intrinsic and extrinsic parameters is referred to as calibration, which should not to be confused with the calibration of the whole measurement system described in this work.

2.3 3D points calculation

When both cameras of the stereo system are calibrated, the 3D position of a feature point in space may be measured by means of a triangulation algorithm. In this work, the algorithm of the middle point is used for triangulation, as in [5]. The algorithm, assuming the projected points $\mathbf{x}_i$ as known from camera calibration, finds 3D points $\mathbf{X}_{1,s}$, $\mathbf{X}_{2,s}$ with the minimum distance, belonging respectively to the preimage lines of cameras 1 and 2. Points $\mathbf{X}_{1,s}$, $\mathbf{X}_{2,s}$ define a segment orthogonal to the two skew preimage lines. Middle point $\mathbf{X}_m$ of this segment is selected as the measured 3D point of the feature.

Equation (4) can be used to find two generic points $\mathbf{X}_1$, $\mathbf{X}_2$, each belonging to the corresponding preimage line and expressed in the reference frame associated with the corresponding camera:

$$ {}^1\mathbf{X}_1 = \lambda_1 \cdot \mathbf{x}_1 \quad , \quad {}^2\mathbf{X}_2 = \lambda_2 \cdot \mathbf{x}_2 \qquad (6) $$

Both points can be expressed in the reference frame 1 attached to camera 1:

$$ {}^1\mathbf{X}_1 = \lambda_1 \cdot \mathbf{x}_1 \quad , \quad {}^1\mathbf{X}_2 = \lambda_2 \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 + {}^1\mathbf{P}_{21} \qquad (7) $$

where ${}^1_2\mathbf{R}$ is the rotation matrix from frame 2 to frame 1, and ${}^1\mathbf{P}_{21}$ is the origin of frame 2 with reference to the origin of frame 1 and expressed in frame 1. Points $\mathbf{X}_{1,s}$, $\mathbf{X}_{2,s}$ with minimum distance are calculated minimizing the following cost function: $g = \left\| {}^1\mathbf{X}_1 - {}^1\mathbf{X}_2 \right\|^2$. The following values of $\lambda_{1,s}$, $\lambda_{2,s}$ are obtained:

$$ \lambda_{1,s} = \frac{\left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2\right) \cdot \left(\mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12}\right) + \left(\mathbf{x}_2^T \cdot \mathbf{x}_2\right) \cdot \left(\mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21}\right)}{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1\right) \cdot \left(\mathbf{x}_2^T \cdot \mathbf{x}_2\right) - \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2\right)^2} \qquad (8) $$

$$ \lambda_{2,s} = \frac{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1\right) \cdot \left(\mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12}\right) + \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2\right) \cdot \left(\mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21}\right)}{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1\right) \cdot \left(\mathbf{x}_2^T \cdot \mathbf{x}_2\right) - \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2\right)^2} $$

,

where ${}^1\mathbf{P}_{12}$=-${}^1\mathbf{P}_{21}$; ${}^2\mathbf{P}_{12}$ is the origin of frame 1 with reference to the origin of frame 2 and expressed in frame 2. Thus, the extreme points ${}^1\mathbf{X}_{1,s}$, ${}^1\mathbf{X}_{2,s}$ of the minimum distance segment and the middle point $\mathbf{X}_m$ associated with the point feature are:

$$ {}^1\mathbf{X}_{1,s} = \lambda_{1,s} \cdot \mathbf{x}_1 \quad , \qquad (9) $$

$$ {}^1\mathbf{X}_{2,s} = \lambda_{2,s} \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 + {}^1\mathbf{P}_{21} \quad , $$

$$ {}^1\mathbf{X}_m = \frac{{}^1\mathbf{X}_{1,s} + {}^1\mathbf{X}_{2,s}}{2} $$

For convenience, all obtained 3D points are expressed in the frame 1 attached to the first camera.

The uncertainties of 3D points are needed for the evaluation of stereo system displacement and rotation as described in section 2.4. The uncertainty evaluation for triangulated 3D points becomes an uncertainty propagation problem, which employs the functional model between input and output quantities:

$$ {}^1\mathbf{X}_m = f\left(\mathbf{x}_1, \mathbf{x}_2, {}^1_2\mathbf{R}, {}^1\mathbf{P}_{21}, \mathbf{K}_1, \mathbf{K}_2\right) \qquad (10) $$

Where $\mathbf{K}_1$, $\mathbf{K}_2$ are the intrinsic parameter matrices of camera 1 and 2 of the stereo system. To calculate the propagated uncertainty of the triangulated position, taking into account the contributions of all uncertainty sources, the Kline-McClinton formula is used in this phase of the measurement algorithm, see GUM [8]. This method is selected, instead i.e. of the Monte Carlo propagation approach, since the calculation is embedded in the algorithm that calculate the displacement and rotation of the stereo camera, and is performed for all detected features. Thus, the use of a Monte Carlo simulation could lead to unacceptable time delays in the position and attitude evaluation of the stereo camera. The uncertainties of input quantities (intrinsic and extrinsic parameters, image feature positions) are evaluated as described in section 4.

2.4 Stereo system displacement and rotation

As described above, when the stereo system is calibrated (intrinsic parameters of both cameras and the position and orientation of camera 2 with reference to camera 1 are known), the middle points ${}^{P1}\mathbf{X}_m$ can be calculated for all features detected by both cameras when the vision system is in an initial position P1. The notation ${}^{P1}\mathbf{X}_m$ means that these points are expressed in the reference frame attached to the first camera, when the vision system is in the initial position *P*1. When the vision system is moved (cameras are rigidly connected) from the initial position *P*1 to a second position *P*2, the same procedure can be used to compute the 3D vectors ${}^{P2}\mathbf{X}_m$ of the features detected by both cameras in the second position *P*2 and expressed in the new frame 1 attached to the first camera. For each feature that is detected by both cameras in both positions *P*1, *P*2, the following equation can be written:

$$ {}^{P1}\mathbf{X}_{m,i} = {}^{P1}_{P2}\mathbf{R} \cdot {}^{P2}\mathbf{X}_{m,i} + {}^{P1}\mathbf{P}_{P2,P1} \qquad (11) $$

Where ${}^{P1}_{P2}\mathbf{R}$ is the rotation matrix from frame 1 in the second position *P*2 to frame 1 in the initial position *P*1; ${}^{P1}\mathbf{P}_{P2,P1}$ is the origin of frame 1 in *P*2 with reference to the origin of frame 1 in *P*1 and expressed in *P*1; i= 1, …, *n*, with *n* being the number of common detected and matched features. Vector ${}^{P1}\mathbf{P}_{P2,P1}$ and the Euler angles that define matrix ${}^{P1}_{P2}\mathbf{R}$ are the numerical output values of the whole measurement procedure and, in this work are evaluated in two steps in a similar way as in [5]: first, a less accurate motion is estimated by least squares estimation embedded within a random sample consensus (RANSAC) process to remove outliers; then, a maximum likelihood motion estimation is performed minimizing a non linear problem. The main differences with [5] are the more advanced feature detection and matching algorithms and the minimization procedure that comprises the Levenberg-Marquardt algorithm.

3 Calibration procedure

The first step to calibrate the whole measurement system is the determination of intrinsic and extrinsic parameters of the stereo system. The selected method is described in [18] and comprises a first closed form analytical evaluation of parameters for both cameras. The obtained values are then used as the starting point for a nonlinear optimization technique based on the maximum likelihood criterion. This nonlinear minimization problem, which is solved with the Levenberg-Marquardt algorithm, takes into account the lens distortion, especially radial distortion, for each camera. In our case the extrinsic parameters are the relative position and orientation between the two cameras.

After the intrinsic and extrinsic parameters of the stereo system are evaluated, the calibration of the measurement system can be performed. To calibrate the whole measurement system, the stereo cameras are mounted on a rotary motorized stage, and/or on a linear motorized stage. The stereo system position and orientation are changed by known quantities and measured by the encoders of the linear and

rotary stages which are used as reference instruments; in each position both cameras acquire an image; the imposed movements are evaluated analyzing images in different positions and orientations. Figure 1 depicts the employed experimental set-up.

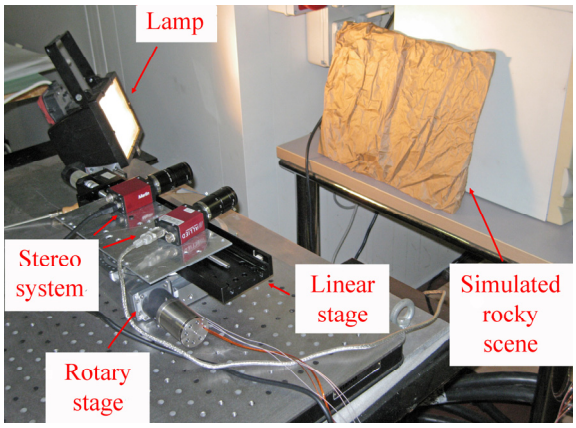


Fig. 1 Experimental set-up.

The calculation algorithm employs the camera model described in [6]. The procedure begins with the detection of image features (keypoints) and the calculation of their position in each image. Then, a triangulation phase allows to compute the 3D coordinates of the detected features and, finally, the displacement vector and rotation (expressed by Euler angles) are calculated. Both cameras are aimed to a simulated planetary scene obtained with crumpled brown paper. In experimental tests the allowed movements are pure translations, rotations and several combinations of linear displacements and rotations.

4 Uncertainty analysis

Besides the experimental calibration tests, a detailed uncertainty analysis is carried out according to the metrological procedures described in [8] and [17]. A Monte Carlo propagation is used to evaluate the uncertainty of linear displacement and rotation, which are treated as the output quantities of an indirect measurement. In this analysis phase the aim is to evaluate uncertainties of position and attitude of the stereo system as precisely as possible, and not to calculate displacements and rotations in a fast way. Moreover, the whole measurement algorithm is highly non linear. Thus, a Monte Carlo simulation is employed instead of the propagation formula embedded in the calculation algorithm as described in section 2.3 and that is an approximated method.

Several uncertainty sources are analyzed and evaluated using experimental tests also with the simulated planetary scene. Particularly, the uncertainty associated with the following quantities are taken into account: intrinsic and extrinsic parameters of the stereo system, whose uncertainties are evaluated using a camera calibration procedure, and the positions of image features, whose uncertainty is affected by several contributions. In this paper three main contributions are evaluated for image features: the reading uncertainty (image noise) of the cameras; the lighting variations of the simulated scene, the finite resolution of the cameras.

The contribution of image noise could be particularly tricky since it depends on the scene. For this reason, its contribution is evaluated using the same simulated rocky scene acquired during measurements. When the same scene is acquired in the same conditions, different images are obtained due to the uncertainty (commonly named image noise in computer vision) associated with the image sensor and its electronics. This uncertainty depends on the acquired scene. Thus, for each camera, 200 images of the same scene were acquired with both cameras in fixed positions; for each pixel, the difference of the read level of gray and the corresponding mean value is calculated and the frequency histogram of all these differences of all pixels is used as an evaluation of the Probability Density Function PDF of the reading uncertainty. The obtained PDFs are used to evaluate the position ($\mathbf{x}'_i$) uncertainty of the image features detected and matched by the detector and descriptor employing a Monte Carlo simulation.

The lighting uncertainty contribution to feature positions $\mathbf{x}'_i$ is carefully evaluated moving the lamp in the experimental set-up in order to reproduce the sun movements on the surface of a planet. Since for the particular simulated scene that is used, the position of the

lamp may cause wide variations of feature positions $\mathbf{x}'_i$, the maximum allowed movement of the lamp was estimated assuming the time t_m required for the stereo system to move from an initial position to a final one. Time t_m should be representative of the average moving velocity of a rover on a planetary surface and allows to estimate a maximum angular variation of the lighting if the planet (e.g. Mars) rotation is known. Once the lamp is moved of the estimated maximum amount, the PDF of the image feature movements is evaluated and is used as an additional uncertainty source associated with the feature positions $\mathbf{x}'_i$.

The uncertainties associated with the intrinsic and extrinsic parameters of the stereo system are evaluated applying a Monte Carlo simulation to the method described in [18] . Also in this case, the Monte Carlo simulation is selected due to the highly non-linear algorithm employed.

According to [8] and [17], all uncertainty sources are expressed by a probability density function (PDF) and are then propagated to the output displacement of the stereo system. Thus, the PDFs of the uncertainty contributions associated with the intrinsic and extrinsic parameters, and with 2D feature positions $\mathbf{x}'_i$ (due to image reading uncertainty, lighting variations and resolution) are propagated to the displacement of the stereo system using a Monte Carlo simulation. In this way, both the calibration curve and its uncertainty are evaluated for the output measured displacement.

5 Results

The results are the calibration curves of the instrument obtained with the set-up depicted in Fig.1. Figure 2 shows the curve and its evaluated uncertainty for longitudinal displacements (substantially parallel to the optical axes), Fig. 3 for transverse displacement (substantially orthogonal to the optical axes), Fig.4 for rotation, Fig.5 for rotation and transverse displacement. In all graphs, the imposed motion is illustrated along the horizontal axis, while the measure one along the vertical axis.

The uncertainty evaluated for longitudinal displacements is larger than that obtained for all other movements. This result can be explained considering that the uncertainty of 3D points acquired by the stereo system is wider along the direction parallel to the optical axis than along a transverse direction, due to the small distance between the two cameras. This disadvantage along longitudinal direction is partially compensated by the fact that the number of

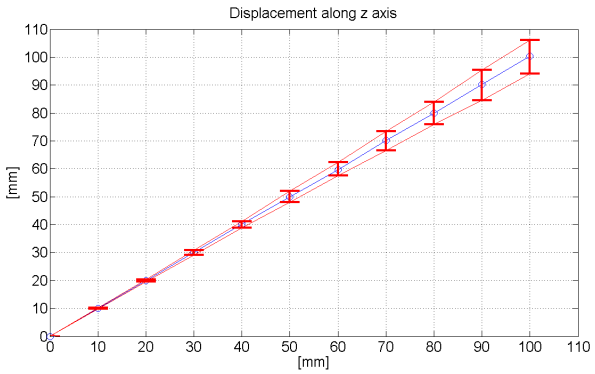


Fig. 2 Calibration curve for longitudinal displacement.

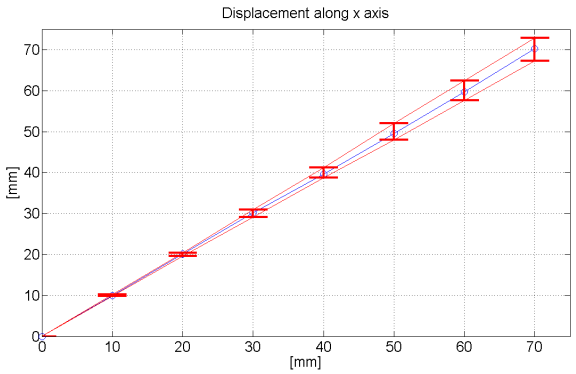


Fig. 3 Calibration curve for transverse displacement.

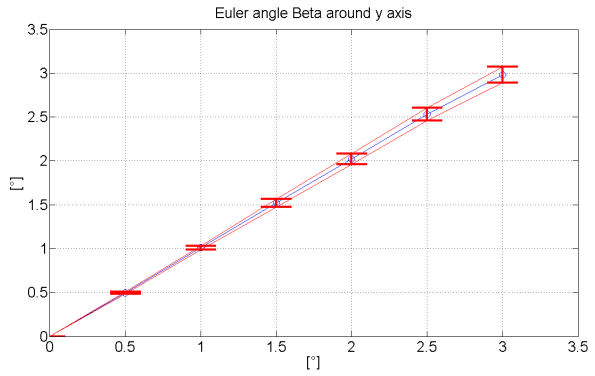


Fig. 4 Calibration curve for rotation.

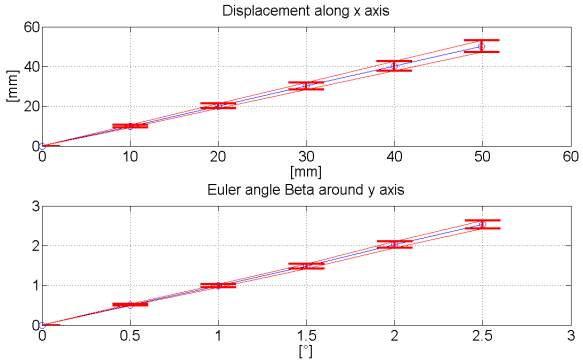


Fig. 5 Calibration curve for rotation combined with transverse displacement.

image features correctly matched in case of axial displacement is generally greater than the number of matched features in case of other movements.

For all displacements and rotations, the evaluated uncertainty increases with the measured motion. A possible reason of this behavior is that the larger the movement and the fewer the correctly matched features among images; thus, the averaging effect associated with a large number of matched features decreases with displacement/rotation. This effect is less evident for movements substantially parallel to the optical axis (longer distances can be measured in this direction than in the transverse one).

6 Conclusion

The calibration of a stereo system as an instrument for displacement and rotation measurement was described. The calibration was carried out using a simulated rocky scene and different types of motion. This procedure allowed to highlight advantages and drawbacks of the considered algorithm for different motion types.

References

[1] Arun K. S., Huang T. S., Blostein S. D., "Least-squares fitting of two 3-d point sets", *IEEE Trans on PAMI*, Vol.9, No.5, 1987, pp. 698–700.

[2] Eggert D., Lorusso A., Fisher R., "Estimating 3-d rigid body transformations: a comparison of four major algorithms", *Machine Vision and Applications*, Vol.9, No.5-6, 1997, pp.272–290.

[3] Umeyama S., "Least-squares estimation of transformation parameters between two point patterns", *IEEE Trans on PAMI*, Vol. 13, No.4, 1991, pp. 376 – 380.

[4] Olson C. F., Matthies L. H., Schoppers M., Maimone M. W., "Rover navigation using stereo ego-motion", *Robotics and Autonomous Systems*, Vol. 43, 2003, pp. 215–229.

[5] Cheng Y., Maimone M. W., Matthies L., "Visual Odometry on the Mars Exploration Rovers", *IEEE Robotics & Automation Magazine*, 2006.

[6] Ma Y., Soatto S., Kosecka J., Sastry S. S., *An invitation to 3-D Vision*, Springer, 2004.

[7] Dubbelman G., Groen F. C. A., "Bias Reduction for Stereo based Motion Estimation with Applications to Large Scale Visual Odometry", *IEEE CVPR*, 2009.

[8] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, *Guide to the Expression of Uncertainty in Measurement*, Geneva, Switzerland: International Organization for Standardization, ISBN 92-67-10188-9, 1995.

[9] Harris C., Stephens M., "A Combined Corner and Edge Detector", *Proc. Alvey Vision Conf.*, 1988, pp. 147-151.

[10] Mikolajczyk K., Schmid C., "Indexing Based on Scale Invariant Interest Points," *Proc. 8th Int. Conf. Computer Vision*, 2001, pp. 525-531.

[11] Lowe D., "Object Recognition from Local Scale-Invariant Features", *Proc. 7th Int. Conf. Computer Vision*, 1999, pp. 1150-1157.

[12] Lowe D., "Distinctive Image Features from Scale-Invariant Keypoints", *Int. J. Computer Vision*, Vol. 2, No. 60, 2004, pp. 91-110.

[13] Mikolajczyk K., Schmid C., "Scale and Affine Invariant Interest Point Detectors," *Int. J. Computer Vision*, Vol. 1, No. 60, 2004, pp. 63-86.

[14] Mikolajczyk K., Tuytelaars T., Schmid C., Zisserman A., Matas J., Schaffalitzky F., Kadir, VanGool T., "A Comparison of Affine Region Detectors", *Int. J. Computer Vision*, No. 65, 2005, pp. 43–72.

[15] Matas J., Chum O., Urban M., Pajdla T., "Robust wide-baseline stereo from maximally stable extremal regions", *Image and Vision Computing*, Vol. 22, 2004, pp.761–767.

[16] Mikolajczyk K., Schmid C., "A Performance Evaluation of Local Descriptors", *IEEE Trans On PAMI*, Vol. 27, No. 10, 2005.

[17] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data—Supplement 1 to the Guide to the Expression of Uncertainty in Measurement—Propagation of distributions using a Monte Carlo method*.

[18] Zhang Z., "A Flexible New Technique for Camera Calibration", *IEEE Trans on PAMI*, vol. 22, n.11, 2000.

CEAS 2011

Space Missions Design and Modelling

3rd CEAS Air&Space Conference
21st AIDAA Congress

ISS: first step in the Human Space Exploration

Piero Messidoro, Claudio Ferro, Eugenio Gargioli
Thales Alenia Space Italia, Torino, Italy

Keywords: *ISS, Enhancements, Exploration, Deep Space*

Abstract

Europe has now entered the full exploitation phase of its contribution to the International Space Station: Node 2 has been on-orbit since October 2007, Columbus, including its internal and external P/Ls, since February 2008, Node 3 and Cupola were launched in March 2010. The Italian Permanent Multipurpose Module, obtained by the enhancement of Leonardo, has been attached to the Node 1 Nadir Port since March 2011. Meanwhile, eleven ISS Increments took place and two successful missions of the Automated Transfer Vehicle were performed, three European Astronauts, Leopold Eyharts, Frank De Winne, Paolo Nespoli, have been part of the ISS Crew and performed a number and variety of experiments on board ISS.

Italy and TAS-I have played a major role in the development of the ISS, and the role remains central now for the operations. After three years, time is mature to draw a line and evaluate the overall team performance when faced with the unique challenges of on-orbit operations. Particularly worth to be highlighted is the change of attitude embedded in transitioning from Design and Development to real Operations on-orbit: this leap is particularly evident in the shorter reaction times, in the flexibility obtained by stretching system performances beyond design and qualification boundaries to adapt to the flight conditions and meet operational needs, in the identification of operational constraints to cope with the resource budgets, in their translation from engineering idiom to operations language through on-board Data Files and Flight Rules, in the integrated management of technical issues raising from anomalies within Columbus and/or on ISS level, in the consequent daily negotiation within IOT/ESA and with the International Partners.

The original ISS architecture has been revised along the assembly Phase with some modules deleted (e.g. the Centrifuge Accommodation Module by JAXA) because of budget constraints, other important elements finally baselined or added, like ERA, Cupola and PMM. The Shuttle retirement has requested a revision of the logistic scenario, which led to the commercialization of the Cargo transportation service to NASA: two consortiums have developed in parallel new transportation systems like Dragon and Cygnus, which are next to their demonstration mission. Similarly, NASA is going to start a program for a commercial crew transportation system. Now, with assembly complete, the ISS horizon is widening: its lifetime has been extended at least to 2020 and talks are on-going to extend it till 2028. Enhancements are on going, to improve the ATV Cargo transportation capability and Columbus functionality via a direct communication link. Call for Ideas have been released by ESA to exploit ISS for experiments for global climate change studies and for exploration. The ISS will become the stepping stone for exploration beyond LEO; new infrastructures have been identified as building blocks of several future exploration scenarios: they will take benefit from the experience gained in developing and operating the ISS.

The paper describes the ISS at its current stage, recalls the TAS-I role in the development of the various elements and in the exploitation activities, describes in detail the improvements already implemented (PMM) and the on-going enhancements (ATV Dry Cargo, Columbus Inmarsat Communication Link), highlights the ISS potential for research and technology demonstration for missions beyond LEO, gives some hints on the Explorations Studies and Architectures, with focus on new infrastructures for deep space missions.

1 Introduction

Europe has now entered the full exploitation phase of its contribution to the International Space Station (ISS): since Node 2 assembly in October 2007, fifteen ISS Increments took place; three European Astronauts, Leopold Eyharts, Frank De Winne, and Paolo Nespoli have been part of the ISS Crew and performed a number and variety of experiments on board ISS.

Italy and TAS-I have played a major role in the development of the ISS, and the role remains central now for the operations. After three years, time is mature to draw a line and evaluate the overall performance and results, from development to utilization, and extrapolate from the acquired know-how and expertise which contribution Europe shall get prepared to provide in the frame of the future exploration scenarios.

1.1 The ISS Assembly

Last year, ISS celebrated the 10th anniversary from Expedition 1, when William Shepherd, Yuri Gidzenko and Sergei Krikalev started the continuous presence of astronauts on-board. In 1998, the construction had started with Node1-Unity and Zarya conjunction. Since then, the ISS has passed through several stages with more than 40 assembly flights, following the sequence originally established and only slightly adapted on the way to the evolving scenario and strategies, to the final assembly in 2010.

The build-up of the ISS marked several achievements for Europe and Italy.

In March 2001, just after integration of the US Laboratory Destiny, the Italian Multi-Purpose Logistic Module (MPLM) Leonardo reached the ISS on the first of several missions. Then, in April, Umberto Guidoni, the first Italian and European Astronaut, visits the ISS on-board the shuttle Endeavour, in the frame of the STS-100 mission.

In October 2007, Node2-Harmony, produced by TAS-I under ESA contract, was attached to the forward port of Destiny: the ISS was finally ready to expand its research capability.

The European Laboratory, Columbus, arrived in February 2008, followed three months later by the Japanese Lab, Kibo. In the mean time, the European Logistic Carrier Jules Verne, the first of the Automated Transfer Vehicles (ATV) fleet, reached the ISS in an autonomous free flight and docked to the ISS in April. Both Columbus and Jules Verne were activated during Increment 16 under the supervision of Leo Heyarts, the first European astronaut member of an ISS crew; since then, the Columbus and ATV Control Centers in Europe have been fully operative in monitoring and commanding from ground the European operations on ISS.

In May 2009, Europe was in the end ready to take over the lead on the ISS, with the first European ISS commander, Frank De Winne, in Expedition 22. The ISS at that time had just reached the capability to support a permanent crew of 6 astronauts.

In March 2010, Space Shuttle Endeavour, in the frame of STS-130 mission, delivered Node3-Tranquility and Cupola, so completing the ISS assembly and reaching its full capability. In December, Paolo Nespoli started the mission MAGISSTRA, becoming the first Italian to join an ISS crew in Increment 26-27.

NASA and the International Partners have agreed to extend ISS lifetime at least to 2020, and talks are on-going to extend it till 2028, and enhancements are under introduction to improve the capabilities. Call for Ideas have been released by ESA to exploit ISS for experiments for global climate change studies and for exploration. The ISS has already become in the end the stepping stone for exploration beyond LEO.



Fig. 1, ISS at Assembly Complete [NASA]

In this endeavor, Italy and Europe have played an important role during the development providing important elements for the ISS completion, now they actively support operations and research, contribute to the enhancements, establish plan for cooperation in future Exploration missions: all in all, we now sit among the Space Powers.

1.2 TAS-I and the ISS

In this environment, TAS-I is one major actor. More than 50% of the pressurized living compartment of the non-Russian section of the ISS has been produced in their Turin plant: Columbus, Node2-Harmony, Node3-Tranquility and Cupola. Out of the four Cargo resupply systems the ISS logistics relied on so far, the Shuttle and ATV make use of pressurized compartments, respectively MPLM and the Integrated Cargo Carrier (ICC), both developed and manufactured by TAS-I. Looking forward, TAS-I has just delivered to Orbital the first Pressurized Cargo Module (PCM) for integration in Cygnus, a new commercial transportation system, bound to a demo mission to the ISS beginning of next year.

A brief overview of the TAS-I elements in the ISS is provided in the following.

The **MPLM** is a large pressurized container, integrated on dedicated logistic missions in the Space Shuttle cargo-bay, to load and transfer dry cargo to and from the ISS. It just terminated its operational phase, due to the Shuttle retirement.

After Shuttle docking to the ISS, the MPLM was transferred with the arm of the Shuttle and ISS Remote Manipulator Systems (RMS), to be berthed to the Unity or Harmony modules. From there, after opening of the hatch, supplies were offloaded, and finished experiments and waste were reloaded. The MPLM was then sealed off again and transferred back with the reverse procedure in the Shuttle, for return to Earth.

The MPLM cylindrical section allows integration of a combination of active and stowage racks in the available 16 bays; during the operative phase, several improvements have been introduced to increase the loading capability. A Common Berthing Mechanism (CBM) at one of the end cones allows berthing to the ISS. A fleet of three flight units (Leonardo, Raffaello and Donatello) was built by TAS-I under a contract with the Italian Space Agency (ASI).

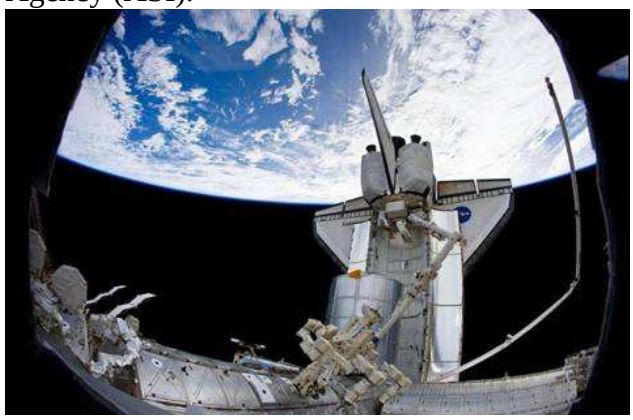


Fig. 2, MPLM as being relocated by the RMS [NASA]

In the 11 years of service, the MPLMs supported 12 successful missions, during which they just consumed a part of their life cycle. This leaves the opportunity to still plan their usage in different mission scenarios: this has already happened to Leonardo, as described later.

The **ICC** has been derived by TAS-I from the MPLM experience: it is similarly composed by a Pressurized Module (PM) for upload of dry cargo and ISS trash disposal, but includes also an External Module (EM) for fluids transport. The EM accommodates tanks for upload of Oxygen and Nitrogen, Potable Water, Propellant; it allows also disposal of ISS Waste Water.

The ICC is designed for assembly to a Service Module produced by Astrium ST GmbH, to compose the ATV. Launched by Ariane5, the ATV is able to reach the ISS in an

autonomous flight; the ICC docks to the Russian module Zvezda, so that it implements a Russian Docking Adapter (RDA) instead of a CBM. A fleet of 5 ATV was ordered by ESA: the first one, Jules Verne, was launched in 2008 and successfully completed the test flight, after 6 months on the ISS. The data gathered during the mission have suggested several improvements that were studied and implemented on Johannes Kepler for the second ATV mission, from February to June 2011.



Fig. 3, ATV Johannes Kepler approaching ISS [ESA]

Edoardo Amaldi, the third ATV, is already in Kourou at the Centre Spatiale Guyennaise preparing for its mission: the launch is planned in February 2012. The fourth ICC is being integrated in these days in TAS-I Turin plant, while the primary structure of the fifth ICC is already under welding in the workshop.
The **PCM** is conceptually similar to the ICC, but scaled down to cope with the Taurus launcher, assigned as vector for the Cygnus program. A dedicate hatch had to be developed in order to optimize the design to the passageway required by the typical cargo transfer for the maximum payload volume allocation.

The **Node2-Harmony**, developed under ASI contract and bartered between ESA and NASA, acts as a building block to connect ISS system elements. It provides a pressurized passage way between berthed elements, and distribute / transfer commands and data, audio and video, electrical power, atmosphere, water, thermal energy to adjacent elements. The internal volume of Node2 is mainly dedicated to the crew passageway; it can be considered in two halves. One half has a single docking port on one of the end cones where Node2 is docked to the station. This half also accommodates eight standard-sized racks, which will house relevant systems and equipment. The other half consists of an additional five docking ports, one on the other end cone, to attach the Pressurized Mating Adapter (PMA) for Shuttle docking, and four arranged around the circumference of the cylindrical main body of Node2, to allow the expansion of the ISS.

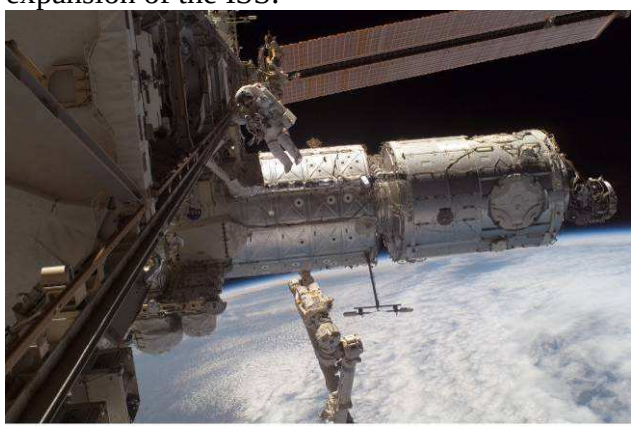


Fig. 4, Node2-Harmony attached at ISS [ESA]

Node2 was launched by the Shuttle mission 10A in October 2007 and docked to the ISS during Increment 16; it has operated since then, supporting 14 visits of the Shuttle, docking and operation of two major elements, Columbus and Kibo, and two missions of the Japanese carrier H2 Transfer Vehicle (HTV).

Node3-Tranquility is similar, but includes a specific functionality dedicated to complement and improve the ISS habitability and operability. The **Cupola**, attached to the Nadir port of Node3, provides a panoramic control tower for the ISS, a dome-shaped module with windows through which operations on the outside of the station can be observed and guided. It is a pressurized observation and work area that accommodates command and control workstations and other hardware. Through the robotics workstation, astronauts are able to control the space station's robotic arm, which helps with the attachment and assembly of the various station elements. Spacewalking activities can be observed from the Cupola along with visiting spacecraft and external areas of the station, with the Cupola offering a viewing spectrum of 360 degrees.



Fig. 5, Paolo Nespoli in Cupola [NASA]

The Cupola is also an observation deck for scientific applications in the areas of Earth observation and space science for Earth and space viewing; this capability is also very appreciated by the crew. Its dome is a single forged unit with no welding. This gives it superior structural characteristics. The Cupola is a "shirtsleeve" module with six trapezoidal side windows and a circular top window of 80 cm in diameter, making it the largest window ever flown in space. Each window is built using very advanced technologies to defend the sensitive fused silica glass panes from years of exposure to solar radiation and debris impacts.
Node3 and Cupola were produced under ESA contract and bartered to NASA; they reached ISS with the Shuttle mission 20A and were integrated during Increment 22 in February 2010.

The **Columbus** laboratory is the cornerstone of the European Space Agency's contribution to the International Space Station (ISS) and is the first European laboratory dedicated to long - term research in space. Columbus supports sophisticated research in weightlessness, having internal and external accommodation for numerous experiments in life sciences, fluid physics and a host of other disciplines. Ten of the 16 internal bays are fully outfitted with resources (such as power, cooling, video and data lines), to be able to accommodate American and European experiment facilities. This extensive experiment capability of the Columbus laboratory has been achieved through a careful and strict optimization of the system configuration, making use of the end cones for housing subsystem equipment. The central area of the starboard cone carries system equipment such as video monitors and cameras, switching panels, audio terminals and fire extinguishers. Although it is the station's smallest laboratory module, the Columbus laboratory offers the same payload volume, power, and data retrieval, for example, as the station's other laboratories. In addition, the External Payload Facility (EPF) accommodates up to four payload pallets for experiments directly exposed to space. Columbus has been delivered to ISS during the Increment 16 by the Shuttle 1E mission in February 2008.



Fig. 6, Working in Columbus [NASA]

Columbus has been always up and running, supporting basic and applied research, experiments, and screening tests for crew physiology, up to the current Expedition 29. The Columbus Control Center (COL-CC) has operated the European Laboratory from ground since the launch day, all year around in a 24x7 shift. The European ISS operation services have been delegated to an **Industrial Operations Team** (IOT), with Astrium ST GmbH as Prime Contractor and TAS-I as major co-Contractor. The IOT contributes to the definition of the Increment Requirements with NASA and the other International Partners, performs the Increment Analysis, to confirm the objectives and allocate the resources to operate the Payload complement in Columbus, populates the Flight Control Team that operates Columbus from COL-CC in Munich-D, provides the necessary engineering support from the Engineering Support Centers in Bremen-D and Turin-I. IOT also takes care of the European logistics, managing the Spare Pool and the up-down traffic of European resupplies.
Tenth of experiments have been performed in Columbus so far thanks to the 9 pressurized facilities, 2 of them, the **European Drawer**

Rack (EDR) and the **Fluid Science Laboratory** (FSL), have also been developed and assembled in Turin, as well as **SOLAR**, an externally exposed platform for Sun observation, all under ESA contract.

2 The ISS Enhancements

Being fully entered in the operational phase and in view of the extended lifetime, it shouldn't surprise so much that the ISS needs already some improvements. For example, the Shuttle retirement dictates the need of a successor for crew transfer, to complement the Soyuz capability, which is remaining the only transportation system in the gap.

The ISS communication link relies on S-Band and KU-Band Antennas, so that no broad band communication is possible. From the 90's, when the ISS design was consolidated, the information technology has progressed a lot: just to say, at that time the World Wide Web was just becoming a mass phenomenon and modems were running at 56 kbps. Astronauts nowadays use Internet to relate with families, to answer questions from students and to twitter with space fans in general. Although the current provisions are good enough to communicate between ground and crew and exchange commands and data between Control Centers and ISS, a state of the art communication link would be desirable to sustain properly Internet traffic between astronauts and Earth.

But one of the most urgent issues to solve is related to stowage, which requires volumes that were underestimated in the architectural studies of the ISS. PMM is meant to fix the problem.



Fig. 7, Typical ISS Stowage before PMM! [NASA]

2.1 PMM

The 24th of February, the PMM has started its journey to the International Space Station aboard the Space Shuttle Discovery, on its last trip before retirement. After docking on the 1st of March to the Space Station Node 1 nadir port, it became a permanent part of the Station.



Fig. 8, Discovery at launch [NASA]

The PMM program is the result of a bilateral agreement between ASI and NASA as a continuation of their past excellent cooperation, consolidated via the "Barter Agreement" on the development and exploitation of the MPLM. The idea of enhancing an MPLM to become a permanent element of the ISS has always been viewed favorably for the benefits it could bring to the on-board operations; and it became even more solid considering that, as a consequence of the Shuttle retirement, the MPLM life cycle would have been completed without reaching the design life limit. It finally materialized in the PMM Program, meant to fully exploit the MPLM fleet even beyond its original purpose.

PMM resolves the ISS stowage issues, improving onboard operations and saving crew time in favor of utilization. Several studies, carried out by TAS-I on behalf of ASI and NASA, demonstrated the feasibility of converting a MPLM flight unit in a module that can remain permanently in orbit, docked to the ISS for at least 10 years.

The MPLM fleet (Leonardo, Rafaello, and Donatello) typical mission foresaw several (up to 25) short duration (max two weeks each) journeys to the ISS. Although significant features need to be adapted to enhance the MPLM to become a permanent element, the

incredible robustness and flexibility shown along all the missions reinforced the confidence in the community about the feasibility and final achievement of the transformation. The commitment to conduct the conversion was finally taken in September 2009, with the start-up of studies to identify the specific changes needed.

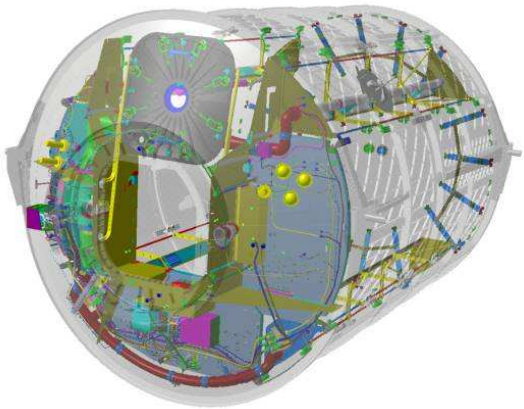


Fig. 9, PMM Cross Section [NASA]

Leonardo was selected as candidate because of its perfect operational record, achieved during its eight successful missions conducted between 2001 and 2010; Leonardo, in particular, had been already modified in 2008 to augment the cargo transportation capability, by introduction of dedicated structures in the aft cone. The PMM project, finally authorized by ASI, foresees mainly three types of changes.

The first change is required by the need to grant the crew access to the equipment, for module operation and maintenance on orbit, considering that for MPLM the maintenance operations were only performed on the ground. So, for example, all the wiring of the module on the Fwd cone structural panels have been repositioned allowing the on-orbit equipment panels opening and closing in a short time without electrical disconnection needed.

The second modification concerns the tailoring of the system configuration to the specific PMM mission, by removing those parts not needed for a long stay at the ISS. The resulting tare mass reduction allowed releasing payload mass for the STS-133 mission manifest.

The third set of modifications resulted as the most critical, including all the analyses and adaptations required to make the module compatible to the space environment over the long term mission. To achieve this, the primary structure has been re-certified as well as all equipment and materials to extend their operational life. This process has requested engineering analysis for all the subsystems on board plus the replacement of some components with other of new construction (for example all the seals toward the open space have been replaced). In this context, certainly the most significant change was the strengthening on the module external of the Meteoroid and Debris Protection System (MDPS), to guarantee the required level of Probability for No Penetration (PNP), in case of possible impacts of Micro-Meteorites, Objects and Debris (MMOD), i.e. high-speed fragments in general along the 10 years of operational life in orbit. Rather than change the external metal shields, which would have had unacceptable mass, cost and schedule impacts, it was agreed to place a reinforcement blanket under the metal shielding and in between the thermal insulation and the primary structure: the blanket is made out of two layers of Nextel and 3 layers of Kevlar assembled in a double "bag" of aluminized Kapton. The reduced thickness of this assembly (about 3 mm) and low density mass (about 1.5 kg / m^2) have helped minimizing the overall impacts on the module. A PNP analysis was performed with simulation techniques to tune the configuration to the required MDPS performances; the relevant mathematical models have been validated by a Hyper Velocity Impact (HVI) test campaign on representative samples. The selected configuration foresees 50 Kevlar-Nextel blankets, distributed in the areas with higher probability of impact.



Fig. 10, Leonardo as PMM in the SSPF [NASA]

The design modifications were established and baselined by TAS-I. They were intended as a modification kit, produced in Turin (Italy), for later implementation in the Space Station Processing Facility (SSPF) at Kennedy Space Center. In the period from April to July 2010, upon Leonardo return from the 19A mission, a TAS-I Team reconfigured the module as planned, in close cooperation with NASA and Boeing team.

The PMM project was carried out by TAS-I as Prime contractor under an ASI contract, with the involvement of ALTEC, Selex Galileo and CGS.

2.2 ATV Cargo Enhancement

Following the successful Jules Verne mission, ESA and industry started an initiative to improve the upload performance of the ATV. Over the development time of more than 10 years, the requirements in dry cargo, but also propulsive support to the ISS have been changed. The delays in the development led to a much later ATV launch schedule and due to the lower solar activities in these later years the propulsive support requirements are reduced and higher need for dry cargo has been identified. The modified requirements push for further design optimization of the dry cargo accommodation in order to make best use of the ATV capability, allowing a better satisfaction of the ISS logistics needs, hence a higher ESA compensation value for each single ATV mission. The ICC layout has been based on the Rack concept for dry cargo accommodation. The Rack is conceived as a multi-purpose, highly flexible structure, able to accommodate every type of container, from Crew Transfer Bags (CTB) to Mid Deck Lockers (MDL), or hard-mounted Cargo. Whereas at the beginning of the ATV development hard mounted cargo items were expected to be placed in the racks, experience has shown that soft cargo bags do nowadays fulfill the ISS logistic needs.

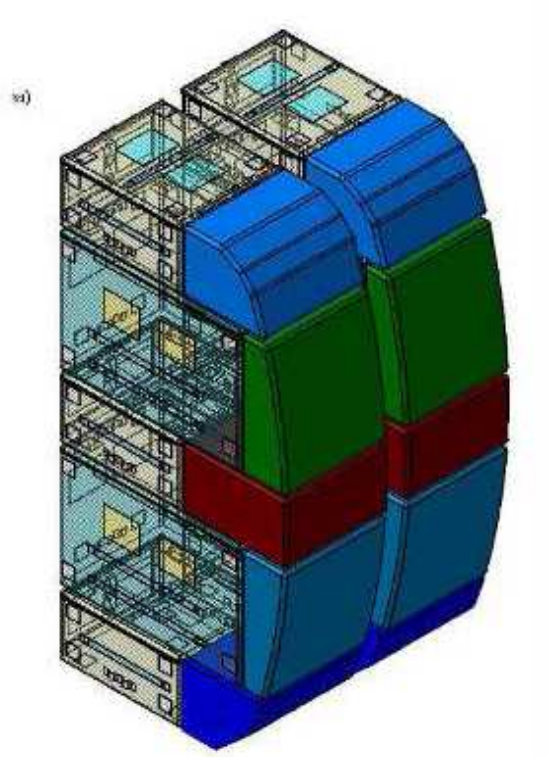


Fig. 11, ATV Cargo Rack

In addition, the average Cargo mass transported by the first two missions has been much smaller (roughly 11 kg/CTBE[1] vs. the design requirement of 25 kg/CTBE), and projections for the next missions do not show a significant increase; such density is also indicating that likely the anisotropy (acting in terms of CoG offset) is much lower than what considered for the Rack verification. This resulted in an overestimation of the Rack and ICC Dry Cargo capability, which is actually constrained by the available volume, measured in CTBE. Already for ATV Johannes Kepler, the existing Racks have been redesigned to gain about 60 kg of tare mass per rack and 30% of cargo usable volume. For the future ATV missions the transportation efficiency can be further increased by introducing alternative secondary structures designed to more specific requirements. The aim is to combine a set of different optional structures to tailor the ICC internal layout to the specific ATV Mission Manifest, reducing the

[1] Crew Transfer Bag Equivalent

ICC tare mass and exploiting volumes left currently free.

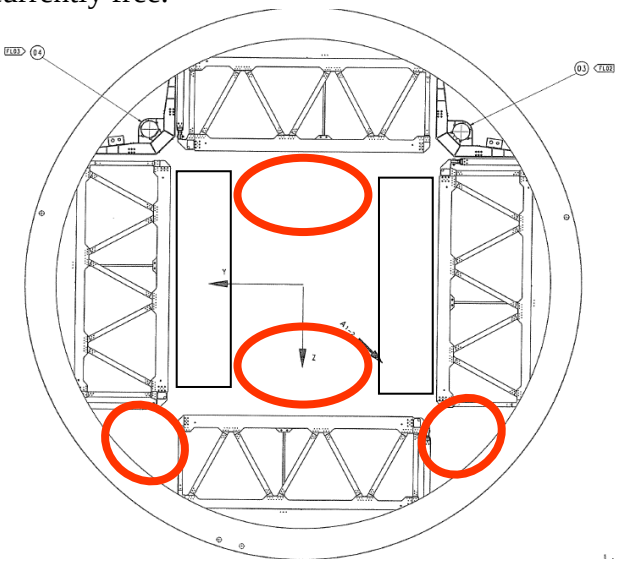


Fig. 12, Volume free in the ICC Cross Section

This is the core of the ATV Dry Cargo Enhancement program, which was awarded by ESA to TAS-I in the first half of 2010. A challenging and aggressive development and implementation logic had to be adopted. In fact, The ATV Phase E activity flow is strictly based on an "L minus" template: the launch dates for the various ATV missions are already booked, depending on the ISS logistics needs and the visiting Vehicles traffic, as part of an Inter-Agencies high level commitment. The ATV Phase E Schedule is then not flexible to accommodate delays; the ATV Enhancement is fit into the planning on non-interfering basis, i.e. no impact shall result to the ATV production schedule or the mission preparatory activities. Coherently, the implementation approach foresees a parallel ATV Enhancement Program, kicked off in May 2010, that involves three major design changes:

1. The Integrated Support Platform Plus (ISP+), derived from a similar application (the ISP) designed by TAS-I and flown already in Node3 and MPLM missions, which is able to substitute the Rack in some selected bays of the ICC, with similar volume capabilities (21 vs. 24 CTBE) but lower average Cargo density (13 kg/CTBE), hence considerably lower tare and higher transportation efficiency (4.0 kg of Cargo per kg of tare);
2. The Standoff Support Frame (SSF), a new Cargo secondary structure, allowing full exploitation of the lower standoffs volume;
3. The MDPS Lite, benefiting from the less severe environment predicted by state-of-the-art tools like ORDEM$_{2000}$.

The ISP+ has been developed maximizing the recurrence to the ISP and, on the other side, complying with the interface and frequency requirements applicable to the ATV Rack: the ICC Primary Structure will regard the ISP+ as a lighter, less loaded Rack, thus avoiding the need of modifications or system delta qualification.

The ISP+ structure is basically a metallic platform on both sides of which ATV standard CTBs can be accommodated. It is composed by following main items:

- Two main longitudinal longerons withstanding the main inertial loads induced by the dry cargo and providing the interfaces with the primary structure;
- Four transversal frames joining the two longerons, which stiffen the entire structure and transfer the inertial loads from the dry cargo to the longerons;
- Three horizontal panels that, beyond the primary function of providing accommodation on both faces for the CTBs, contribute to the structure overall stiffness. These panels are removable on-orbit by the Crew in order to get access to the underneath bags.

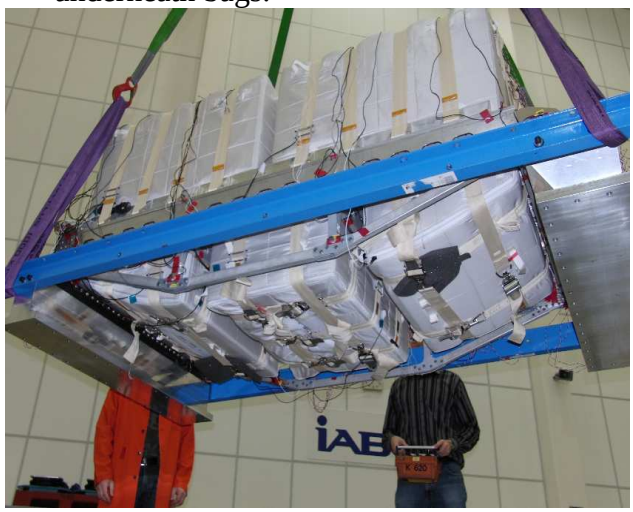


Fig. 13, The ISP+ loaded for the Sine Test

The replacement of ICC with ISP+ in the ICC deck removes the accessibility issues that

prevented the use of the lower standoffs volume on Jules Verne and Johannes Kepler.

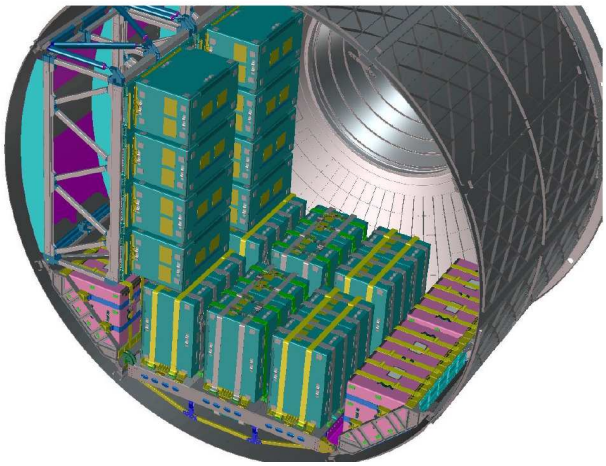


Fig. 14, typical ICC cross section with ISP+ and SSF in the deck

The volume that becomes available is cumulatively equivalent to one ILR, it is then worth thinking of implementing a dedicated structure to make it available for Cargo: the SSF is aimed to exploit the empty volume present in the lower standoffs, resulting in an increase of the capability by roughly 20 CTBE.
The SSF is composed by:

- Three machined aluminum frames that interface with the brackets installed on the longerons;
- Panels providing support to the CTBs;
- Stiffeners.

In order to make the optimum use of the volume available, a new “non-standard” CTB with dedicated shape has been conceived, hence maximizing the volume allocated for the dry cargo.

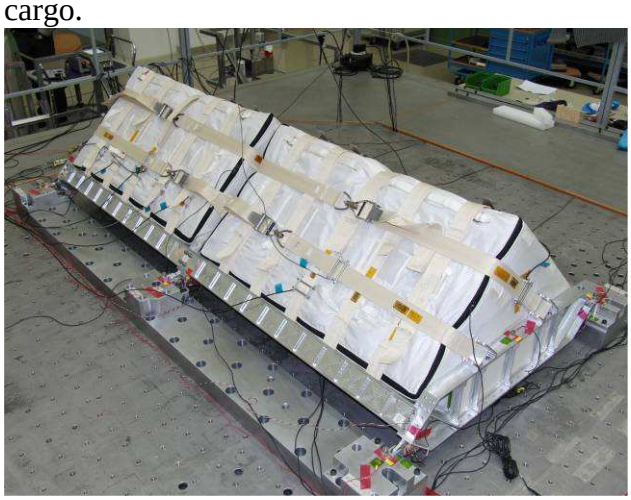


Fig. 15, The SSF loaded for the Sine Test

The MDPS Lite is based on the demonstrated technological capability to reduce the uniform thickness of the aluminum bumper from 1.6 mm to 1.2 mm.

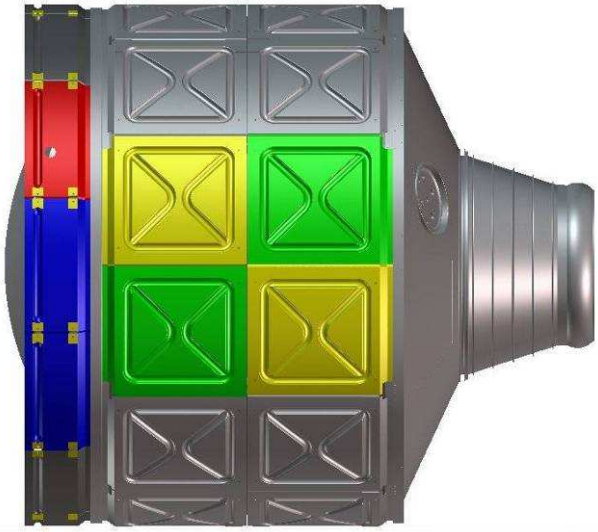


Fig. 16, the ICC Revised MDPS Layout

Samples of the thinner/lighter panels were subject to Hyper Velocity Impact (HVI) tests in order to establish Ballistic Limit Equations for revising the ATV modeling and simulate the reduced MDPS thickness. The Probability of No Penetration (PNP) analysis, run with ESABASE2 and $ORDEM_{2000}$, has given positive indications that the performance of the ATV MDPS, originally predicted by ESABASE1 and $NASA_{90}$, can be confirmed also for the reduced thickness. PNP predictions concluded that the MDPS Lite can be implemented on the Pressurized Module (PM) cylinder section and on the External Module (EM), resulting in an overall tare mass reduction of 60 kg.

The Qualification Review was successfully held in February 2011, positively concluding on the qualification status of the ISP+, the SSF and the MDPS Lite and releasing the manufacturing of the Flight Models. The MDPS Lite is already introduced in the Edoardo Amaldi for the third ATV mission. Four ISP+ and 2 SSF have been completed in these days and they now sit together with the Racks on the shelves of the ATV Cargo Structures. Based on the manifest consolidation for the next missions, ESA is in the process of deciding what to embark in the ATV4 or ATV5 missions.

2.3 CICL-Terminal

The need of increasing the ISS communication capabilities, particularly for the European Missions, was at the origin of several investigations in the recent years for the implementation of a communication terminal on Columbus. Looking forward to the future operational phase, ESA agreed to implement during the development phase hooks and scars on the module for a later introduction of a terminal.

In parallel, an industrial consortium (COM DEV, Inmarsat, and Broad Reach Engineering) is currently developing, in the frame of the ESA ARTES program, a communications terminal for LEO spacecrafts (SB-Sat), with the communication link making use of existing Inmarsat satellites and Inmarsat ground infrastructure.

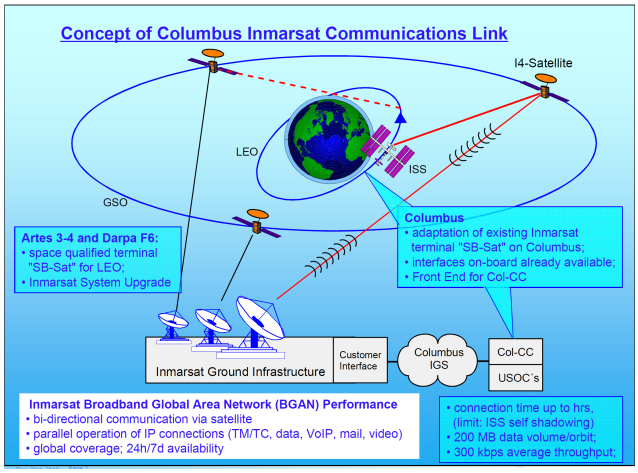


Fig. 17, SB-Sat Flight and Ground Architecture

Therefore, as per the current state of the art, it becomes appealing the possibility to exploit the SB-Sat for Columbus needs, synergizing the efforts of ARTES and ISS Exploitation programs. ESA has then awarded TAS-I, with Astrium ST as Prime Contractor, a feasibility study for a Columbus-Inmarsat Communication Link Terminal (CICL-T).

The architectural concept for the complete system consists of two SB-Sat terminals allowing the link handover between satellites without break of the connection: the Main Terminal (including Antenna, SB-Transceiver, Terminal Controller and Power Distribution System) and the Slave Terminal (with Antenna and Transceiver only).

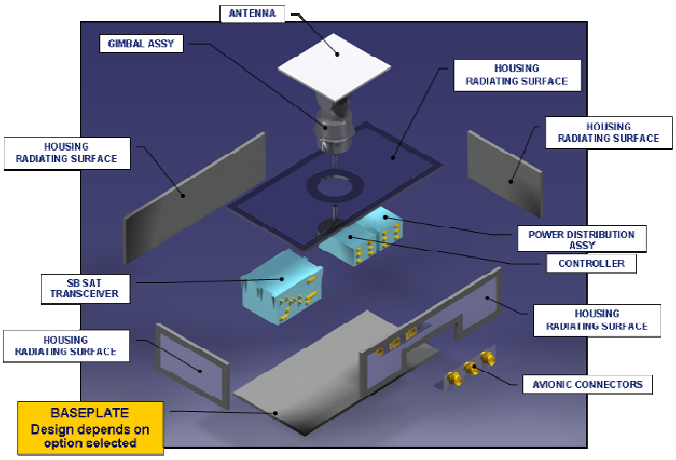


Fig. 18, CICL-T Assembly – Exploded view

The housing containing the components will accomplish the double scope of mechanical support and thermal rejection surface, this last function being found the driving aspect for the overall terminal dimensions. The structural Baseplate will provide the mechanical interface with Columbus, withstanding the launch (if transported hard-mounted) and on-orbit loads. Finally, the Power & Data resources will be provided through dedicated harness cables to be installed on the external module and interfacing with the existing avionic Columbus Stanchions feed through.

The preliminary study regarding the adaptation of the SB-Sat terminal on Columbus was focused to identify the possible accommodation options for the CICL-T and to select the preferable one through a dedicated trade-off.

Five different technical solutions have been conceived and then compared on the basis of functional, mechanical, ergonomic, operational and programmatic criteria.

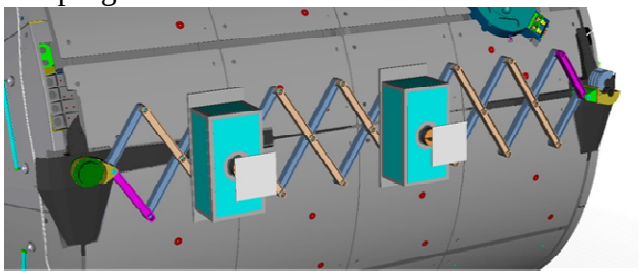


Fig. 19, one of the accommodation options

Since the Terminal will be necessarily integrated on orbit on Columbus, a crucial role was given to the aspects relevant to accommodation at launch, stowage and transfer

on ISS, deployment and installation of the assembly. For this reason, dedicated assessments have been performed to evaluate the compatibility with the Carriers, the possibility to stow the hardware inside standard bags, the potential tasks to assembly on orbit the terminals and the complexity of the required Extra Vehicular Activities (EVA).

The outcome of the trade-off has recommended the implementation of the option exploiting the mechanical interfaces of the MDPS panels (locking device) to secure on top of them the terminal Baseplate.

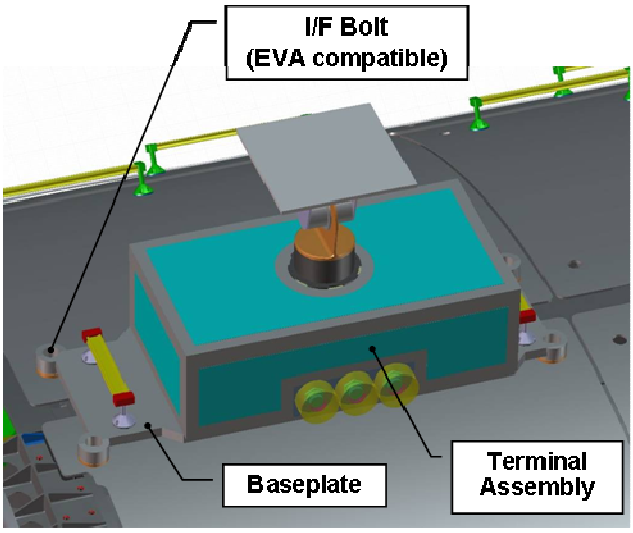


Fig. 20, selected accommodation option

The existing interface bolts will be replaced with longer ones, to be compatible with the new thickness of the entire package (Columbus structure, MDPS Panel and terminal Baseplate).

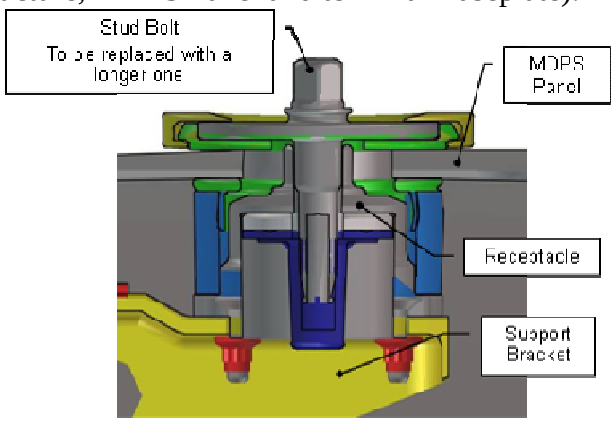


Fig. 21, the MDPS Locking Device

In addition, the overall interface layout of the Baseplate is required to reflect the MDPS panel concept, providing the needed clearances and slots adopted at MDPS level I/F to prevent integration misalignment & thermo-elastic distortion (3 slots and one fixed hole).

The main advantages offered by the selected solution are the reduced mass and volume of the overall system, the simplicity of the mechanical design, the compatibility with all Carriers and with both pressurized and unpressurized transport, the feasibility of the deployment and the relatively low complexity of the EVA for installation. Moreover, it offers flexibility on the choice of the two Terminals locations, giving the possibility to select the best accommodation from thermal, field of view, operational standpoint and to maximize the distance between the two Terminals, with consequent advantage for the Satellites link handover. It would also allow changing the Terminal positions during the program development and, if really needed, to relocate the equipment during their lifecycle.

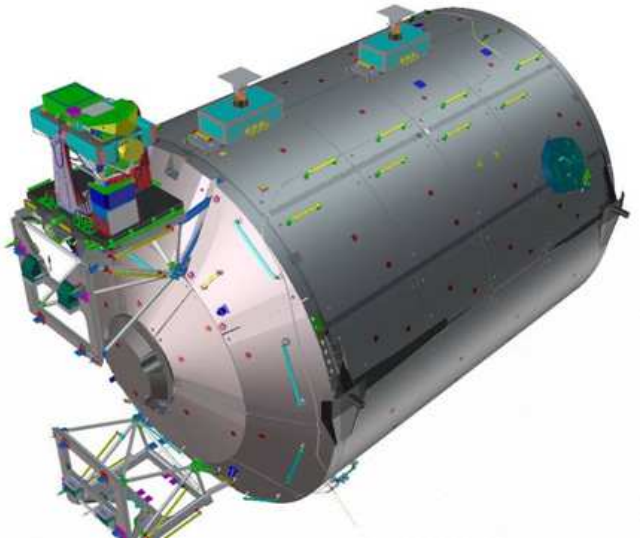


Fig. 22, Overall CICL-T Configuration on Columbus

As a reference, the configuration with both terminals on Columbus Zenith area has been found the preferable one, since allows to maximize the Terminal field of view, to reduce as much as possible the required thermal rejection surface and to leave about 2.5 m of distance between Main and Slave Terminal.

The study is aiming now to the definition of the System requirements first, planned for completion by the end of October 2011, and then of the Sub-System requirements, to be settled by the end of December 2011. In the same timeframe, the design concept will be

consolidated to a maturity adequate to start a C/D phase already in 2012. The CICL-T is planned to become operative in early 2014.

3 ISS as test bed for Exploration

The present configuration of the ISS (with the possible addition of new attached modules) could be envisioned to accommodate and support technological demonstrations, as a preparatory test bed for applications to be exploited in the frame of long duration exploration missions. Various subjects could be considered potential fields for technological enhancement. Some of them appear particularly interesting for their implications, and are briefly addressed hereafter.

Rapid prototyping: the complexity of the systems operated in long human space missions is expected to require innovative methods to face failures, satisfy mission unexpected needs or conditions. Maintenance tasks could require significant storage of spare parts with huge allocation of dedicated mass and volume on the spacecraft. A possible concept to improve the logistic criticality relies on the capability to have directly on-board the necessary design/manufacturing tools and the associated raw materials. This approach requires the development of dedicated machines to be operated with clean and safe processes, with low power consumption and limited waste of material. A demonstrator could be tested on the ISS for the fabrication/replication of parts. The design itself of the future exploration vehicles could be oriented to take these new prototyping capabilities into account.

Collaborative robotics: various possibilities of experimentation can be envisioned, linked with the different phases of an exploration scenario.

- A possible application is related to the testing of a robotic system in tele-presence mode in the conditions typical of a “human arrival imminent scenario”. A robotic device on the planet surface (in this case on ground) could be controlled by a human operator resident in a spacecraft orbiting around the planet (in this case on ISS). The on ground robotic system should include: human robotic hand, anthropomorphic arm(s), mobile platform.



Fig. 23: the EUROBOT Ground Prototype

The robotic hardware could be derived by the re-use / upgrading of existing demonstrators (e.g. Eurobot) complemented by new development where necessary (e.g. hand exoskeleton). The control system on-board ISS should include Augmented Reality devices for immersive environment in addition to the classic laptop based Man Machine Interfaces.

- Another significant field of investigation would be represented by the testing of an IVA robotic support, through incremental steps, towards the development of humanoid robotic assistants for exploration. This could include not only arms and effectors, but also artificial vision (objects and human recognition and tracking), multimodal naturalistic HMI (e.g. voice, gesture control), adjustable autonomy (from semi-autonomous behavior to manual control depending on task).

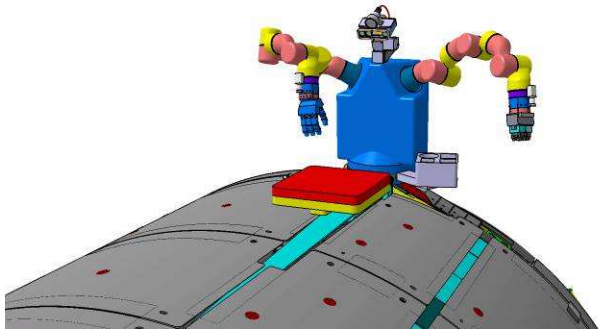


Fig. 24: example of EVA robotic assistant

- Finally the capabilities of robotic assistance could be also tested in EVA operations, as a continuation of the technological effort

developed in Europe in the frame of the Eurobot project.

A flight demonstrator can be envisioned with incremental configurations, evaluating the human-robot cooperation aspects in the various on-orbit operational situations (thus alleviating the astronauts' workloads) and the relevant safety implications.

Power generation: a (possibly scaled) demonstrator for a regenerative fuel cells system could be tested in 0g conditions, so enabling the assessment of new types of electrolizers / fuel cells.

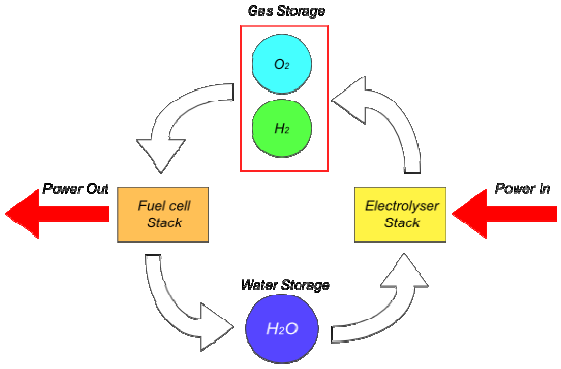


Fig. 25: the water – power closed loop

Inflatable technologies: in order to cope with the volume limitations induced by the current and future launcher capabilities, inflatable structure technologies are considered a promising technological challenge.

ISS could offer the possibility of docking an experimental pressurized inflatable element, endowed with all necessary provisions for Meteoroids and Debris Protections, thermal insulations, and functional monitoring. The Station could also provide a suitable base point for the demonstration of inflatable equipment applications, like antennas, booms, radiators.

Bio-regenerative Life Support: some dedicated volumes of ISS (that could be obtained also through the aforementioned inflatable provisions) may accommodate research activities devoted to the food production and management.

Activities have already been conducted by TAS-I on a ground based facility (EDEN), a demonstrator of a rack-like cultivation system. These functions could be coupled as well with the recycling of water and regeneration of O2 from CO2.



Fig. 26: Lettuce cultivated in the EDEN growth chamber

Environmental protections: besides MMOD protections, also techniques of hybrid (active and passive) radiation shielding and monitoring could be located at the ISS elements (one module or part of it).

4 Exploration Scenarios

While preparing for exploration, strategies for future missions are under definition throughout the world and also at ESA. Since the Apollo program conclusion, the most important international space agencies developed scenarios aimed to progressively expand the human presence beyond the Low Earth Orbit. The ultimate goal of all the various exploration scenarios is a human mission to Mars, but the high complexity and risks associated to a Mars manned endeavor force to conceive intermediate missions to develop the required technologies and to validate them in space.

Recently, the Agencies participating to the development of the ISECG Global Exploration Roadmap have agreed on the following exploration destinations of interest for human visits:

- Low Earth Orbit;
- Cis-lunar Space;
- Moon;
- Near Earth Asteroids (NEA's);
- Mars and its moons, Phobos and Deimos.

The ISS is the first important step towards human expansion into space and, following a

step-by-step approach, both Asteroids and the Moon are considered important destinations to contribute preparing the future human mission to Mars. The Global Exploration Roadmap identifies two pathways for human exploration missions after ISS:

1) Moon First
2) Deep Space First

which mainly differ with regard to the sequence of sending humans to the Moon and Asteroids.

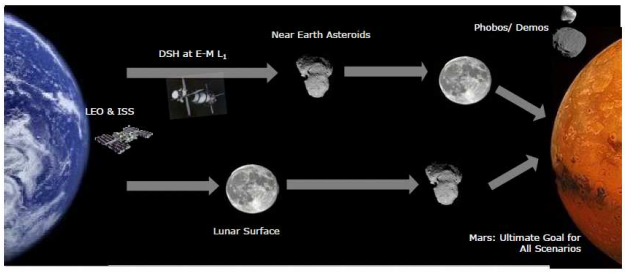


Fig. 27, Optional Strategies for Future Human Exploration Missions

The identification of the two optional pathways provides an initial framework for assessing, in synergy with the major European space industries, the priorities of activities needed to prepare the human exploration missions and for identifying the possible international cooperation to be established to achieve the proposed goals.

An industrial consortium led by Thales Alenia Space - Italia has been awarded an ESA Study, aimed to develop a coherent exploration scenario, in terms of building block elements, development approaches and operational aspects. In addition, the two different strategies imposed the definition of two separate approaches to develop and demonstrate the capabilities ultimately required for the human exploration of Mars. In light of the different technologies and capabilities that have to be acquired to implement the Mars mission, four main frameworks were identified and analysed: ISS, post-ISS, Moon and Deep Space.

4.1 ISS framework

An extension of the ISS lifetime until 2025 or 2028 is considered mandatory by the European Space community, to support a full exploitation of the existing ISS focused on the development, testing and demonstration of the technologies required to achieve the further exploration steps: fields like automation and robotics, cryogenic management, advanced ECLSS, novel energy sources, advance re-entry technology, and inflatable habitat, will decisively benefit from the usage of the ISS as a stepping stone towards exploration. In particular, Europe may contribute with the following projects:

- A robotic crew assistant (i.e., Eurobot);
- An Advanced Life Support Demo;
- An Inflatable demonstrator element possibly followed by an operational habitat;
- The MPCV Service Module development;
- Participation to the provision of logistic services system (i.e. based on the MPCV Service Module).

4.2 Post-ISS framework

In parallel to the ISS utilization, the development of a man-tended orbital infrastructure will secure continuity of Low Earth orbit research in the European priority fields and will serve as test-bed for future exploration technologies after the ISS disposal.
Three different approaches may be envisaged:

- Building-up a minimal infrastructure composed by a free-flyer infrastructure operated by periodic crew visits;
- Building-up of a complete post-ISS infrastructure in LEO with possible usage of existing ISS modules being not beyond lifetime;
- Building-up of a complete new post-ISS infrastructure in LEO for space exploration.

Even if major achievements deriving from a complete LEO infrastructure exploitation are doubtless, budget considerations leads to prefer a minimal infrastructure, securing potentials for subsequent expansion. Potential European contributions to the Post-ISS architecture are:

- Service Node: It is a combination of a Node and a Service module and it will function as a connection point for the other modules providing the initial resources for the entire infrastructure (i.e. power, crew quarters, attitude control, communications);

- Laboratory Module: based on the Columbus design, the module will permit the execution of the required experiments (i.e. life science research and microgravity effect counter measures in preparation of future space exploration missions, study of the behavior of a bio-regenerative life support system under long-term exposure to the space environment);
- Habitation Module, Green House and Inflatable Module.

4.3 Moon framework

The proposed lunar exploration scenario merges the already foreseen agencies robotic missions (i.e. GRAIL, LADEE and CHANG'E 2), private entities intentions (i.e. Google Lunar X Price missions) and the results achieved by the international exploration groups in a coherent architecture aimed to maximize international collaborations (i.e. sharing capabilities and minimizing useless redundancies).

The already planned robotic missions will enhance the current knowledge on the Moon environment and validate new technologies and capabilities (i.e. the ESA Moon Lander) that will be used afterwards to prepare the following human mission. After the robotic phase, a man-tended period will precede the first human mission that has been foreseen after 2030. During the initial man-tended phase, an unpressurized rover will be landed on the lunar surface to provide mobility capabilities to the future astronauts. Afterwards, in order to extend the exploration range, two pressurized lunar rovers (i.e. the NASA SPR or the ESA PLR) will be delivered to the Moon surface via a man-rated large lander.

Potential European contributions to the Moon exploration architecture are:

- Large size robotic landers (1000-1500 kg), namely the ESA Moon Lander;
- Lunar Rovers, both pressurized and unpressurized;
- Lunar Communication/Navigation satellites
- Power provision systems;
- Man-rated ascent vehicles (i.e. Constellation Altair Lander size).

In parallel with the definition of the landers, the considered transportation chain foresees the presence of already existing launchers (i.e. ESA Ariane 5, Soyuz) and future heavy lift rockets.

4.4 Deep-Space framework

Besides the Moon, the Near-Earth Objects become the primary targets of human exploration in order to understand the key secrets of the evolution of life and to prepare the future Manned Mars Mission.

Deep Space Missions are extremely expensive and technologically challenging and the establishment of international cooperation is mandatory. Different strategic schemes have been considered:

- "technology dependent" - each partner is concentrated on a specific set of technologies/capabilities that can be shared among different elements (e.g. life support systems);
- "function dependent": each partner is concentrated on the development of particular elements that can depend on functions provided from the elements of another partner (e.g. ISRU Plant);
- "mission dependent": each partner develops all the elements that are relevant to accomplish its mission (e.g. Mars Orbiter), that work independently from the ones provided by the other international partners that are part of the mission.

The proposed scenarios foresees the implementation of a robotic precursor mission to a Near Earth Asteroid previously identified as possible candidate for a subsequent human mission, in the 2022-2025 timeframe. In parallel, the development of a Deep Space Habitat will be crucial to perform the manned mission to the selected NEA. The spacecraft will be initially stationed in cis-lunar/EML1 space to test system and subsystems and then used for the first human missions to a Near Earth Asteroid in the 2028-2030 timeframe.

4.5 ESA Positioning in the Exploration

Coherently with the above, in the frame of the two exploration pathways ("Moon First" vs.

"Deep Space First"), the following European initiatives have been already activated:

- a 1 mT Cargo Lunar Lander;
- a Lunar Un-Pressurized Rover;
- a Service Module for the Multi-Purpose Crew Vehicle (MPCV), both scenarios, as a barter with NASA for compensation of ESA contributions to the ISS Common System Operations Costs beyond 2016.

In addition, both scenarios foresee the presence of a key element for the future manned exploration activities: the Deep Space Habitat, for which, in order to finally secure a leading international role, Europe shall exploit the capabilities and the heritage gained in participating to the ISS endeavour.

5 Deep Space Habitat

The Deep Space Habitat is then the core of any future space exploration mission. The ISS experience shall be exploited to develop a module able to support different human missions towards deep space targets. The module will have some specific characteristics deriving from the peculiarities of the mission and of the environment it has to withstand: this strongly influences the design of the pressurized habitat, where the astronauts have to live and operate for quite a long period.

According to the necessity of a habitation module to enable travels beyond LEO, TAS-I has carried out a preliminary analysis of a possible architecture for the Deep Space Habitat. It can be conceived either as part of a cis-lunar orbital infrastructure or of a space-ship for deep space exploration missions. It would represent the first human outpost beyond LEO; in particular, it is likely that at first it will to be deployed in the first Earth-Moon Lagrangian point (EML1), as a platform for research and meant to demonstrate a set of critical technologies and associated operations required by a deep space human exploration mission (e.g. to a NEO).

In fact, an infrastructure in EML1 would allow, more than in LEO, to experience the conditions typical of the journey towards an asteroid or Mars, thus guaranteeing long term testing of specific technologies. For example, one of the major issues of human space missions is represented by the long exposure to space radiations; beyond the protection of the Van Allen belts, the Deep Space Habitat would allow analyzing the effects of radiations on human body as well as testing the effectiveness and duration of the radiation protection systems. In addition, the psychological effects of a long permanence far from Earth can be properly analyzed, before moving further towards deep space targets.

The Deep Space Habitat in EML1 would also offer the possibility to test and check new long term autonomous systems, e.g. regenerative ECLSS, which represents an important point in the design of long duration missions due to the strong constraints in terms of mass. Moreover, in-situ diagnostic and maintenance capabilities could be improved in view of more challenging missions, where coming back to Earth is not a possible option in case of failure of any equipment. Another crucial aspect for missions very distant from Earth is that any acute illnesses or injuries that might happen to the crewmembers have to be dealt with on board of the spacecraft. For this reason, tele-operated surgical robotic systems shall be adopted: the Deep Space Habitat deployed in EML1 could allow testing of these technologies, in order to identify the most significant criticalities and improve them before their adoption in a real deep space mission.

Besides being a test bed for new technologies, the Deep Space Habitat deployed in EML1 is envisioned to support lunar human exploration missions, providing a staging post and a safe haven for crew working on the Moon surface, increasing the science return from lunar surface robotics, providing servicing of transportation system elements.

The Deep Space Habitat is intended as a human-tended facility, and visits of crew of four astronauts are periodically foreseen. The experiments loaded on board are controlled and monitored from ground from the control and support centre. Continuous direct communications with ground, with no need for satellites, are envisaged, relying on dedicated antennas. Furthermore, the module is conceived as an expandable infrastructure allowing for

later docking of additional modules, such as logistics storage modules, laboratories for scientific research or a module for tourism.

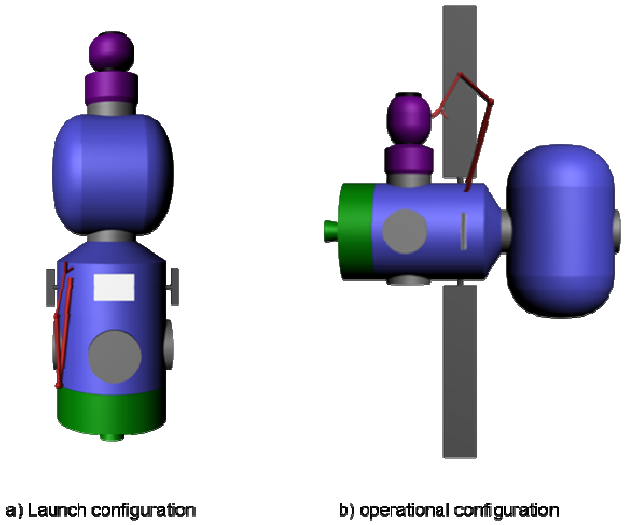


Fig. 28, Deep Space Habitat architecture

The overall architecture of the module has been derived from a set of system trade-off performed accordingly to the objectives to be accomplished. The Deep Space Habitat is composed of a Rigid Node attached to an Inflatable Module. The presence of the node with its 4 radial ports ensures the possibility to have 3 visiting vehicles simultaneously attached and to eventually expand the module. The implementation of the inflatable technology is foreseen since, in view of very long missions, the comfort of the crew becomes a more and more significant design parameter. The rigid node is axially attached to a Propulsion Module (depicted in green in Figure 1), not considered as part of the Deep Space Habitat system, which is in charge of providing orbit/attitude control.

One of the four radial ports of the node is used for attaching the airlock, which has been introduced in the architecture because different EVAs are to be performed (for external maintenance, for exploration, for managing external payloads). Moreover, additional EVA support items are envisaged, such as Enhanced-Manned Manoeuvring Units (E-MMU), EVA tools, etc. The airlock is characterized by a rigid Equipment Lock and an Inflatable Crew Lock. E-MMUs and EVA tools are stored in dedicated compartments on the external surface of the Equipment Lock.

For protecting the crew against radiation a passive shielding is envisioned. In particular the protection provided by the structure and equipments is sufficient for protecting against GCR, while a dedicated high density polyethylene shelter is envisaged for protecting the crew against SPE.[2]

A robotic arm is introduced in the architecture to reconfigure the module from the launch to the operational configuration and to support external maintenance.

The habitat is sized to ensure an overall crew habitable volume of at least 80m^3 (that means optimal 20m^3/crew member) and an overall pressurized volume of ~240m^3. The main features of the pressurized elements can be synthesized as follows:

- The Inflatable Habitat, characterized by a rigid core and multi-layer wall, has a total pressurized volume of ~165m^3, with an external size of ~8m x 5m.
- The Rigid Node is sized to guarantee a pressurized volume of ~84m^3, with an external size of ~ 4.5m x 5m.
- The Airlock is characterized by an equipment lock of ~ 2m x 1m and an inflatable crew lock of ~ 2m x 2m.

The overall mass of the Deep Space Habitat amounts to about 25 tons, including resources and crew systems sized for 20 days of maximum stay of a crew of four astronauts. The module is deployed with this amount of resources and its resupply is foreseen with the periodic visits of the crew.

In a transportation system, the Deep Space Habitat becomes the crew quarter where the astronauts live and operate during the transfer to an asteroid or to Mars. In this respect, the module described above can be regarded as a precursor for an exploration mission. The experience gained and the technological demonstrations achieved in EML1 will be exploited to adapt the design of the common core for the deep space journey. Architectural changes are expected, in order to integrate the

[2] These evaluations refer to a NEO reference mission lasting 1 year.

Deep Space Habitat in the transportation system. As a matter of fact, in a NEO (or a Mars) mission the habitation module will be part of more complex transportation architecture and will likely have different interfaces with the propulsive module to which it is attached. In addition, the three free radial docking ports will not be necessary, while an additional axial docking port would be necessary for safety and operational complexity reasons.

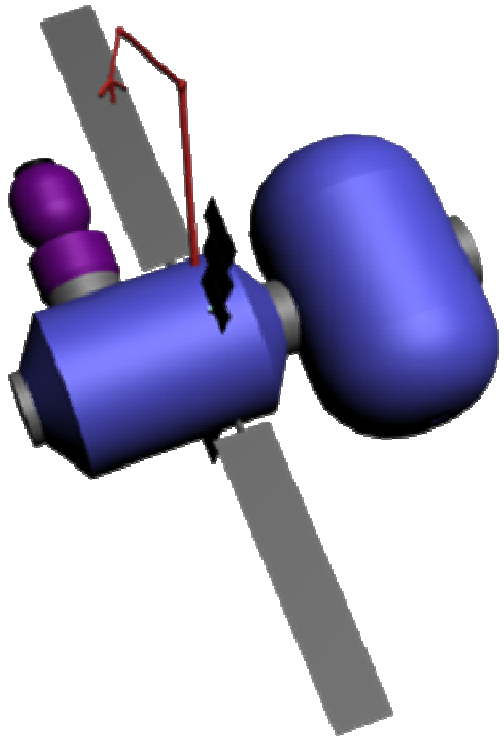


Fig. 29, Deep Space Habitat in a Transportation System

6 Conclusions

The ISS has been decisive for Italy and TAS-I to step up and join the Space Powers.

The role achieved in the ISS development, build up, operations and utilization will be finally exploited to participate together with the International Partners to the initiatives under consolidation in the frame of the future missions beyond LEO, the Deep Space Habitat being the natural successor of the ISS.

References

[1] M.A. Viscio, F. Fenoglio, "*Exploration Research Habitat concept*," SD-RP-AI-0735, TAS-I, 2011.

[2] M.A. Viscio, F. Fenoglio, "*Crew Transportation BB*," SD-TN-AI-1348, TAS-I, 2011.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Solar Array Simulators for Low Power Space Missions

F. Topputo, F. Bernelli-Zazzera, and F. Ercoli-Finzi
Politecnico di Milano
Dipartimento di Ingegneria Aerospaziale
Via La Masa, 34 – 20156 – Milano (MI), Italy
francesco.topputo, franco.bernelli, amalia.finzi@polimi.it

F. Lazzari
CBL Electronics
Voc. Bodoglie, 148/P/3 – 06059 – Todi (PG), Italy
fabrizio.lazzari@cblelectronics.com

Keywords: *Solar array simulator, power subsystem, solar cells.*

Abstract

In this work we discuss issues related to the simulation of low power systems with hardware means. Simulating low power systems is a challenging task as the possible low-intensity, low-temperature environment, together with possible dust deposition and ice condensation, worsen not only the production of power but also make it difficult to predict it. To overcome these problems, we have developed solutions in terms of software and hardware tools for power estimation and simulation. The developed low power, hardware solar array simulator system is briefly discussed in this paper. Although this solution is reported for the case of Rosetta lander Philae, it applies also to possible low power future missions aimed to perform in-situ operations on comets and asteroids.

1 Introduction

Rosetta is the third of ESA's cornerstone missions within the science program Horizon 2000. The ambitious goal of this mission is the injection into an orbit around a comet and the delivery of a lander that will perform a detailed in-situ investigation of a comet nucleus for the first time. The prime scientific objective is to study the comet 67P/Churyumov-Gerasimenko to help understand the origin and the evolution of the Solar System. The comet will be reached in 2014 after a journey involving several close encounters with minor and major bodies. After a phase of close comet investigation, a safe and scientifically important site will be selected for in-situ investigations. The lander delivery is foreseen in November 2014 at a distance of about 3 AU from the Sun.

The Rosetta probe is made up by two spacecraft: an orbiter, Rosetta, with 11 scientific instruments on-board, which will orbit the comet 67P/Churyumov-Gerasimenko, and a lander, Philae, with 10 scientific instruments, which will land on the comet to perform in-situ analysis. This large array of instruments will perform the most extensive study of a comet to date. Philae can be considered as an independent spacecraft, although it is the main payload of Rosetta. Philae has a mass of 97.9 Kg including 26.7 Kg for scientific payload.

At the time this note is being written, Rosetta is approaching the comet 67P, at about 5 AU from the Sun. Both Rosetta and Philae have been placed into hibernation due to the weakness of the sunlight, which prevents from producing enough power to fully operate the probe. The wake-up is foreseen for January

2014, six months before the rendez-vous with the comet.

Politecnico di Milano is involved in the Rosetta mission through the activities on the lander Philae. In particular, the Aerospace Engineering Department is responsible for the activities of the solar generator and the instrument SD2 (Sampler Drill and Distribution subsystem [1]).

2 Philae Power Subsystem

The Philae Power Subsystem (PSS) manages all the electrical power needed by the lander during its entire lifetime. The main source of power is represented by the primary and secondary batteries, with capacity of 1000 Wh and 130 Wh at comet arrival (1200 Wh and 150 Wh at launch, respectively). The primary battery will allow the operations of Philae during the main scientific phase (about 5 days) just after the landing on the comet, during which all instruments will be operated at least once. The secondary battery will be recharged with the power produced by the solar arrays.

The power distribution is basically performed using the Philae Primary Bus. The main subsystems, the Command Data Management System (CDMS) and the Thermal Control Units (TCU), are directly connected to the primary bus via dedicated DC-DC converters. The other subsystems and all the experiments are connected through switches to the Primary Bus or directly to the Secondary Bus, which is stabilized. The Wake-Up System provides for the exact and safe start of the Philae operations. It monitors if the temperature and the available power are in the expected ranges before switching power to start the system.

Rosetta is the first deep-space mission that will go beyond the main asteroid belt relying only on solar cells for power generation. During the comet in-situ investigations, the solar generator of Philae, made up by 2 m^2 of solar arrays, will produce about 8 W of peak power (Fig. 1). To optimize the in-situ power production, the solar generator was built using new Low-Intensity Low-Temperature (LILT) solar cells that are able to produce energy in the very tough environment in which Rosetta and Philae will survive and work [2]. In total, there are 1224 silicon solar cells with dimension 32.4 mm x 33.7 mm, 200 μm thick (Fig. 1).

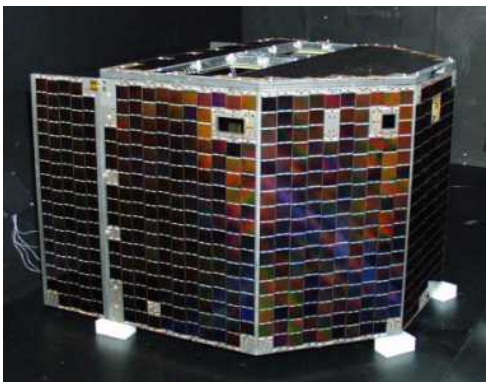


Fig. 1 Philae's solar generator.

Excluding the bottom, all but one side (i.e., the Lid) of Philae are covered with solar cells to maximize the produced power. With reference to Fig. 2, the solar generator is made up by 8 panels: Wall 1 to Wall 5, Balcony 1 and 2, and the Lid. The two balcony panels are connected to Wall 1 and Wall 5, respectively. Thus, the solar generator is made up by *six* separate electrical sections, or Solar Arrays (SA). It can be seen that each array is orientated towards another direction, to enable the Lander to generate power whenever possible without solar arrays movements [4].

The six solar panels are connected to five individual Maximum Power Point Trackers (MPPT), with SA1 and SA5 connected to the same MPPT as they are opposite and never exposed to the Sun simultaneously.

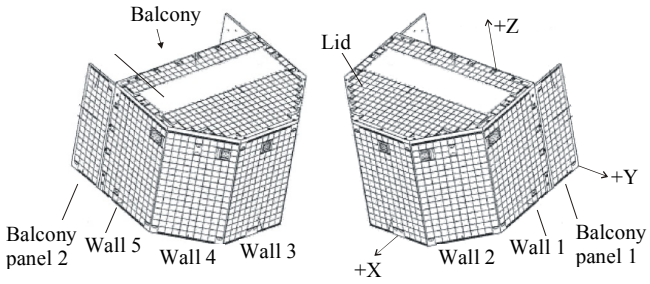


Fig. 2 Philae's solar arrays.

Politecnico di Milano is in charge of developing models and software tools to estimate the produced power, as well as of developing ad hoc hardware simulators to operate the Philae's Ground Reference Module (GRM). These are described in the remainder.

3 Predicting Solar Arrays Performances

During its nominal lifetime, Philae will experience large variations of operating conditions, which strongly affects the performances of solar arrays. As an example, the Sun intensity will vary from 0.11 SC (at 3 AU) to 0.69 SC (at 1.2 AU) in case the extended mission will be confirmed. The characteristic temperatures will vary accordingly, and the accumulated radiation dose will increase as the comet approaches the perihelion. In order to correctly simulate the power produced by solar arrays in these conditions, a software simulator has been implemented [5]. This simulator computes the I-V curves of each of the six SA as a function of

- the Sun aspect angles (azimuth and elevation)
- the Sun distance
- the solar arrays temperature
- the fluence (10^{14} MeV equivalent rad. dose).

The characteristic I-V curve is described by four main parameters (see Fig. 3):

- the short-circuit current (Isc)
- the open-circuit voltage (Voc)
- the maximum power point current (Imp)
- the maximum power point voltage (Vmp).

Given the shape of the I-V curve, it is possible to calculate the current as a function of the voltage (i.e., I=I(V)), or vice-versa, once the four parameters are computed as a function of the environmental conditions. This is done by using a simple model.

A screenshot of the SW Solar Array Simulator is reported in Fig. 4. In the first two columns the six I-V curves are reported; the third column reports the power profile, the Sun aspect angles, and the Lander geometry with respect to the Sun. These are updated at each simulation step.

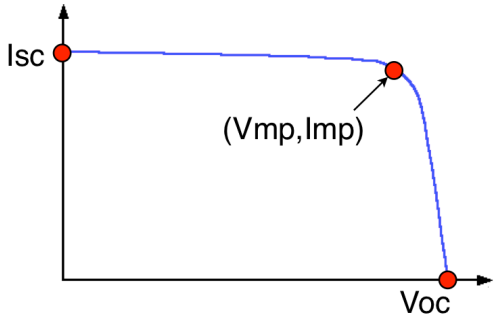


Fig. 3 I-V characteristic curve.

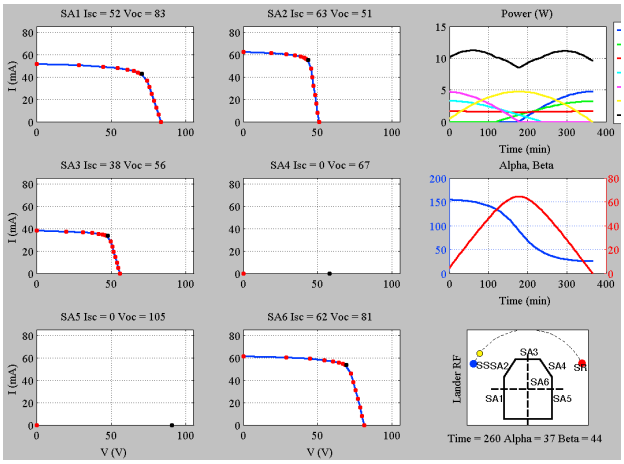


Fig. 4 Philae's solar array simulator software.

4 Simulating Solar Arrays Performances

Testing on-comet operations is not trivial from the power production point of view. In this case, the on-comet I-V curves profile produced by the solar generator has to be mimicked using on-ground equipment. It is known that solar cells currents scale with the cosine of the Sun incidence angles. Thus, the low solar intensity coupled with high values of Sun incidence angles produce very low currents (below 10 mA) that have to be simulated. The variations of SA temperature during the comet day produce changes of the I-V curve, too (for increasing temperature, Voc increases and Isc decreases). The effect of temperature variations on SA1 is shown in Fig. 5, whereas Fig. 6 reports the variations due to varying Sun distance on the same SA [6].

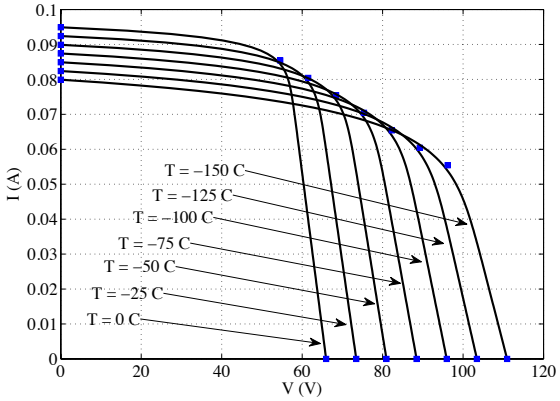


Fig. 5 – I-V curves of SA1 for varying temperatures (Sun perpendicular to SA1, 0.11 Solar Constants).

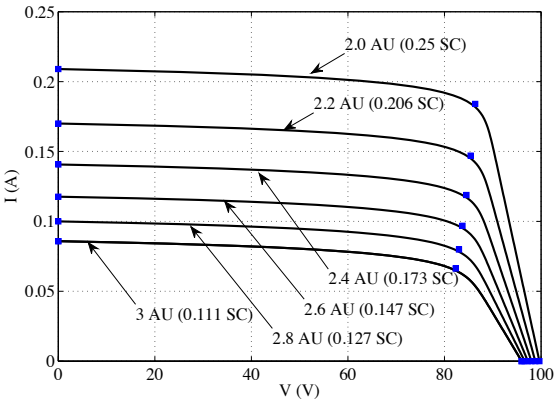


Fig. 6 I-V curves of SA1 for varying solar constants (Sun perpendicular to SA1, T = -100 °C).

Other problems arise due to the presence of the MPPT. In order to track the maximum power point on the solar panels IV-curve, Philae's five MPPT move around the knee-point with a tracking frequency of approximately 30 – 70 Hz. This tracking frequency must not couple with the frequency at which the on-ground equipment provides power to the MPPT (i.e., the frequency at which the current is given for varying voltage).

Previous experiments [7] have shown that using existing, off-the-shelf components to simulate the SA behaviour causes two major problems:

- it is not possible to simulate currents below 100 mA due to hardware limitations of the simulator;
- operating the simulator and the MPPT simultaneously causes instabilities due to the coupling between the MPPT tracking frequency and the frequency at which the I-V curve is represented.

5 The Solar Array Simulator System

In order to both simulate low currents and allow the operation with the GRM, a dedicated Solar Array Simulator (SAS) has been designed and developed. The SAS requirements have been agreed between Politecnico di Milano and the Lander Control Centre (LCC) at Deutsche zentrum für Luft- und Raumfahrt (DLR). The operational architecture has been conceived by Politecnico di Milano (PoliMi), whereas the SAS has been designed and produced by CBL Electronics Srl supported by PoliMi.

5.1 Minimal Requirements

The minimal hardware requirements of the SAS have been agreed by DLR and PoliMi [9]. The main hardware requirements were:

- six independent channels
- voltage in the range 30 – 120 V
- current in the range 0 – 200 mA
- fully analog system
- bandwidth higher than MPPT tracking frequency

The six channels were asked to simulate the six SA of Philae independently; the current and voltage ranges were assessed through parametric analyses [6]; an analog simulator was required in order to stabilize the operations with the GRM.

The main software requirements were:

- static/dynamic simulation scenarios
- I-V curves imported with text files
- log files with 10 ms minimum sampling rate.

The first requirement is related to the possibility of simulating frozen I-V curves as well as time-varying curves (i.e., to mimic the SA behaviour during a comet day); the I-V curve data had to be imported with text file in order to use the SAS SW (Fig. 4); the requirement on the log file was wanted to sample the output I-V data at a frequency higher than that of the MPPT.

5.2 Operational Architecture

The whole simulation system's architecture is sketched in Fig. 7. Beside the SW SAS, two HW SAS have been developed: Fast Loop SAS (FL-SAS) and Diode SAS (D-SAS). These two modules are complementary and can be used independently. For reasons that will be explained later, the FL-SAS is the nominal simulator and the D-SAS is used as backup.

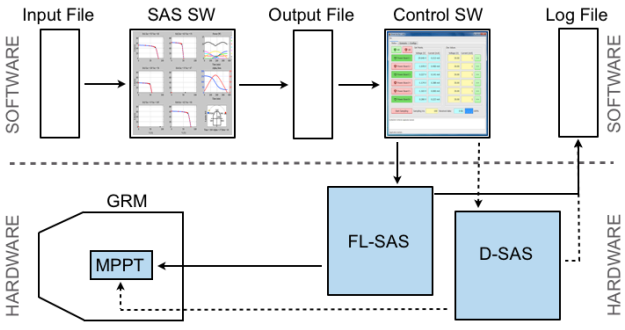


Fig. 7 – Sketch of the operational architecture.

Input file. This file contains the values of the environmental parameters (Sun aspect angles, Sun distance, SA temperatures, fluence) as a function of time.
SAS SW. This software calculates the I-V curves data as a function of time. For each time label, the curves are discretized into 15 (I,V) coordinate points and stored [11].
Output file. This file contains the data on the discretized I-V curves as well as summary information on the simulation (e.g., the estimation on the power produced by each SA).
Control SW. This software imports the data to set the two HW SAS, and it is also used to operate them.
Log File. This files contains the information on the simulation; i.e., the sampled values of the output current and voltage. The sampling rate is specified in the control SW.
FL-SAS. This is the HW Fast Loop SAS. It reproduces the I-V curve as a union of 14 straight lines with varying slope.
D-SAS. The Diode SAS reproduces the I-V curve using a chain of diodes. There is no control on the shape of the curve.
GRM. This is the Philae's GRM. It is equipped with all Philae's subsystems and payloads but the solar arrays. The SAS is connected to the GRM's **MPPT**.
Figure 8 shows the Philae's GRM operated with the D-SAS.

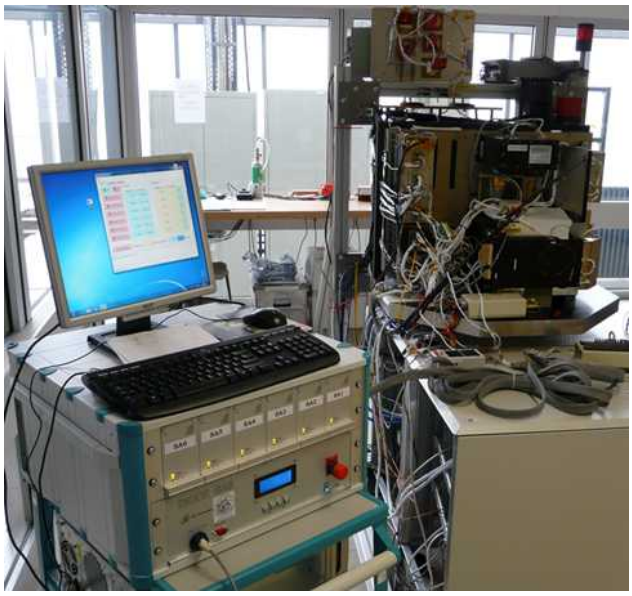


Fig. 8 D-SAS (bottom left) and GRM (top right).

5.3 The Fast Loop SAS

The FL-SAS is made up by six independent electrical boards, each one representing the I-V curve of the SA of Philae (Fig. 9). The boards consist in a number of programmable potentiometers used to discretized the I-V curve into a set of 14 lines with varying slope (15 discretization points with non-uniform distribution are used). The potentiometers are summed up so to construct the whole I-V curve.

The flexibility with which the I-V curve is constructed allows us to reproduce cases with varying maximum power point, beside the mere variation of Voc and Isc. A sample, low current curve with this feature is shown in Fig. 10 [12].

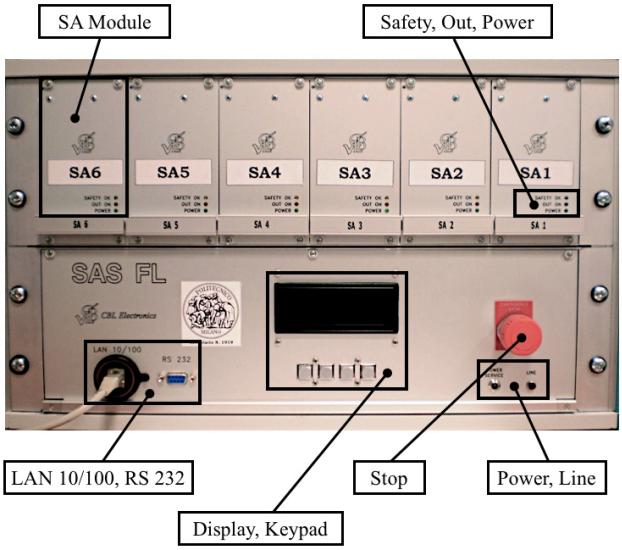


Fig. 9 Overview of FL-SAS.

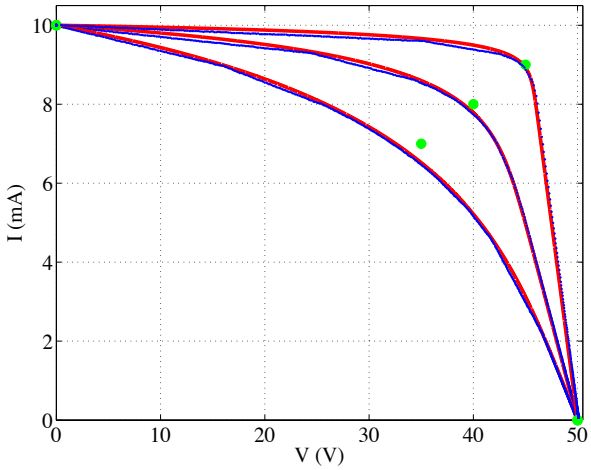


Fig. 10 Three I-V curves (Isc = 10 mA, Voc = 50 V) with varying MPP produced by the FL-SAS' SA2; red: theoretical curve, blue: real curve.

5.4 The Diode SAS

The D-SAS is a device with six independent, electrical channels, too. The I-V curves in this case are represented by a chain of diodes connected in series. The Isc is set by controlling the diodes current, whereas the Voc is achieved by varying the number of diodes. It is straightforward that the D-SAS suffers from several limitations:

- the setting on Voc is poor and depends on the diode forward voltage drop (about 0.7 V);
- the diode's I-V characteristic varies with diodes temperature;
- the shape of the I-V curve depends on the shape of the I-V characteristic of the diodes;
- there is no control on the maximum power point.

For the reasons above the D-SAS has been chosen as a back-up simulator, although it guarantees stable operation when connected to the MPPT. Figure 11 shows an I-V curve obtained with the D-SAS.

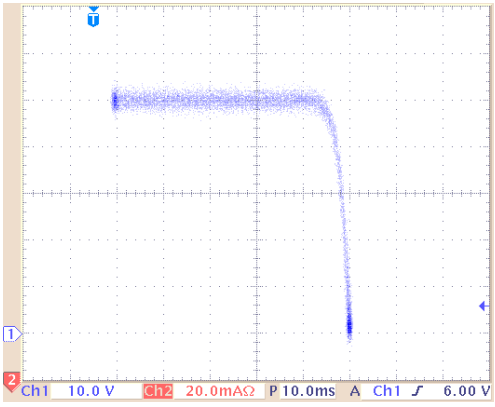


Fig. 11 An I-V curve (Isc = 100 mA, Voc = 50 V) achieved with the D-SAS.

6 Operation of FL-SAS with MPPT

The SAS system described in this paper has been mainly developed to overcome the instability arose when another SAS was used to feed the GRM [7]. Thus, the proper working of FL-SAS connected to the MPPT is crucial. To assess this, a dedicated test has been performed.

The experiment, shown in Fig. 12, consists in connecting the SAS to the MPPT board, whose output is represented by the bus, kept at a fixed voltage by the power sink. A 660 µF capacitor is introduced to simulate Philae's bus capacitors.

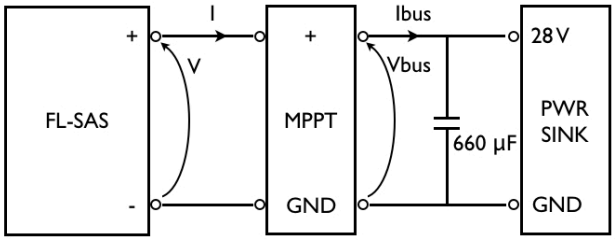


Fig. 12 System used to tests the FL-SAS with MPPT.

The FL-SAS operates at variable current and voltage (I,V); these values are logged by the FL-SAS built-in logging system. This allows us to report the (I,V) points at which the FL-SAS operates and the MPPT is fed. The output of the MPPT board (Ibus, Vbus) is seen on the scope, and eventually recorded.

In Fig. 13 a sample test case is reported. In this case the bus voltage is set to 28 V, whereas the I-V curve is specified by Isc = 80 mA, Voc = 80 V, Imp = 70 mA, Vmp = 70 V. The red line is the I-V curve; the blue points are the (I, V) values sampled by the FL-SAS. These represent the output of the FL-SAS and therefore the values at which the MPPT is fed.

A smooth, stable operation has been observed, so indicating the effectiveness of the FL-SAS when connected to the MPPT. It can be seen that the MPPT works in the region about the "knee" of the curve, which is the region where the maximum power is extracted from the SAS. As the blue points move away from the MPP (green point), the efficiency of the MPPT decreases. A rough estimation of the efficiency can be achieved by measuring Ibus (an 80% efficiency has been observed in the test).

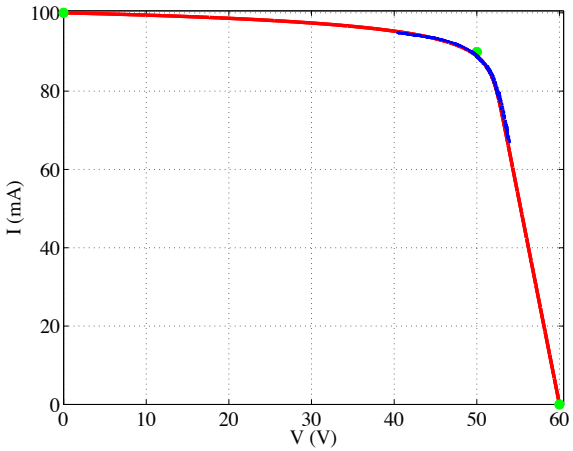


Fig. 13 FL-SAS' SA3 with MPPT (Isc = 100 mA, Voc = 60 V); red: theoretical curve, blue: real (I,V) points.

7 Scenario Simulation

An on-comet scenario has been tested. From the FL-SAS' point of view, simulating a scenario means reproducing given I-V curves as a function of time. A few-minute scenario, which is representative of a typical six hours comet day, has been prepared and loaded on the FL-SAS. Simulating a longer scenario does not add any other qualitative feature that can be obtained from the short scenario. The scenario has been sampled using the built-in sampling feature with 10 ms of sampling rate. From the settings it is possible to reconstruct the theoretical I-V curves, whereas from the sampled (I,V) points it is possible to check the values at which the MPPT works.

Figure 14 reports the theoretical and the sampled (I,V) points of the scenario simulation for SA5. It can be seen that the MPPT works smoothly in the region about the MPP, whose width depends on the current. For low currents (e.g., currents below 10 mA) the MPPT stops tracking and works in "pass-through" mode. It can be concluded that the FL-SAS operation in conjunction with the MPPT has been successful.

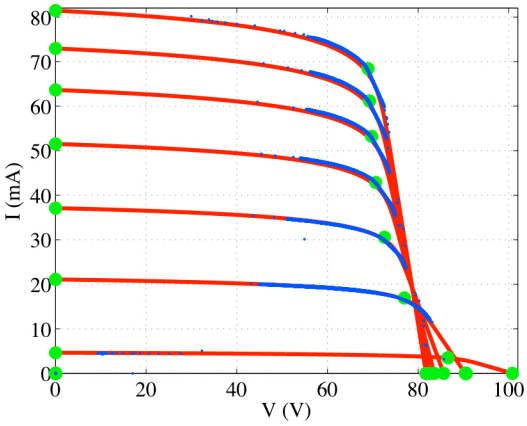


Fig. 14 Theoretical (red) and real (blue) I-V points of SA5 recorded during the scenario simulation.

Both the FL-SAS and the D-SAS performed nominally when connected to the GRM through its built-in MPPT (Fig. 15). In particular, a stable Lander operation has been observed by analysing the telemetry data. This allows us to use the two SAS to execute and assess GRM operations before commanding Philae for its real, on-comet operation.

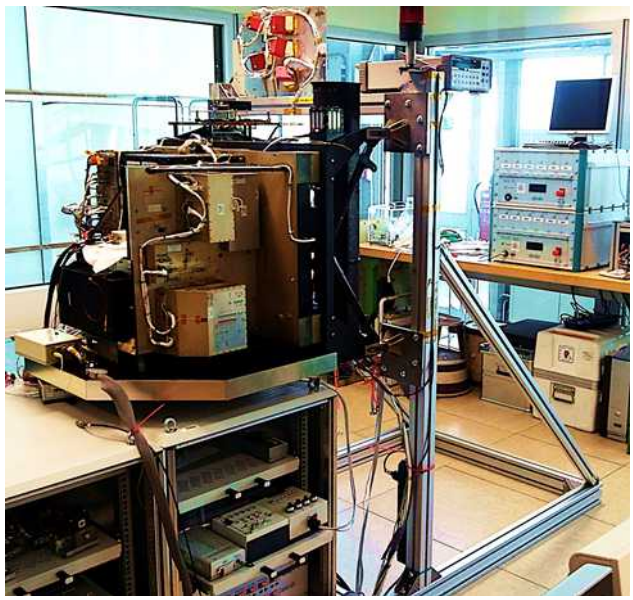


Fig. 15 FL-SAS and D-SAS (top-right) connected to the GRM (centre-left) at LCC.

8 Conclusions

In this note we have described the Solar Array Simulator tools that have been developed to work in conjunction with the Philae's Ground Reference Module. The simulator is made up by software and hardware pieces. The software SAS is used to predict the produced power and to generate the theoretical I-V curves of the six SA once some environmental parameters are provided. Another task of the SW SAS is to control and set the HW SAS. The two HW SAS, i.e., the FL-SAS and the D-SAS, are conceived to replace the Philae's six SA in terms of electrical behaviour. Dedicated tests have proved that low currents can be simulated, and that the two SAS operate in stable conditions when connected to the MPPT. This is a great departure from existing, off-the-shelf Solar Array Simulators. The case of the Rosetta lander Philae is discussed, although the same concepts apply in similar context.

Acknowledgments

The authors would like to acknowledge the Italian Space Agency (ASI) that supported the activities described in this paper through the contract I/062/08/0. The authors are also grateful to the colleagues at LCC that supported this project with useful feedback and valuable suggestions.

References

[1] Ercoli-Finzi A., Bernelli-Zazzera F., Dainese C., Malnati F., Magnani P.G., Re E., Bologna P., Espinasse S., and Olivieri A., "SD2 - How To Sample A Comet", *Space Science Reviews*, Vol. 128, 2007, pp. 281-299.

[2] D'Accolti G., Beltrame G., Ferrando E., Brambilla L., Contini R., Vallini L., Mugnuolo R., Signorini C., Fiebrich H., and Caon A., "The Solar Array Photovoltaic Assembly for the Rosetta Orbiter and Lander Spacecraft's", *6th European Space Power Conference*, Porto, Portugal, 6-11 May, 2002.

[3] Pesco M. and Ercoli-Finzi A., "Rosetta Space Mission: The Solar Array Photovoltaic Assembly", *XIX Congresso Nazionale AIDAA*, Forlì, Italy, 17-21 September, 2007.

[4] VV. AA., "Lander User Manual, 2.2-9b Solar Arrays", RO-LAN-UM-3100-SA, Deutsche zentrum für Luft- und Raumfahrt, 2007.

[5] Topputo F., Bernelli-Zazzera F., and Ercoli-Finzi A., "On-Comet Power Production: the Case of Rosetta Lander Philae", EPSC2010-96, *European Planetary Science Congress 2011*, Rome, Italy, 19-24 September, 2010.

[6] Topputo F., "Sample I-V Curves to Reproduce with the New Solar Array Simulator", PHILAE-SA-SAS-005, Politecnico di Milano, 2009.

[7] Geurts K., "Rosetta Lander Solar Array Simulator & MPPT Operation Instability", RO-LAN-TN-3224, Deutsche zentrum für Luft- und Raumfahrt, 2009.

[8] Geurts K., "Rosetta Lander MPPT Operation", RO-LAN-TN-3225, Deutsche zentrum für Luft- und Raumfahrt, 2009.

[9] Topputo, F., "New Philae Solar Array Simulator: Technical Requirements", PHILAE-SA-SAS-004, Politecnico di Milano, 2009.

[10] Geurts K., "Rosetta Lander Solar Array Simulator Requirements", RO-LAN-TN-3304, Deutsche zentrum für Luft- und Raumfahrt, 2009.

[11] Topputo F., "Philae Solar array Simulator 1.0 User's Manual", PHILAE-SA-SAS-006, Politecnico di Milano, 2010.

[12] Topputo F., "Acceptance and Test of CBL Fast Loop Solar Array Simulator", PHILAE-SA-SAS-007, Politecnico di Milano, 2011.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Thermal modelling and testing of the baffles of SIMBIO-SYS scientific suite of BepiColombo

E.Friso[1], **C.Bettanini**[1], **F.Cucciarrè**[1], **P.Ramous**[1], **M.Zaccariotto**[1], **S.Debei**[1]
[1]*CISAS "G.Colombo", Università di Padova, Italy*

I.Ficai Veltroni[2], **M.Rossi**[2]
[2]*Selex Galileo S.p.A., Italy*

Keywords: *thermal analysis, baffle, Stavroudis, BepiColombo*

Abstract

BepiColombo is a European Space Agency mission to Mercury which will investigate the interior, surface and environment of the planet. SIMBIO-SYS (Spectrometer and Imagers for Mpo Bepicolombo Integrated observatory SYstem) is an integrated scientific suite which has been selected to perform imaging and spectroscopic investigation of the surface of Mercury. This work presents the thermal analysis carried out to support the design of the baffling system of the payload which is exposed to the harsh environment of Mercury. Detailed thermal mathematical model are presented and the results of the mathematical model show the thermal behaviour in orbit and demonstrate the feasibility of the design solution. Thermal vacuum test campaign have been carried out in order to validate the mathematical model and the experimental data shown an overall good correlation.

1 Introduction

BepiColombo will be the third mission to Mercury, the least explored planet in the inner Solar System, following the NASA Mariners 10 and MESSENGER. It is an interdisciplinary mission carried out jointly by the European Space Agency (ESA) and the Japanese Aerospace Exploration Agency (JAXA), comprising two different orbiters which will be studying the planet and its environment. The Mercury Planetary Orbiter (MPO) is led by ESA and will focus on a global characterization of Mercury through the investigation of its interior, surface, exosphere and magnetosphere [1].

An integrated suite of instruments, Spectrometer and Imagers for MPO BepiColombo Integrated Observatory SYStem (SIMBIO-SYS), which includes a stereo imaging system (STC). a high-resolution imager (HRIC) and a visible-near-infrared imaging spectrometer (VIHI) has been selected for the BepiColombo mission. SIMBIO-SYS will scan the hermean surface with the three channels and map the physical, morphological, tectonic and compositional properties of the entire planet [2]. One of the main technical challenges of SIMBIO-SYS is to withstand the harsh environment at the planet. In particular the intense thermal environment generates significant difficulties in controlling the payload temperature and specific effort must be putted in the thermal design of those elements directly exposed to Sun and Mercury such as the telescope baffles.

2 Mercury thermal environment

The orbit selected by ESA is mainly a compromise between science objectives and thermal load on the spacecraft. The MPO spacecraft will fly on a nadir pointing polar orbit at 400 x 1508 km with a 2.3h period with its apoherm on the equator on the Sun side when Mercury is in its perihelion. Half a Mercury year later, at aphelion, the sub-solar point occurs when the spacecraft is at its minimum distance to the planet [1].

Since the orbit of Mercury is the closer to the Sun with the higher eccentricity of the solar system planets, the direct solar irradiance varies from 6290 W/m^2 when Mercury is in its aphelion up to 14490 W/m^2 at perihelion. Furthermore the temperature distribution of the planetary surface varies from a minimum of 100K in the shadowed side up to about 690K at sub-solar point at perihelion.

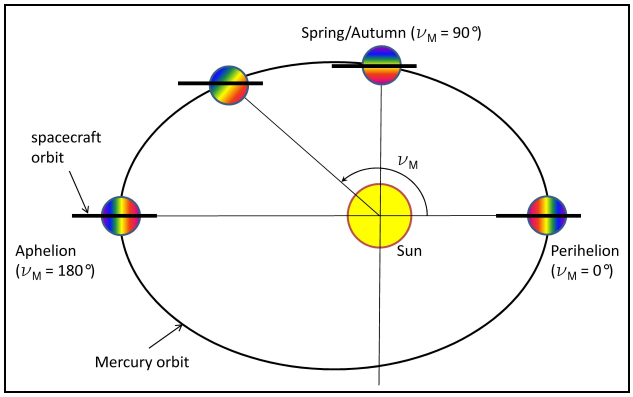


Fig.1: Orbit of the BepiColombo MPO spacecraft for different true anomaly (ν_M) of Mercury

At the early stage of the analysis, a dedicated mathematical model has been developed and simulations has been carried out in order to evaluate the solar, albedo and planetary infrared components of incident fluxes on the front end of the payload during the orbiting of MPO spacecraft around Mercury at different values of planet true anomaly (for the details of the mathematical model see [4]). The following figures 2, 3, 4 show the solar, albedo and the planetary infrared heat fluxes incident on a nadir pointing unitary surface as a function of Mercury and spacecraft true anomalies.

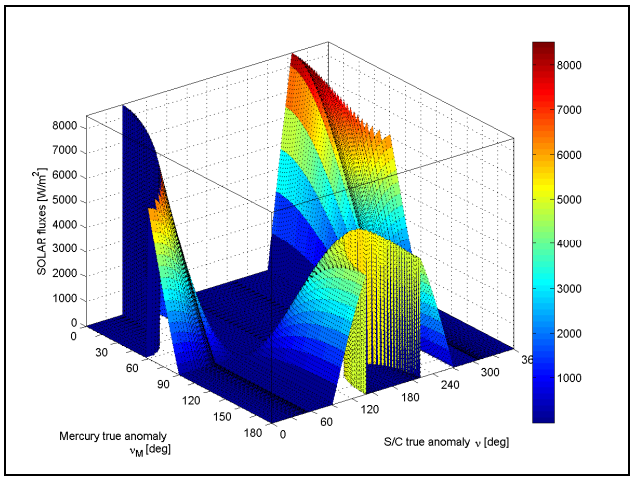


Fig.2: The solar fluxes incident on a surface with normal pointing toward nadir as a function of the true anomaly of the S/C orbit and of the Mercury orbit. The hollows of the surface corresponds to eclipses occurrence.

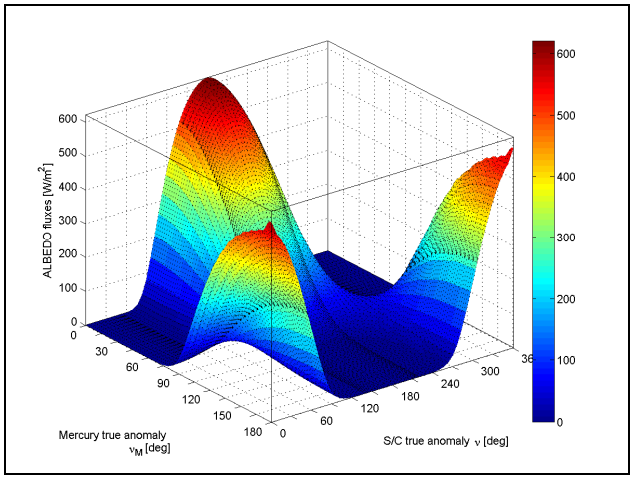


Fig.3: The albedo fluxes incident on a surface with normal pointing toward nadir as a function of the true anomaly of the S/C orbit and of the Mercury orbit.

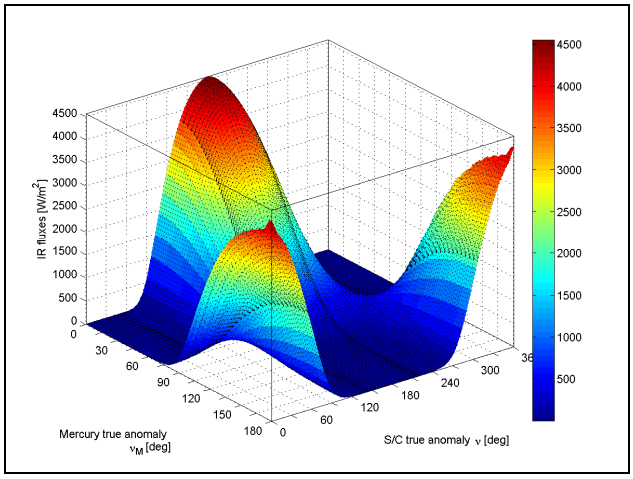


Fig.4: The infrared planetary fluxes incident on a surface with normal pointing toward nadir as a function of the true anomaly of the S/C orbit and of the Mercury orbit.

It could be recognized that the thermal environment is critical not only in terms of amount of thermal input coming from space but also because of the variation of the different contribution of thermal input. In order to sign out the overall most critical radiative cases to be analyzed for the design of the baffle, the total energy incident on the unit surface during one spacecraft orbit are defined in Eq. (1), (2), (3) as a concise parameter to evaluate the criticality of the specific orbit.

$$E_{IR} = \int_0^T q_{IR}(t)\, dt \qquad [\mathrm{J/m^2}] \qquad (1)$$

$$E_S = \int_0^T q_{SUN}(t)\, dt \qquad [\mathrm{J/m^2}] \qquad (2)$$

$$E_A = \int_0^T q_{ALBEDO}(t)\, dt \qquad [\mathrm{J/m^2}] \qquad (3)$$

$$E_{env} = E_S + E_A + E_{IR} \qquad [\mathrm{J/m^2}] \qquad (4)$$

The parameters E_{IR}, E_S, E_A represents the total incident infrared, solar and albedo energies respectively and T is the spacecraft orbital period. The variable E_{env} represents the overall thermal input coming from space. Figure 5 shows this total energies as a function of Mercury true anomaly. The peak values of the total heat loads occur at perihelion (ν_M = 0°) and at Mercury true anomaly ν_M = 141.8°, which will be considered as hot cases. At perihelion the average contribution of infrared fluxes from the planet is much stronger than the solar component of incoming heat flux since, even though the solar constant has its maximum in this season of Mercury, the duration of sunlit phase is shorter. The solar contribution of incoming heat flux has its maximum at around ν_M = 141.8°, which corresponds to the beginning of seasons which presents apoherm eclipse. In fact, in this season of Mercury, the nadir pointing surface will never be eclipsed by the planet and also the angle between the surface normal and the Sun direction will reach its minimum. The minimum of heat loads coming from space occurs when Mercury is at true anomaly ν_M = 90° (hermean Spring/Autumn). Another significant analysis case will be at Aphelion where the instrument is foreseen to be operative and the thermal input coming from space is still very high.

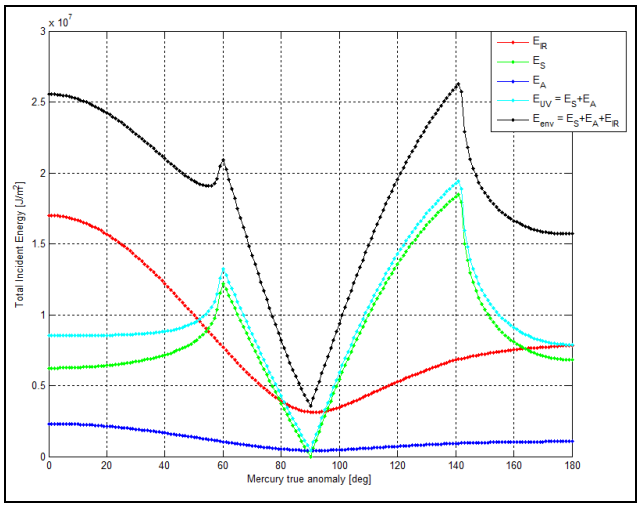


Fig.5: the curves graphs the total incident energy (per unit area) on a nadir pointing surface during one orbit of the S/C as a function of Mercury seasons.

This study and the simulations carried out allowed to identify the most critical cases for the thermal design of the payload and the baffles, providing a detailed knowledge of the thermal environment that the payload will face during BepiColombo mission.

3 Thermal design and modeling of SIMBIO-SYS baffles

The SIMBIO-SYS scientific suite design is depicted in Fig. 6. It is composed by three optical channel: the high resolution camera (HRIC) with its reflective Stavroudis baffle and its proximity electronics on the top (HRIC PE), the stereo camera (STC) with its two black absorbing baffles (H/L) pointing ±20° forward/backward-along-track respectively and the proximity electronics on top of the honeycomb structure (STC PE) and finally the imaging spectrometer (VIHI) with its black absorbing baffle and its own proximity electronics (VIHI PE).

Due to its larger aperture, the HRIC channel will see the largest amount of heat entering telescope from the environment, then a Stavroudis reflective baffle has been selected by the SIMBIO-SYS team. In contrast to the classic approach of having highly absorbing baffle, a reflective baffles tries to reject, as

much as possible, the power entering the aperture back through the front aperture thus minimizing the absorbed power. A Stavroudis baffle consists of an alternated series of highly reflective ellipsoids (oblate spheroid) and hyperboloids having common foci located at the edge of entrance aperture; this geometry would ideally reject all the stray light entering the aperture back out of the entrance [3]. The STC and VIHI channels will use classical internally black absorbing baffles.

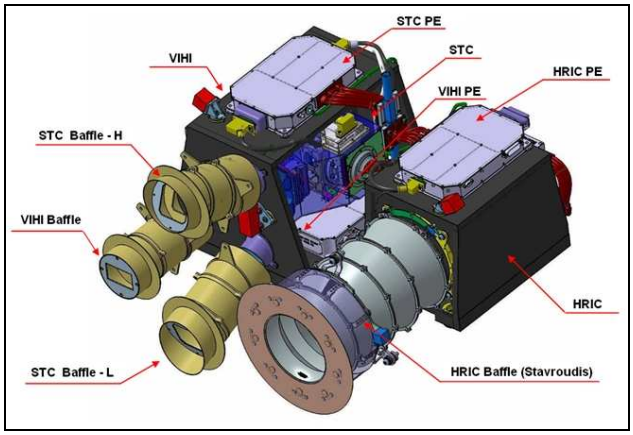


Fig.6: The SIMBIO-SYS scientific suite.

Thermal analyses of baffles have been afterwards carried out using lumped parameter thermal network method and implemented in ESARAD/ESATAN software. This allows to predict the main heat transfer mechanisms and temperature distribution and to estimate the performance of baffles in terms of heat rejection capability. Radiative heat transfer plays an important role in space application and the selected software implements Monte Carlo Ray Tracing algorithms for the computation of radiative exchange factors and absorbed power of thermal nodes in orbit. In particular, a study of the appropriate implementation of the Stavroudis baffle's geometry into thermal analysis software has been performed in order to effectively model this particular type of geometry. The main technical problem with modelling the Stavroudis geometry in ESATAN is that the program does not support the ellipsoids and hyperboloids geometry which therefore must be approximated in an appropriate way. In order to assess how the surfaces has to be approximated, a sensitivity analysis was performed by comparing different approximation with an indicator. The Stavroudis geometry suggests that an appropriate model should minimize the absorbed flux on the internal surface of the frontal ring and the direct solar absorbed fluxes by a disc at the rear side of the baffle [4]. The primitive shapes available to describe the surface of the Stavroudis baffle are basically spherical segments and frustum of cones. The approximation of the generatrix of the surface of revolution has been made with dedicated algorithms such that a more dense meshing occurs where the variation of the surface normal is higher, in order to best fit the curvature of the baffle geometry. It was found out that such a non-equally sized meshing approach, together with the use of spherical segments to approximate the ellipsoids and frustum of cones to approximate hyperboloids, will significantly increase the modelling effectiveness, minimizing the overestimation of the environmental absorbed fluxes and greatly reducing the computational time [4]. The architecture of the thermal model of SIMBIO-SYS baffling sub-system is described in Fig.7.

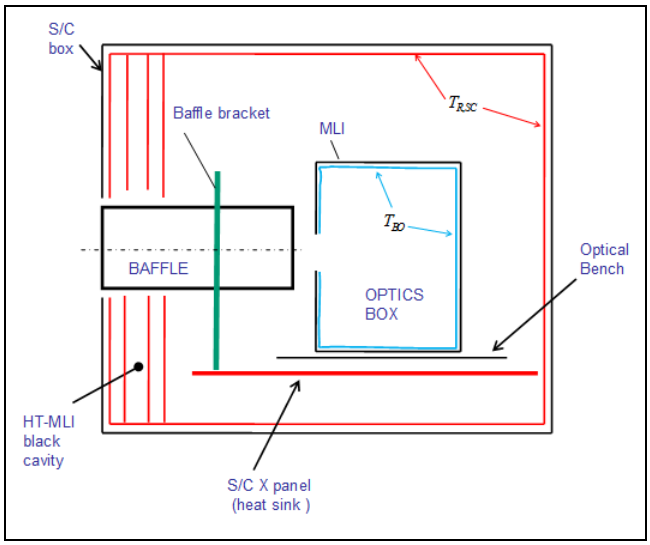


Fig.7: Thermal modelling scheme and interfaces of SIMBIO-SYS baffles.

Geometrically the unit optics STC, VIHI, HIRC has been considered as a box with a hole corresponding to the optical aperture of the instrument. The inner side of the unit optics box has been assumed as a blackbody cavity at boundary temperature, constant over the orbit and representing the average temperature of the components of the optical unit. The value of the temperature T_{BO} has been established on the basis of thermal results available from

preliminary thermal models and a sensitivity analysis shown that the temperature of this boundary condition does not greatly affect the baffles' temperature distribution and exchanged heat power with interfaces. This simplification allows to focusing on the detailed thermal analysis of baffling subsystem without implementing the whole unit thermal model [4].

The radiative interfaces to the spacecraft has been modeled with two different blackbody cavity as foreseen by specification [5] and are graphed in Fig.7. The first cavity represents the radiative interface with the high-temperature MLI at the nadir pointing surface of MPO spacecraft (HT-MLI black cavity). The boundary temperature of this cavity has been defined as follows: when the spacecraft is exposed to solar or planetary radiation, the cavity temperature is equal to +250°C, otherwise the temperature decrease instantaneously to -80°C during eclipse. The radiative interface with the spacecraft internal environment is an internally black envelope at constant temperature $T_{R,SC}$ of about +50°C.

The conductive interface to the spacecraft consists of a honeycomb aluminium panel (baffle bracket) to which the baffles are mounted and it is conductively linked to a high thermal capacity panel of the spacecraft, which is therefore considered as an heat sink (see Fig.7). The SIMBIO-SYS optical units are conductively isolated to the optical bench.

3.1 Thermal design of HRIC baffle

The baffle of the high resolution camera HRIC will be a highly reflective Stavroudis geometry. Since the baffle's inner surface has to be shaped very accurately and smooth, in the range of a few nanometers root mean square surface roughness, the baffle will be manufactured with a single point diamond turned AA6061-RSP aluminium alloy. The front of the baffle, which is penetrating the spacecraft high temperature MLI, is isolated laterally by two concentric cylindrical titanium foil of 0.1mm thickness. These two cylinders are conductively isolated to the baffle and will be coated with a low-emissivity gold surface, which allows to minimize the heat leakage from the spacecraft to the baffle and to reduce the axial temperature gradient of the baffle that would affect the thermo-optical performance of the reflective Stavroudis geometry.

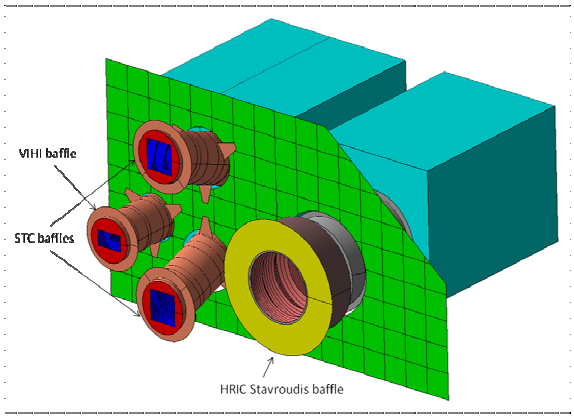


Fig.8: Thermal model of SIMBIO-SYS baffle and the support panel.

A frontal ring ensures that no light penetrates the aperture outside the foci of the first ellipsoidal surface. The external diameter is slightly bigger than the baffle radius in order to avoid trapping light between the baffle and the spacecraft MLI. The ring consist of a 0.4mm tick titanium alloy which is coated with ZrO_2 ceramics on the planet facing side and is conductively decoupled from the baffle, providing a thermal shield for the main body of the baffle. The baffle will be conductively decoupled to the baffle bracket with low conductance titanium joints.

	Solar absorptivity α	Infrared emissivity ε
Diamond turned AA6061 (Stavroudis internal surface)	0.15	0.02
ZrO2 ceramics (frontal ring)	0.25	0.76
Gold coating	0.19	0.03

Table 1: Thermo-optical properties of HRIC baffle.

3.2 Thermal design of STC and VIHI baffles

The stereo camera STC and the imaging spectrometer VIHI will use a classical highly absorbing baffle approach and will be internally black coated (Electrodag502) in order to protect

the telescope from the unwanted light coming from space. The external surface of the frontal vane of the baffles will be highly polished to partially reject the heat flux incoming from space. The mounting screws will conductively link the baffles to the support honeycomb panel of the spacecraft.

	Solar absorptivity α	Infrared emissivity ε
Electrodag502 (internal surface of baffles)	0.90	0.85
Polished (frontal vane)	0.12	0.05

Table 2: Thermo-optical properties of STC and VIHI baffles.

The baffles of SIMBIO-SYS will be made of aluminium alloy which, due to the high value of specific thermal capacitance, minimize the temperature variation of the baffle over the orbit and reduce the temperature gradients thanks to the high thermal conductivity which ensures the necessary geometric stability for the optical performance.

3.3 Results of simulations

The thermal model has been implemented to perform simulations for the most critical orbits singled out after the preliminary assessment of BepiColombo thermal environment and considering different values of boundary conditions at spacecraft interfaces. Here are presented some of the most significant results of simulations; a more detailed and complete description of simulation and data results can be found in [6]. Fig. 9, Fig.10 and Fig. 11 show the temperature of the VIHI, STC and HRIC baffle respectively over the BepiColombo orbit when Mercury is at its perihelion. This case has been found to be the most critical design case for the baffle thermal design. The two peaks in the plots corresponds to the sudden solar illumination while the maximum in the middle of the curve corresponds to the infrared flux coming from the sub-solar point of the planet. The temperature minima occurs at the orbit part when the instrument is in eclipse.

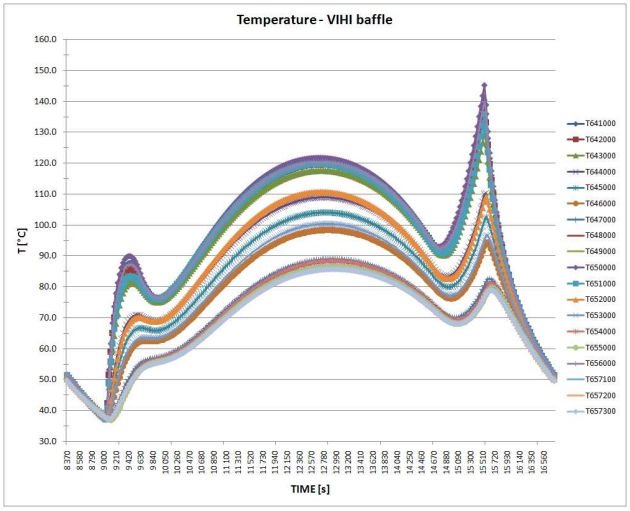


Fig.9: Temperature distribution of VIHI baffle's nodes during perihelion orbit

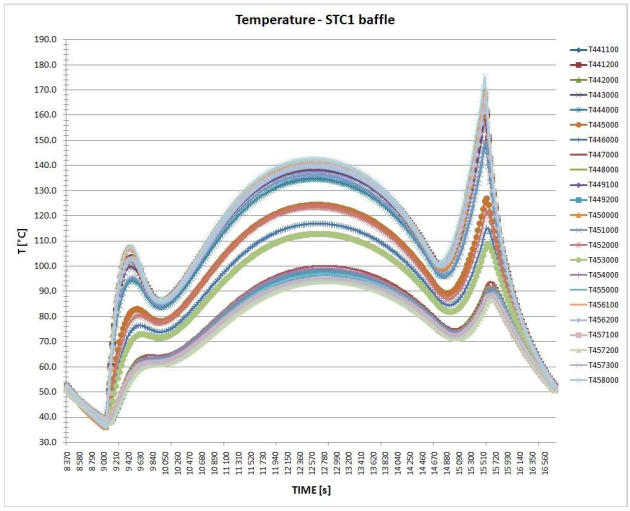


Fig.10: Temperature distribution of STC-H baffle's nodes during perihelion orbit

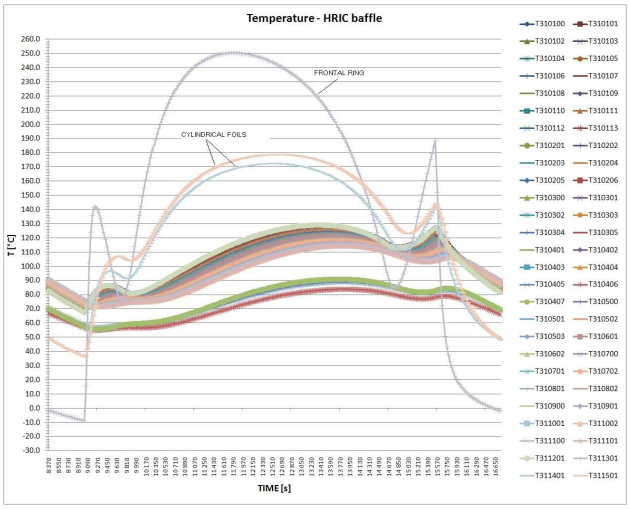


Fig.11: Temperature distribution of HRIC baffle's nodes during perihelion orbit

Heat balance calculation has been made to estimate the heat rejection efficiency of the

baffles, which has been defined as the ratio between the heat rejected directly to space (by reflection or by re-emission of absorbed heat) and the total incident heat flux from the environment. Table 3 reports the values of heat rejection efficiency for perihelion orbit.

	VIHI	STC	HRIC
Total environmental heat flux (W)	13.8	36.6	106.2
Heat rejected by the baffle (W)	10.9	26.9	100.7
Heat rejection efficiency (%)	79.0%	73.5%	94.8%

Table 3: Heat power rejection of SIMBIO-SYS baffles at Mercury Perihelion (average orbital values).

The SIMBIO-SYS baffles demonstrate an overall good performance in term of heat rejection capability and guarantee that the heat loads introduced by the unit to the spacecraft interfaces will be within the limit established by the spacecraft thermal control system [5].

4 Thermal testing of SIMBIO-SYS baffles and thermal model validation

Thermal verification tests and subsequent correlation to analytical models has been carried out in order to reduce the uncertainty of numerical prediction. Due to the extreme thermal environment of Mercury, test requirements should be tailored to the specific BepiColombo program and an experimental set-up has been developed at CISAS "G.Colombo" research center of the University of Padova to simulate the orbital fluxes and the spacecraft thermal interfaces. At present the test-bed consist of a thermal vacuum chamber (air pressure in the 10^{-6} mbar range) with heating and cooling controlled heat sources. Moreover, a solar simulator is under integration at CISAS premises which consist of an optical system powered by a 10kW Xenon arc lamp that will collect a uniform light beam into the chamber and will irradiate up to around 14kW/m^2, which corresponds to the solar constant at Mercury perihelion.

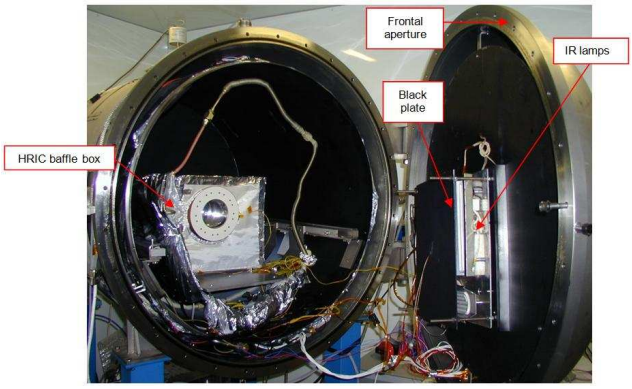


Fig.12: SIMBIO-SYS baffles thermal vacuum testing set-up (here the HRIC baffle)

Thermal Vacuum tests campaign on the structural thermal models (STM) of the SIMBIO-SYS baffles have been carried out, separately for HRIC baffle and STC-VIHI baffles. The baffles have been placed inside an aluminium box covered by high temperature MLI, frontally exposed to the heat source (hot or cold plate). To simulate the thermal interfaces with the spacecraft, the baffles are mounted on an aluminium plate at controlled temperature, which represents the baffle bracket. A cryostat has been maintained at about -180°C and conductively connected to the baffle bracket plate and to the aluminium box by means of copper braids in order to allow the temperature control of these interfaces [7], [8].

Two main steady-state conditions have been simulated during the tests: a hot case condition, by means of a plate heated up by infrared lamps, and a cold case condition, by means of a plate cooled by liquid Nitrogen. The boundary conditions for the heat sources, which simulate the infrared radiation incoming from the planet, have been set in order to obtain a radiative flux incident on the front end of the baffles equivalent to the corresponding one in orbital conditions. During the hot case tests, the temperature of the plate facing the baffles has been controlled at about 220°C by means of PID digital controller, with a stability of at least 1°C/hour for 2 hour, allowing to obtain an incident flux equivalent to the peak of infrared heat flux coming from Mercury at perihelion orbit. In cold case simulation, in order to reproduce the thermal condition when the baffles look at the shadowed cold side of the

planet, the plate in front of the baffles has been maintained at around -150°C with a stability of at least 1°C/hour for 2 hour, which has been the minimum temperature value obtainable at equilibrium during tests. The temperatures has been monitored by means of four wires PT100 resistance thermometers being placed at significant points of the baffle surfaces and at control points of the experimental set-up. Fig.12 shows the temperature results obtained during the hot case test of the STM model of Stavroudis baffle [7], [8].

A thermal mathematical model of the test-bed has been added to the baffle model in order to compare the experimental data with the numerical results. The thermal testing conditions has been numerically simulated and the temperature values obtained by analysis has been compared with experimental data. The assumptions made on thermal parameters has been validated with test results, in particular the contact conductivity values within the baffles vanes has been tuned to best fit the experimental data. The temperature values obtained by analysis showed an overall good correlation with experimental data and the maximum difference in temperature between the experimental data and the numerical results is within 5°C, according to the requirements.

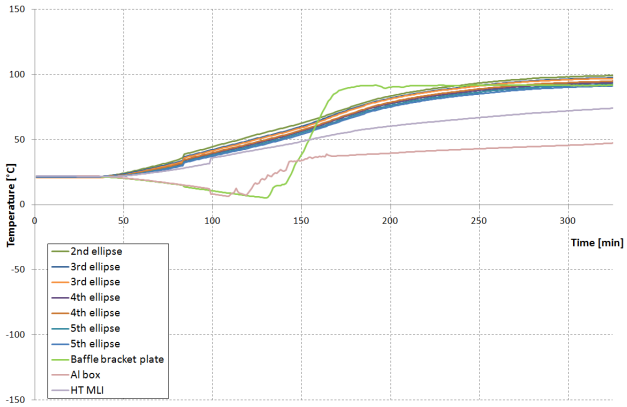


Fig.13: Temperature profile of points of interest of HRIC Stavroudis baffle, internal surface of the HT-MLI and aluminium box during hot case testing

5 Conclusions

Extensive studies have been conducted in order to prove and optimize the thermal design concept of the baffling system for SIMBIO-SYS scientific suite. The preliminary analysis of Mercury thermal environment provided a framework for the thermal design of the payload on board MPO spacecraft allowing to figure out the most critical design cases, supplying the design drivers for the realization of an experimental set-up capable to simulate the main heat loads in orbit. Detailed thermal models of baffles have been afterwards carried out using lumped parameter thermal network method and implemented in thermal analysis software. Special attention has been paid on the implementation of the Stavroudis baffle geometry and appropriate discretization method has been found in order to minimize the overestimation of the absorbed flux and to reduce the computational time. The mathematical model allowed to predict the thermal behavior in orbit and demonstrated the feasibility of the thermal design solution. Thermal balance test on baffles' structural thermal models (STM) and subsequent correlation to thermal analysis had successfully provided physical insight into heat transfer mechanisms and reduced the uncertainty of model prediction.

References

[1] J.Benkhoff et al., BepiColombo Comprehensive exploration of Mercury: mission overview and science goals, *Planetary and Space Science*, Vol.58, 2010, pp. 2-20.

[2] E.Flamini et al., SIMBIO-SYS: the spectrometer and imagers integrated observatory system for the BepiColombo planetary orbiter, *Planetary and Space Science*, Vol.58, 2010, pp. 125-143.

[3] O.N.Stavroudis, L.D.Foo, System of reflective telescope baffles, Optical engineering, Vol.33 (3), 1994, pp.675-680.

[4] E.Friso, "Thermal effects reduction techniques for the SIMBIO-SYS scientific suite of BepiColombo mission", Ph.D. Dissertation, CISAS University of Padova, Italy, 2010.

[5] EID-A BepiColombo, BC-EST-RS-01140, Issue 2, Rev.2, 2010

[6] E.Friso, "Detailed thermal mathematical model of SIMBIO-SYS baffles", Technical Report, CISAS University of Padova, 2011

[7] F.Cucciarrè, "Stavroudis baffle's STM TMM validation", Technical Report, CISAS University of Padova, 2010

[8] F.Cucciarrè, "STC-VIHI baffles' STM TMM validation", Technical Report, CISAS University of Padova, 2010

UML modeling procedure for anticipating cost at system level design

G. Belvedere [1] and P. Gaudenzi [2]
[1]*Ph.D. Student, Dipartimento Ingegneria Meccanica e Aerospaziale – Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italia*
Email: giampiero.belvedere@uniroma1.it
[2]*Full Professor, Dipartimento Ingegneria Meccanica e Aerospaziale – Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italia*
Phone number: +39 0644585304 Fax: +39 0644585670
Email: paolo.gaudenzi@uniroma1.it

Keywords: *UML, complexity, design to cost, fuzzy logic, lunar rover.*

Abstract

The complexity of currently developed systems pushes the need to make use of Model Based System Engineering techniques in order to maintain a holistic view of the problems, being contemporarily focused on the particular aspects of the design. Among the architectural languages, Unified Modeling Language is the most widespread one. The paper addresses the issue of cost and risk estimate describing a procedure that employs UML to formalize the process of the design phase starting from the system requirements down to the product tree description, fulfilling a synergic view of the design process. A novel definition of design to cost is given, based on entropy-based complexity measurement and a fuzzy model of the quality of the estimate. The procedure has been applied to the study of a Lunar Rover mission.

1 Introduction

Model Based System Engineering (MBSE) consists of a set of formal languages and related tools to describe systems requirements, functions, operational context and environment, thus easing the system engineering process by first defining the operational concepts, focusing on the development of an architecture with all its components and mutual relationships, taking into consideration how the system is designed and its life-cycle evolution. Architectures have a strong influence over the life cycle of a system. Today, the increasing complexity of systems drives the need for the practice of system engineering adherence to architectural precepts. Among the architectural languages Unified Modeling Language (UML) is the most widespread one.

The paper is organized as follows: section two will be dedicated to a description of the use of UML as a system engineering tool to model complex systems starting from the requirements; the following subsection will describe how UML has been used to derive system complexity, since complexity is related to cost. In section three a method will be presented, based on a synergic use of UML system representation and a fuzzy model of the human reasoning process, to derive in a quantitative manner the combined cost risk effect intervening in the process of cost estimate. The methodology will be applied to

the preliminary design of the Guidance, Navigation and Control (GNC) subsystem of a lunar rover, as described in section four. Finally the conclusion will be presented.

2 UML Modeling Of Systems

UML is a graphical language widely used to represent both software-based intensive systems and complex processes, whose principal objective in a system description is to provide system architects, software engineers and developers with tools for analysis, design and implementation of software-based systems as well as for modelling business and similar processes [1].

UML analysis is a technique for requirement elicitation. Drusinsky [2] described the process of identifying Natural Language requirements of interest from UML analysis diagrams such as activity diagrams and message sequence diagrams. Emadi and Shams [3] proposed new algorithms that enable a designer to transform use-case and sequence diagrams into executable models based on Petri nets. Meziane, Athanasakis and Ananiadou [4] dealt with the problem of translating software requirements expressed with informal natural language specifications into formal ones. Mindock and Watney [5] used UML to represent behaviors and interfaces using common language and terminology across multiple project elements, throughout different levels of system engineering and among system, software, and controls engineering.

By the use of UML Di Domizio and Gaudenzi [6] formalized the process of concurrent engineering applied to the design of satellites while Belvedere, Filippazzo and Gaudenzi [7] modeled the cost estimate process in a concurrent engineering environment.

2.1 UML model representation of system complexity

UML has been widely used as a mean of deriving software system complexity. The complexity is used since the more complex a system, the more difficult to design, build and use. Exploiting complexity measurement techniques, systems architects might measure complexity earlier in and throughout the engineering lifecycle. In the field of software many complexity measures have been proposed.

Halstead proposed a difficulty measure [8]:

$$D = \frac{L}{V} = \frac{(2+n3)*log_2(2+n3)}{(N1+N2)*log_2(n1+n2)} \qquad (1)$$

Where D is the difficulty, L represents the program level and V is the program volume, n1 is the number of unique operators in a procedure, n2 is the number of unique operands in a procedure, n3 is the number of unique input or output operands in the minimum equivalent procedure; N1 represents the total number of operator occurrences in a procedure, while N2 is the total number of operand occurrences in a procedure.

McCabe [9] proposed a *cyclomatic* complexity measure calculated according to the following relation:

$$v = e - n + 2 * p \qquad (2)$$

where e, n and p are properties of a program control flow graph: e is the number of edges, n is the number of vertices, and p is the number of distinct connected components.

Bertolino, Bonivento, De Angelis and Sangiovanni Vincentelli, starting from models described in UML notation, provided an early estimation of performance related parameters comparing possible alternative implementations in advance [10].

Briand and Wüst [11] empirically investigated the relationship between class size and the development effort for a class, trying to build prediction cost models from size and design measures, and used these models to predict system development effort.

In [12], using effort data from an object-oriented development project, Lavazza and Robiolo investigated the relationship between class size and the development effort for a class, and what additional impact structural properties such as class coupling have on effort.

The complexity of software projects by the help of UML models has been investigated in [13] by Podgorelec and Hericko.

Lankford tried to adapt Halstead difficulty and McCabe cyclomatic complexity, to the measurement of architecture complexity using metrics applied to architecture models based on UML class, sequence and state diagrams by redefining measure factors in terms of UML element [14].

3 A Design to Cost method based on UML, complexity measurement and a fuzzy model

In the present section a methodology will be presented to get an anticipation at system engineering level of the complexity of a system [15]. The methodology is based on the following main points:

- a UML model of the system, starting from mission objectives and requirements, and a calculation of system complexity by means of the information extracted from UML diagrams;
- a fuzzy model of the overall quality of the technical and cost estimate;
- a calculation of the risk of the outcome of the cost estimate process based on the product of the two quantities determined above.

Since Wades and Welch have shown how 70% of the future system costs are anticipated during the conceptual study phase [16], the aim of the proposed method is to point out in a quantitative manner the combined effect of technical and human elements in order to minimize the risk inherent in every estimate starting from the initial design phases.

3.1 UML Model Representation of System Requirements

First step is the translation of the Primary Mission Objectives into a Use Case Diagram. From the main use case derived by the Mission Objectives, a list of dependent use cases (sub use cases) are derived: these ones represent Mission Requirements and Constraints.

From the use case diagrams of Requirements and Constraints, an "*expert judgment*" approach takes out the main functions to be performed by the system. The functions are described by sequence diagrams and the classes represented inside the latter ones are derived by the actors of the use cases. The sequence diagram messages self-directed into one of the participating classes imply a function internally performed by the class itself. From this fact new classes are derived and the represented class is split into its subclasses.

Once a UML model of the system has been laid down, the elements for its complexity calculation are derived from the following characteristics:

- The number of messages in the sequence diagrams exchanged among the participant objects;
- The self-directed messages exchanged by an object, representing its internal functions;
- The attributes and the methods of the classes, derived by the use cases and sequence diagrams analysis.

How these characteristics will be accounted for in order to determine the complexity of the system is described in the next subsection, where concepts of information theory will be used.

3.2 Entropy-based System Complexity Calculation

Information entropy is a concept introduced for the first time by Shannon [17]. For an information source X with alphabet $A = \{a_1, a_2, \dots, a_n\}$ and alphabet probability distribution $P_A = \{p_1, p_2, \dots, p_n\}$, its entropy is defined as:

$$H(X) = -\sum_{i=1}^{n} p_i(x) * log_2(p_i(x)) \quad (3)$$

The more difficult it gets to estimate the outcome at the source output, the larger the entropy becomes.

The entropy concept has been applied as a metric for object oriented design quality. According to Chatzigeorgiou and Stephanides [18] the entropy of the next generation of an object oriented system with enhanced functionality remains close to its previous level in case the added functionality affects a limited number of existing classes; on the other hand a poorly designed system increases entropy drastically.

For software systems UML diagrams are sufficient to represent the complexity of their architecture; the same is not directly applicable to different architecture domains. In the case of space systems both the Technology Readiness Level (TRL) and the design status have a strong influence on system complexity and cost.

If the TRL value increases, the standard deviation about the most likely estimate decreases, thus implying a higher probability of having a low cost variance for the estimate. Values of the Coefficient of Variation (CV), representing the ratio between the standard deviation and the mean value of the estimate, may be found in [19], as a function of the TRL scale.

Equation 3 may be modified introducing the CV, thus obtaining:

$$H = -\sum_{i=1}^{n}(1 - CV_i) * log_2(1 - CV_i) \qquad (4)$$

where n is the number of classes of the system, CV_i is the Coefficient of Variation for the subsystem or element represented by the i-th class. Each element of Eq. (4) represents the baseline factor to calculate the average complexity of the system. The complexity may be divided into intra-complexity and inter-complexity. For the intra-complexity the proposed metrics is as follows [15]:

$$H_{int} = -\sum_{i=1}^{n}\left[NF_i * \frac{1}{SIM_i}\right] DS_i * (1 - CV_i) * log_2((1 - CV_i)) \qquad (5)$$

where SIM_i is a Euclidean distance measure between the attributes of the classes associated with the existing and the new system [15, 20], DS_i is the design status that takes into account the percentage of new design, variable from 0 to 1, and NF_i is the number of functions performed by the class. Similarity coefficients must be weighted according to the relevance of the features by an expert judgment approach.

For the inter-complexity the proposed metrics is as follows [15]:

$$H_{ext} = -\sum_{i=1}^{n} NM_i * (1 - CV_i) * log_2(1 - CV_i) \qquad (6)$$

where NM_i is the number of messages exchanged in the sequence diagram by the i-th element with the other participants.

Total complexity is given by the sum of Eq. (5) and (6):

$$H_{tot} = H_{int} + H_{ext} \qquad (7)$$

3.3 A Fuzzy Model of the Cost Estimate Process

Next step of the proposed method is the derivation of a fuzzy model of the cost estimate process based on heuristic qualitative knowledge. Fuzzy modeling starts from the fact that a precise mathematical model of a complex system or process cannot be obtained [21]. The model links both the experience of the cost estimator and of the technical experts in the field of design with the output quality of the cost estimate. The first two ones constitute the input fuzzy variables. Their fuzzy domain is divided into five sub intervals, each covered by a membership function, represented in Fig. 1. Each sub interval represents the belonging to a particular set of technical or cost engineering skills. The output variable is represented by the "*goodness*" of the cost estimate, defined within the interval [0, 1]. Its fuzzy domain is built with the same type of membership functions as the input variables within five categories, from very bad to very good estimate output, as shown in Fig. 2. The fuzzy rules that link the two input variables with the output one and the membership functions are described in [15].

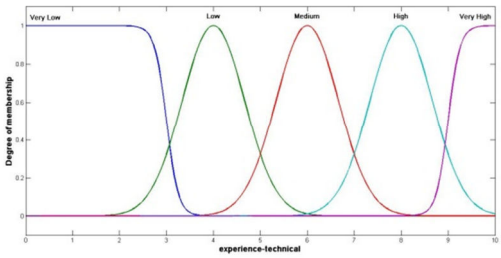


Figure 1 Gaussian and sigmoidal shaped membership functions representing the experience of the technical expert and of the cost estimator

Figure 3 shows the output of the fuzzy model; the independent axes represent the experience of the cost estimator and of the technical expert; the z-axis represents the "*goodness*" of the estimate.

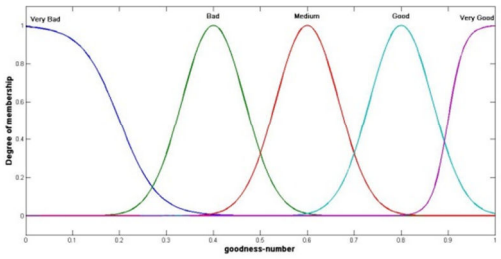


Figure 2 Gaussian and sigmoidal shaped membership functions for the output variable of the fuzzy process

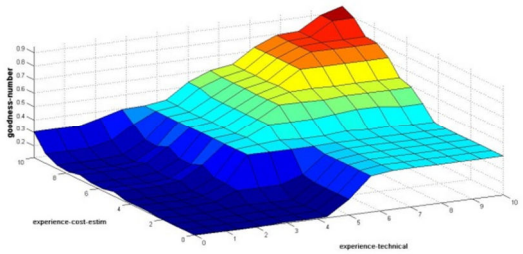


Figure 3 Output Surface representing the "*goodness*" of the estimate according to the proposed fuzzy model

3.4 A Definition Of Design to Cost

After having calculated the complexity of the UML model and having subsequently built the fuzzy rules of the human part of the cost estimate process, a preliminary assessment of the risks associated with the cost estimate are calculated according to the following equation [15]:

$$R = B*H_{tot} \qquad (8)$$

where B is the "*badness*", a fuzzy number derived from the output of the fuzzy model built in section 3.3, calculated as

$$B = 1 - Goodness \qquad (9)$$

H_{tot} is the complexity of the system, given by Eq. (7). A Design to Cost is obtained if the minimization of Eq. (8) is pursued, while trying to satisfy the requirements and being cost-efficient in the employment of the human resources since the beginning of the design phase.

The surface of Fig. 3 defines zones that may give useful indications in order to optimize Eq. (8):

- A zone, in dark blue color, where both the experience of the cost estimator and the technical expert is low; this zone should be used when the complexity is low (recurring activities);
- A zone, represented in yellow and red color, where both the experience of the cost estimator and the technical expert is high; this zone should be used when the corresponding system complexity is high (development of new systems);
- The remaining surface portions could be used in the case of a mix between recurring and not recurring design activities.

4 UML Modeling Procedure of a Lunar Rover

At Italian Space Agency (ASI) a study has been performed by means of the Concurrent Engineering methodology, that has had the aim to define, starting from mission requirements, a system architecture and its associated subsystems in order to evaluate the feasibility of the realization of a Lunar Rover mission (RLI –

Rover Lunare Italiano) within the required time and budget constraints, providing also an evidence on the coherence, completeness and consistency of the above mission requirements. As a final outcome, the study has allowed to define a preliminary configuration for the RLI able to fulfill a set of consistent requirements [22-23].

In this section the methodology described in section three has been applied to the cost risk estimate process.

4.1 Lunar Rover Requirements and Design Drivers

According to the procedure of section 3.1, two use case diagrams have been derived, the first one for Mission Objectives, represented in Fig. 4, the second one for Mission Requirements and Constraints, represented in Fig. 5.

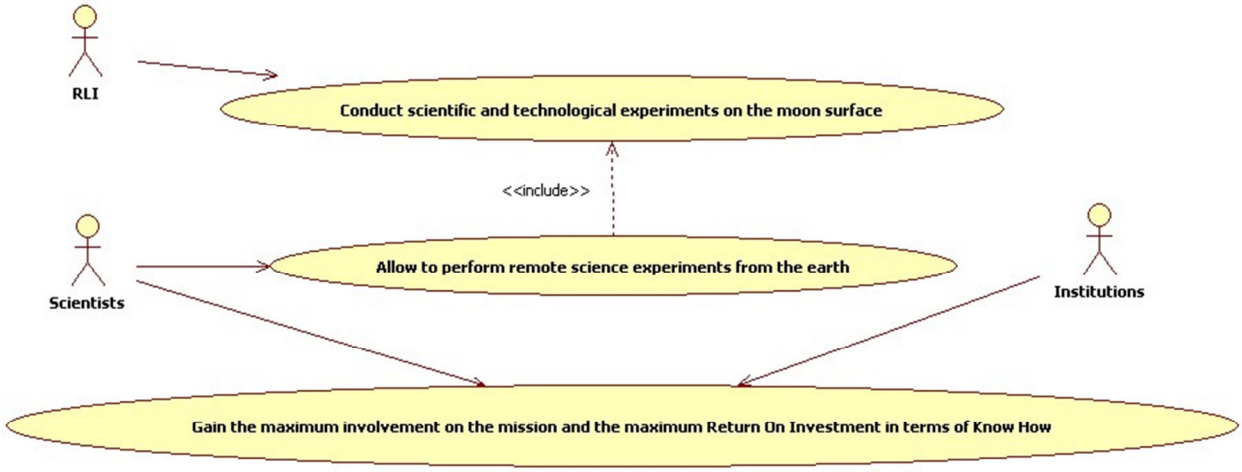


Figure 4 RLI Mission Objectives Use Case diagram

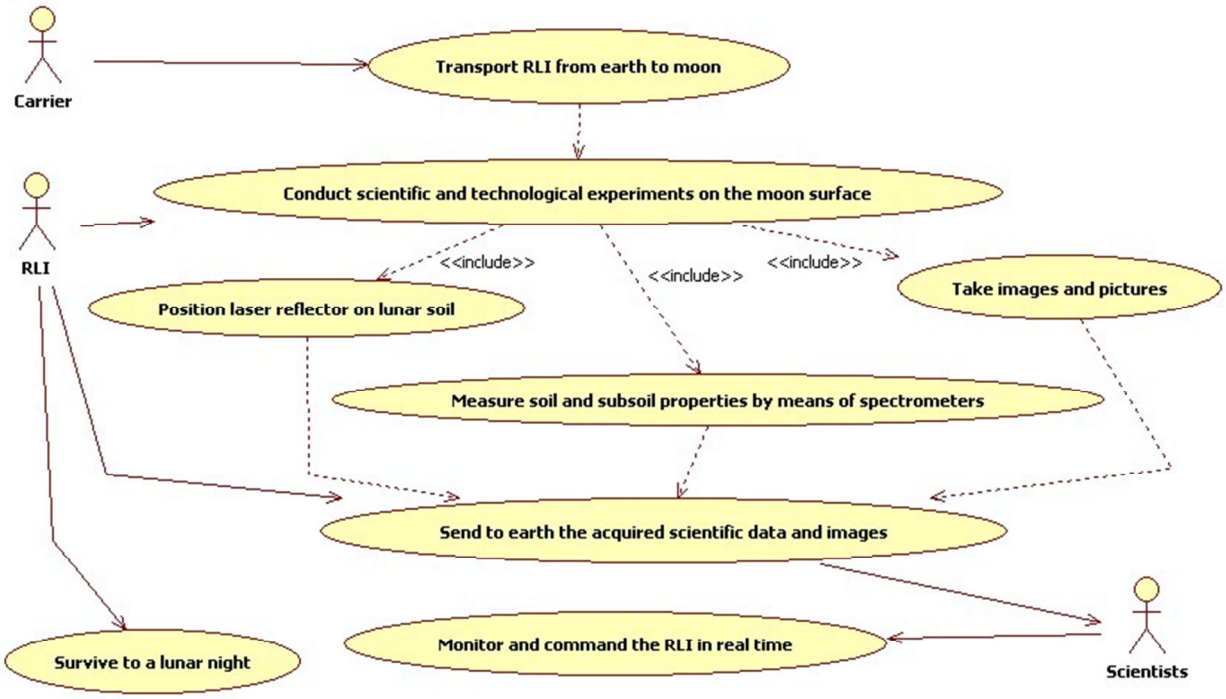


Figure 5 RLI Mission Requirement Use Case diagram

Four main sequence diagrams have been derived, two of them represented in figure 6 and 7.

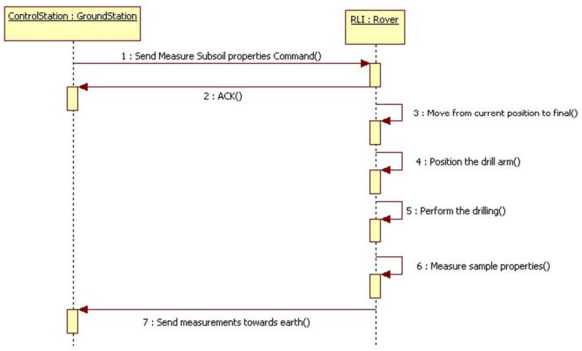


Figure 6 "*Measure subsoil properties*" Sequence diagram

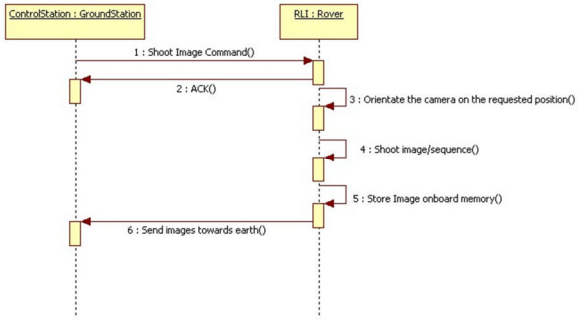


Figure 7 "*Shoot images*" Sequence diagram

4.2 Calculation of RLI GNC subsystem cost-risk factor

The sequence diagrams show how one of the most critical functions necessary for the lunar rover to efficiently manage its mission is the GNC one.

Lunar rover missions have been categorized as long range (cross-country) travels and limited range (local) explorations in terms of their traversing area and the nature of the vehicles tasks. The most conservative requirement for the driving control of a lunar rover is to stop when a perception sensor detects any obstacle ("*move and wait*" approach). A minimum look-ahead distance is defined to be continuously scanned for any obstacles in order to ensure the safety of the vehicle [24]. Approaches for Guidance and Navigation of Lunar Rovers are described in [24 – 28].

The preliminary class diagram of Fig. 8 has been derived by the sequence diagram analysis and the review of the literature.

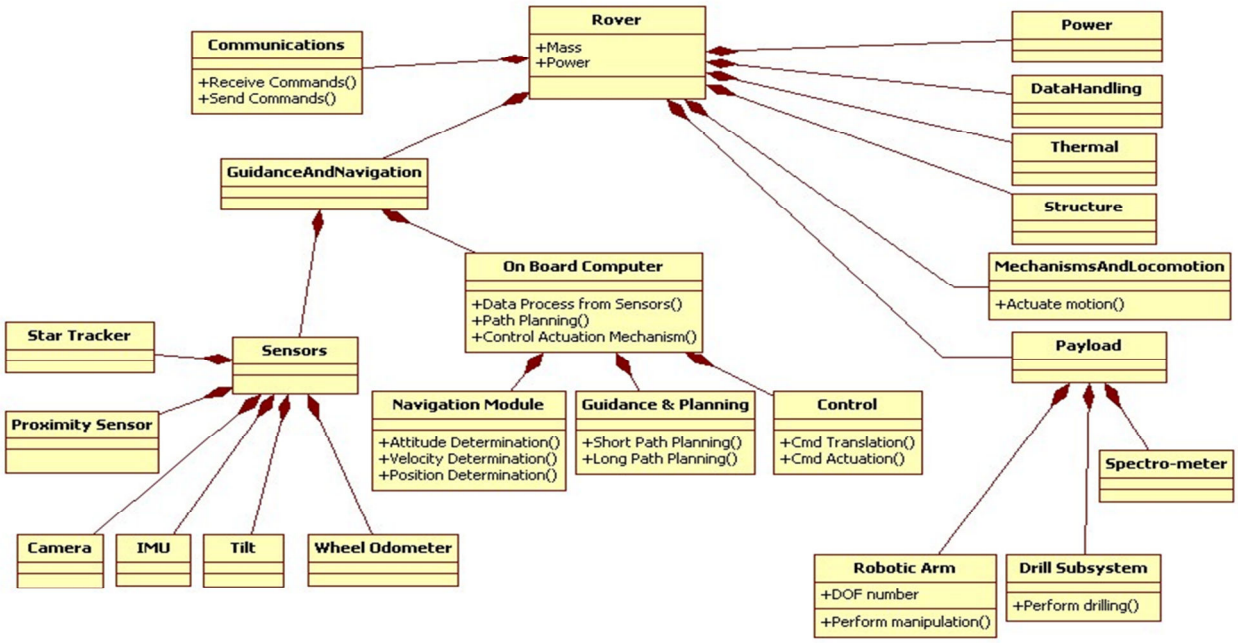


Figure 8 - Preliminary RLI UML class diagram

Four main functions are needed for the autonomous navigation of a rover [28]:

- Feature detection;
- Stereo matching (search the left and right image to find the corresponding locations for features detected);
- Triangulation (compute three dimensional position of each pair matching pixels);
- Feature tracking (After the motion of lunar rover, the features change. The procedure to locate the features detected before motion in the camera view after motion is called feature tracking).

Two sequence diagrams have been derived, the first one relative to a manual or "*wait and move*" procedure (Fig. 9), the other one relative to an autonomous navigation system (Fig. 10). The corresponding complexities have been calculated according to Eq. (7). A TRL value of 5 has been considered for all the elements; the number of functions and messages in Eq. (5) and (6) have been derived by the sequence diagrams of Fig. 9 and 10.

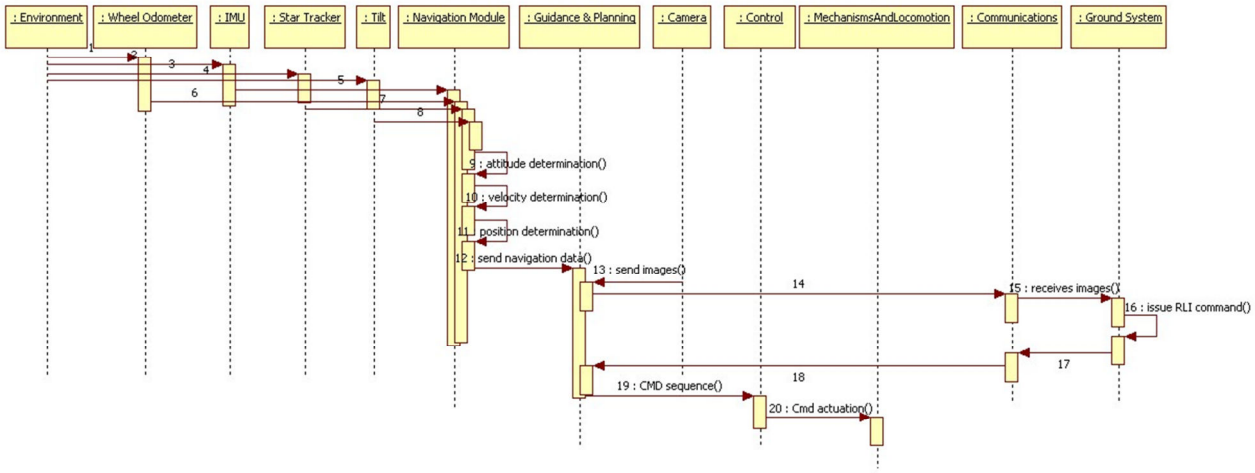


Figure 9 "*Wait and move*" RLI sequence diagram

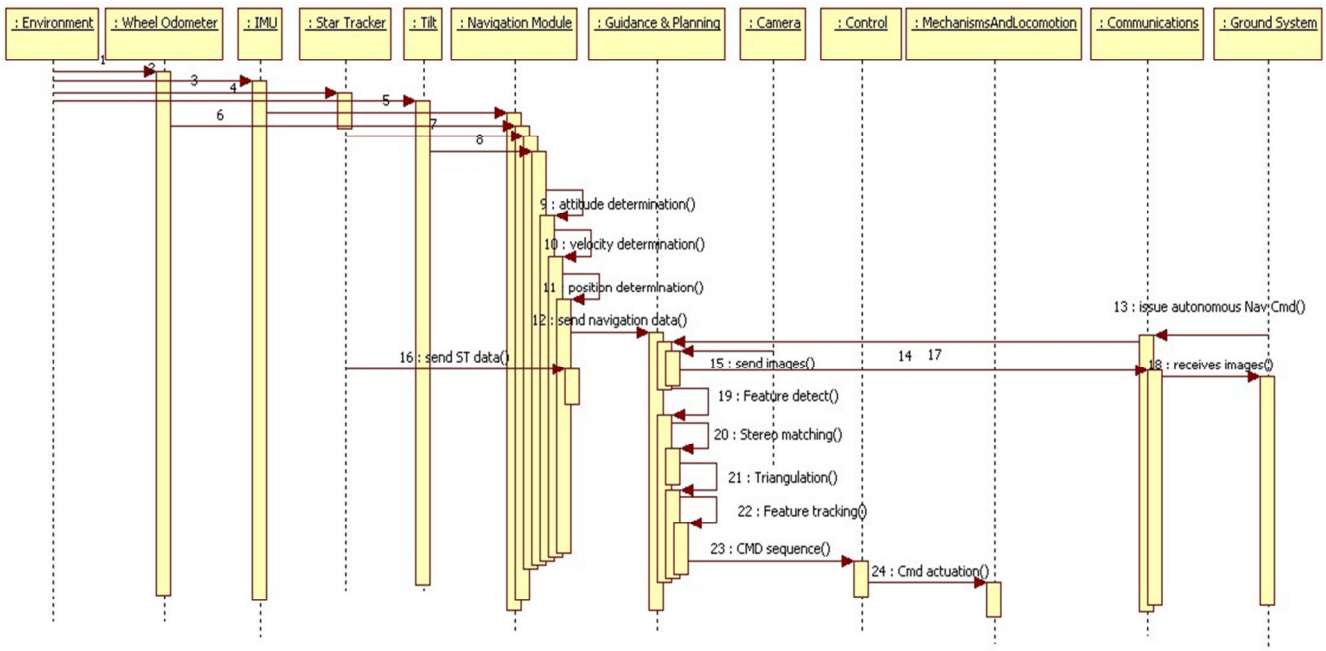


Figure 10 "*Autonomous navigation*" RLI sequence diagram

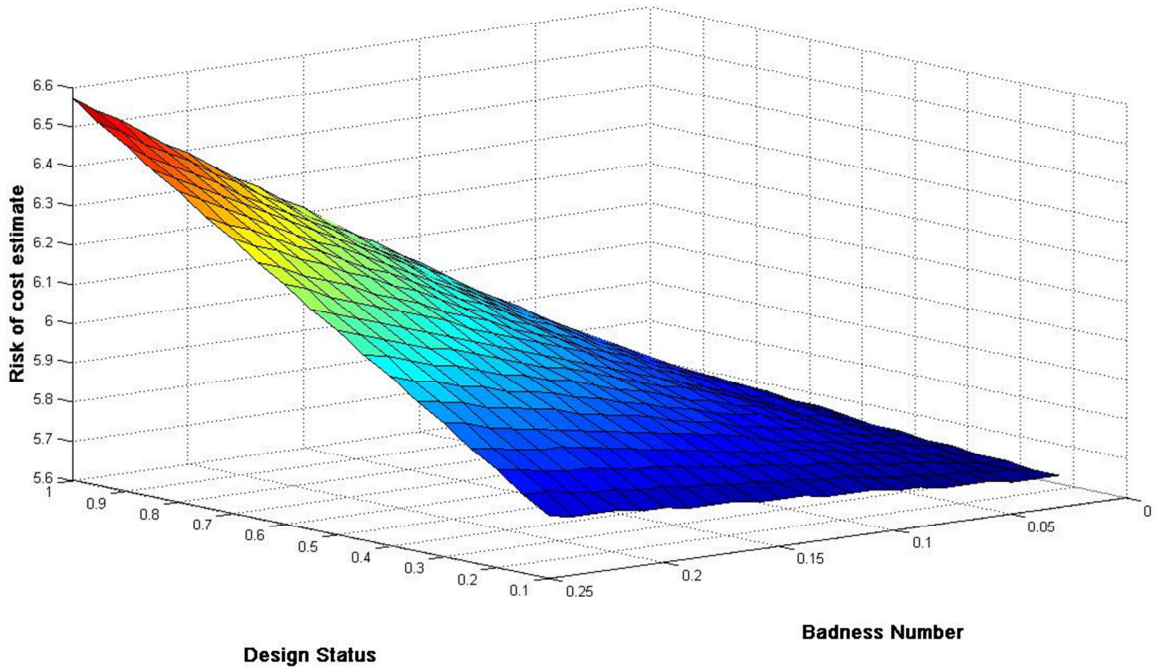


Figure 11 Relative Risk of the cost estimate of GNC subsystem (autonomous navigation) as a function of the design status (ranging from slight modifications to existing equipment – 0.1 – up to totally new design – 1 – and the skill of the people involved in the preliminary design phase.

The resulting risk of the estimate, visualized in the case of the autonomous GNC system, is shown in Fig. 11. The design status, ranging from 0.5 to 0.7, coupled with a badness number estimated at 0.2, put the GNC project in the high part of the risk zone, thus confirming the choice of a prototype model philosophy as the option for risk and cost minimization.

Figure 12, shows the estimated RLI Engineering Model (EM) and Engineering Qualification Model (EQM) costs.

5 Conclusion

A definition of design to cost has been presented, based on a synergic use of UML object modelling, information theory concepts and a fuzzy model of the cost estimate process. The proposed equation may be used as a precursor quantitative measurement of the risks of the estimate itself. Complementing and integrating HW complexity with SW complexity measurements should increase the proposed architectural model ability to capture total system complexity obtaining a more precise early warning indication of possible risks. The mission requirement analysis of the RLI project, used as a study case, has permitted to point out some elements to be taken into account for the development of the lunar rover itself.

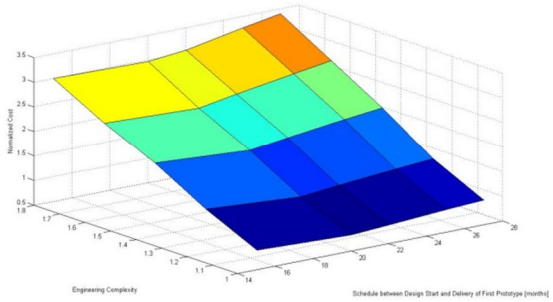


Figure 12 Normalized RLI EM and EQM cost, function of both the engineering complexity and the interval between the beginning of the design phase and the EM delivery. The cost increases dramatically with the engineering complexity. For a fixed value of complexity, there is an optimum cost effective schedule; if the schedule deviates from that point, the cost increases.

References

[1] OMG UML Superstructure, V 2.1.2 Nov. 2007, available at http://www.omg.org/spec/UML/2.1.2/

[2] Drusinsky D., "From UML Activity Diagrams to Specification Requirements", *SoSE '08 IEEE International Conference on System of Systems Engineering*, 2008

[3] Emadi S. and Shams F., "Transformation of use-case and sequence diagrams to petri nets", *CCCM 2009, ISECS International Colloquium on Computing, Communication, Control, and Management*.

[4] Meziane F., Athanasakis N. and Ananiadou S., "Generating Natural Language Specifications From UML Class Diagrams", *Requirements Engineering*, Volume 13, Number 1, 1-18

[5] Mindock, J. ; Watney, G., "Integrating System and Software Engineering Through Modeling", *IEEE Aerospace Conference,* 2008

[6] Di Domizio D. and P. Gaudenzi, "A Model for preliminary design procedures of satellite systems", *Concurrent Engineering*, Vol. 16, n. 149, 2008

[7] Belvedere G., Filippazzo G. and Gaudenzi P., "Cost evaluation of satellite subsystems by means of an integration between a Cost Estimating Relationship software and a concurrent engineering design tool", *3rd International SECESA Workshop*, Rome 15-17 October 2008

[8] Halstead M., *Elements of Software Science*, Elsevier, 1977

[9] McCabe T., "'A Complexity Measure," *IEEE Trans Software Engr*, SE-2 (4): 308-320, Dec 1976

[10] Bertolino A., Bonivento A., De Angelis G. and Sangiovanni Vincentelli A., "Modeling and Early Performance Estimation for Network Processor Applications", *Lectures Notes in Computer Science*, 2006, Vol. 4199/2006, 753-767

[11] Briand L. C. and Wüst J., "The Impact of Design Properties on Development Cost in Object-Oriented Systems", Technical Report *ISERN-99-16*

[12] Lavazza L. and Robiolo G., "Introducing the Evaluation of Complexity in Functional Size Measurement: a UML-based approach", *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*,

[13] Podgorelec V. and Hericko M., "Estimating Software Complexity from UML Models", *ACM Sigsoft Software Engineering Notes*, Volume 32, Number 2, March 2007

[14] Lankford J., "Measuring system and software architecture complexity", *Proceedings of the IEEE Aerospace Conference*, 8(8-15, 2003):3849–3857, March 2003

[15] Belvedere G., "Object Oriented Modeling as a Design to Cost Method for Space Missions", PhD Thesis, Mech. And Aerospace Dept., "La Sapienza" University, Rome 2011

[16] Wade D.I. and Welch C.S., "Spacecraft Manufacturing Implications for Volume Production Satellites", *47th International Astronautical Congress*, Beijing, China, 1996

[17] Shannon C.E., *The Mathematical Theory of Communication*, The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July, October, 1948

[18] Chatzigeorgiou A. and Stephanides G., "Entropy As a Measure of Object Oriented Design Quality", *Proceedings of the Balkan Conference in Informatics*, pp. 565-573, 2003

[19] Wertz J.R. and Larson W.J., "*Space Mission Analysis and Design*", Space Technology Library ed., 2003, Chap. 20, 4

[20] Shepperd M. and C. Schofield, "Estimating Software Project Effort Using Analogies", *IEEE Transactions On Software Engineering*, Vol. 23, No. 12, November 1997

[21] Sugeno M. and Yasukawa T., "A Fuzzy-Logic-Based Approach to Qualitative Modeling", *IEEE Transactions On Fuzzy Systems*, Vol. 1, No. 1, February 1993

[22] Cifani G., "Analisi Tecnico Economica di un Rover Lunare", Master Thesis in Aerospace Engineering, Aerosp. Eng. Dept., "La Sapienza University", Rome 2010

[23] Battagliere M. L. et al., "First Studies of ASI Concurrent Engineering Facility (CEF)", 4*rd International Workshop on System and Concurrent Engineering for Space Applications SECESA 2010*, Lausanne, 13 – 15 October

[24] Yoshioka N. and Wakabayashi Y., "Driving technology and preliminary testing of a lunar rover", *Control Eng. Practice*, Vol. 5, No. 6, pp. 845-850, 1997

[25] Fuke Y. and Krotkov E., "Dead Reckoning for a Lunar Rover on Uneven Terrain", *Proceedings of the 1996 IEEE International Conference on Robotics and Automation*, Minneapolis, Minnesota - April 1996

[26] Kuroda Y., Kurosawa T., Tsuchiya A., Kubota T., "Accurate Localization in Combination with Planet Observation and Dead Reckoning for Lunar Rover", *Proceedings of the 2004 IEEE lnternational Conference on Robotics and Automation*, New Orleans, LA April 2004

[27] Kunii Y., Kuroda Y., Suhara M. and Kubota T., "Command Data Compensation for Real-time Tele-driving System on Lunar Rover: Micro-5", *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, Seoul, Korea . May 21-26, 2001

[28] Pingyuan C., Fuzhan Y., and Hutao C., "Research on Autonomous Navigation of Lunar Rovers for the Moon Exploration", *Proceedings of the 2006 IEEE International Conference on Robotics and Biomimetics*, December 17 - 20, 2006, Kunming, China

CEAS 2011

Poster Session

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Numerical Study of Flow Separation in Advanced Nozzles

Francesca Scaramuzzino, Emanuele Martelli
Second University of Naples SUN, Italy

Francesco Nasuti
University of Rome "La Sapienza", Italy

Keywords: *Nozzle, Dual bell, Inflection region, Flow Transition, Flow Separation*

Abstract

Dual-bell is a kind of advanced nozzle concept. The dual-bell design is such that it can operate at two different altitudes. This double adaptation is obtained through a geometrical discontinuity on the nozzle wall profile. At low altitude the flow is attached throughout the first bell and a controlled separation is reached. At high altitude the nozzle operates with attached flow in the whole divergent section. Different wall shape geometries can be adopted on the divergent part and differently affect the nozzle efficiency. A well proven way to improve performances, when dealing with open-cycle engine, is the injection of the turbine exhaust gases into the divergent section. The aim of this work is to analyze how the injection of a secondary gas into the main flow modifies the chamber-to-ambient nozzle pressure ratio (NPR) at which transition occurs. Steady-state simulations have been conducted either for a configuration without the gas injection (clean case) either for a configuration with the gas injection (film cooling case). Numerical results demonstrate that the film cooling effectively changes the transition NPR, in particular the NPR obtained with the film cooling is higher than that one achieved in the clean case.

1 Introduction

Today's liquid rocket engines use conventional bell-nozzle type. These ones are limited because of their fixed geometry and therefore one optimum design point is achieved. Another unavoidable limit is the onset of side-load forces during start-up and shut-down [13]. To overcome such a drawback, in the last years, new nozzle concepts have been expounded [1, 11]. One of them is the *dual-bell* nozzle. The dual-bell has two bells in the divergent part with a discontinuous wall curvature that controls the flow separation preventing side-load onset [3]. The so-called inflection point causes a twofold type of flow fields. At low altitude the base nozzle is full flowing and completely separated in the rest of the divergent part. At high altitude the flow is totally attached in the whole divergent part. Nevertheless side-loads remain a concern for the dual-bell nozzle design because of their rise during the transition. Actually, in the ideal case a controlled separation is achieved since the separation point is placed about a precise location, the inflection point, for a wide range of nozzle NPR's. In the real case, due to viscosity, the ideal step becomes a region of

finite length in which the pressure gradient is negative as in conventional nozzles. This region, called *inflection region* [3], offers stable and unsymmetrical positions to the separation point leading to the onset of side-loads.

In [3, 4] the influence of different parameters on the inflection region has been investigated.

A current way, used on conventional nozzles, to improve efficiency is the injection of the turbine exhaust gas (TEG) into the main nozzle if a gas-generator open-cycle is adopted. The TEG injection provides a low-temperature gas that it is able to create an insulating effect that reduces the convective heat transfer rate from the hot-gas stream to the exposed nozzle surface. Recent studies have been focused on the advantages of the film cooling application (a secondary jet inflow) on dual-bell nozzles from the point of view of performances and side-loads [2, 5].

In this work the injection of a secondary gas has been simulated and a numerical campaign has been executed to search the transition NPR for the clean case (without film cooling) and for the film cooling case. The effect of the interaction between the two jets on the transition NPR has been evaluated showing that the film cooling effectively increases the transition NPR, for the range of secondary mass flow rate considered.

2 Numerical approach

The numerical analysis is carried-out by a validated solver of Reynolds-Averaged-Navier-Stokes equations, which are integrated by an explicit second-order time-and-space accurate scheme. The code is a 2-D axisymmetric in-house solver [6, 10] and it is based on a lambda-scheme elaborated by Moretti [7].

Shock discontinuities are handled with a hybrid methodology: the differential quasi-linear form is replaced with an integral Godunov form in the field zones where shocks occur [10, 12]. Turbulence is computed by a modified version of the Spalart-Allmaras one-equation model [8].

3 Profile design method

An appropriate tool to draw dual-bell profiles has been adopted [3]. This design tool is based on the characteristics method (MOC).

The dual-bell nozzle design is split into two steps: the base design and the extension design, as depicted in Fig. 1.

For the base a truncated nozzle has been chosen and it has not been changed for the different extensions.

The second bell is designed by means of MOC: once the pressure behaviour of the second operative mode is assigned, the outline is designed as a streamline.

The data required are the base area ratio, $\varepsilon_1 = A_{base}/A_{throat}$, the extension length ratio, L_e/L^* (where L_e is the extension length and L^* is a reference length) and wall pressure behaviour. The procedure is reported in [3].

Three typologies of extension profile can be designed: negative, constant and positive gradient profile.

The second bell profile with negative pressure gradient behaves like conventional nozzles with stable positions of the separation location along the whole profile. Thus a negative wall pressure gradient profile has to be excluded for dual-bell nozzles. Constant and positive pressure gradient would lead, respectively, to indifferently stable and unstable position of the separation location.

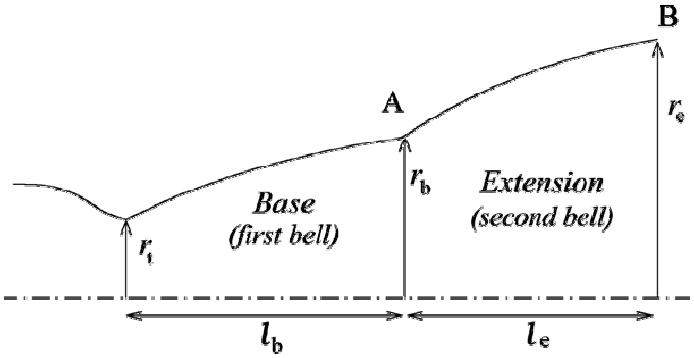


Fig. 1 Dual Bell Geometry

4 Test case design and characteristics

All the simulations have been carried out adopting a 2-D axisymmetric cold gas sub-scale nozzle whose geometry is referred to the experimental study conducted in [9].

Main geometric properties and operating conditions are reported in Table 1.

Throat radius	0.01 m
L* reference length	0.01 m
Base area ratio ε_1	9.61
Base length L_b/L^*	5.2
Chamber temperature	300 K
γ (nitrogen)	1.4
Ambient pressure p_a	0.1 MPa

Table 1 Base Geometric and Operating conditions of the dual bell nozzle

4.1 Validation test case

As already described, the profile shape is identified by the wall pressure behaviour along the extension during high-altitude operation. At first two profiles families have been considered, named COP (constant pressure profile) and LIP (linearly increasing pressure profile). The geometry taken from [9] is related to a COP profile. The same data have been used to design a LIP profile.

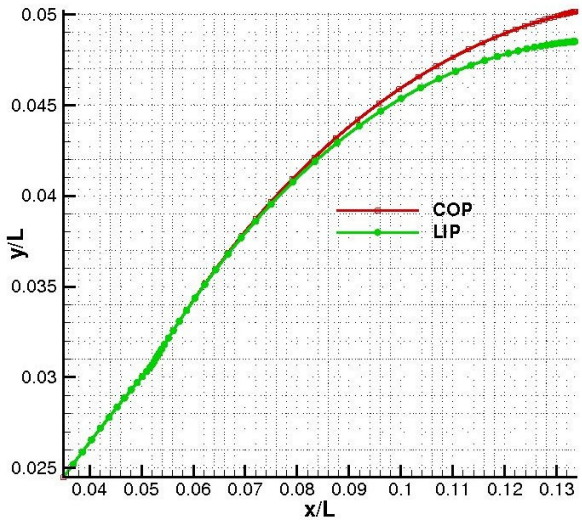


Fig. 2 Comparison between COP and LIP dual bell profiles

Figure 2 shows that, drawing a LIP profile leaving the same L_{tot}/L^* of the COP profile, the area ratio changes, as reported in Table 2.

The COP profile has been employed in order to compare the numerical solution with experimental data taken from a test conducted at DLR's cold gas stand [9].

Figure 3 shows a good correspondence between the numerical solution and experimental data.

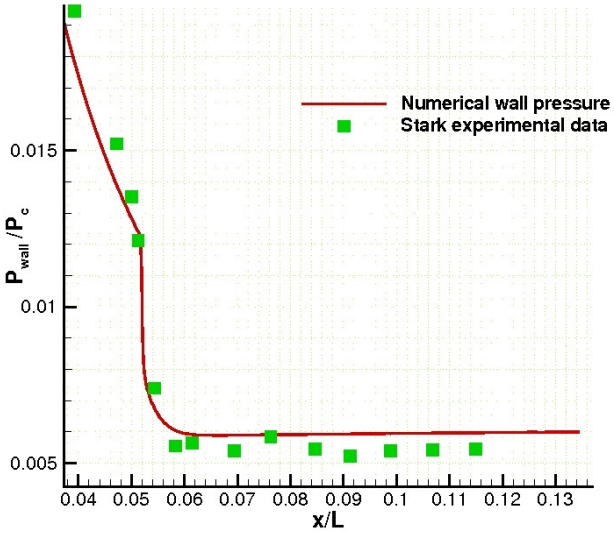


Fig. 3 Comparison between numerical solutions and experimental data taken from [9]

	COP	LIP
Throat radius	0.01 m	0.01 m
L*	0.01 m	0.01 m
Base area ratio	9.61	9.61
Base length L_b/L^*	5.2	5.2
Extension area ratio	25.6	23.06
Total length L_{tot}/L^*	13.5	13.5
Extension length L_e/L^*	8.3	8.3

Table 2 Geometric and Operating conditions of the COP and LIP dual bell profiles

The computational domain is defined by the nozzle geometry and external field.

A multiblock grid is used to suitably distribute cell density according to the zones of major interest. The nozzle is divided in three blocks. The second block includes only the zone of the inflection point in order to enhance the numerical solution in the region with a wall pressure negative gradient. Three rectangular domains are connected to the nozzle domain to simulate the outer field. The nozzle is characterized by subsonic inflow boundary conditions enforcing total temperature and total pressure with the flow direction, symmetric axis

and adiabatic wall. The external domain is characterized by subsonic inflow on the left side, with the same assumptions given for the nozzle domain, assigned pressure on the upper wall (whose height is $4R_{ext}$ where R_{ext} is the radius of the extension cross-section) and non-reflecting boundary conditions on the right side (whose length is $6R_{ext}$) as reported in Fig. 4.

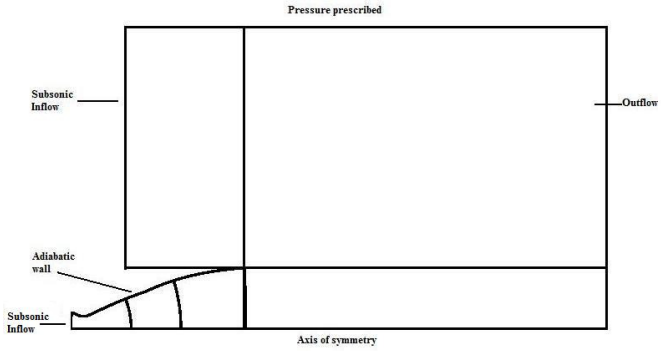


Fig.4 Computational domain and boundary conditions of the dual bell

4.2 Film cooling test case

For the comparison between clean case and film cooling case the LIP profile has been chosen. Actually, it has been demonstrated in [3] that profiles with positive pressure gradient show a better transitional behaviour with a sudden jump of the separation point from the inflection point region to nozzle exit and display only a slight reduction of nozzle efficiency. The secondary flow inlet is made by means of a slot located in the divergent part upstream the inflection point. The slot is simulated through the translation of the wall profile outward from the injection point to the nozzle end. The slot geometry is reported in Table 3 and in Fig. 4.

Slot height [m]	0.001
Slot length [m]	0.01
L*	0.01
Slot initial abscissa, x_{slot}/L^*	0.025
γ of the feeding gas	1.4

Table 3 Geometrical and operating conditions for the slot simulating the film cooling inlet

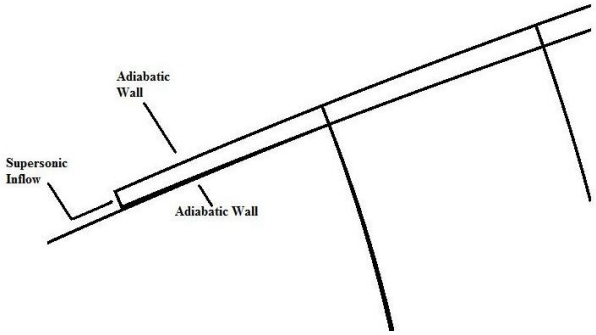


Fig. 5 Enlargement of the arrangement for the film slot

The nozzle is divided in eight blocks. The further blocks downstream the secondary gas block injection are applied to properly describe the mixing layer zone, see Fig. 5 and Fig. 6.

Four blocks have been taken in account to simulate the external field. The inlet and outlet main flow have the same boundary conditions previously enumerated. At the supersonic film inlet static pressure, static temperature, Mach number and flow direction are enforced. The inlet static temperature and static pressure of the film are set equal to that ones of the main flow in order to have an adapted secondary jet. A parametric analysis of dual-bell film cooling has been carried-out by considering four different values of the Mach number film [1.1, 1.4, 1.8, and 2.2].

As can see in Fig. 5, the film cooling is injected into a channel with adiabatic wall with a length of 0.01 m and a height of 0.001 m, see Table 3.

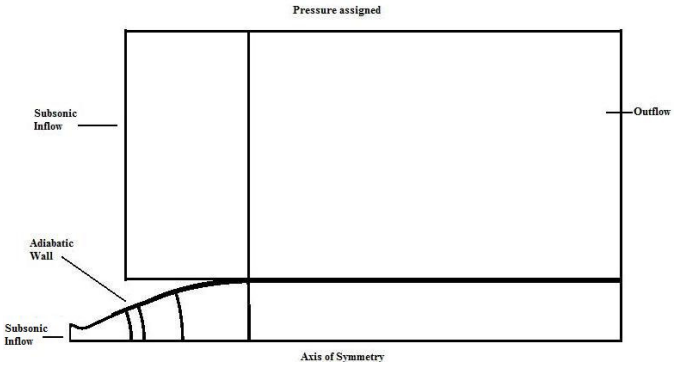


Fig. 6 Film cooling configuration and boundary conditions

5 Numerical results

In the following part numerical results are reported.

5.1 LIP flow field description

Figure 7 shows two pressure profiles for a LIP extension. At NPR=23 the flow is anchored at the inflection region and it works with the first bell (first operative mode, OMI) increasing efficiency by means of the reduced overexpansion. At NPR=30 the flow transition is occurred and the second area ratio becomes operative (second operative mode, OMII). In this way the nozzle operates with a higher area ratio achieving better performances.

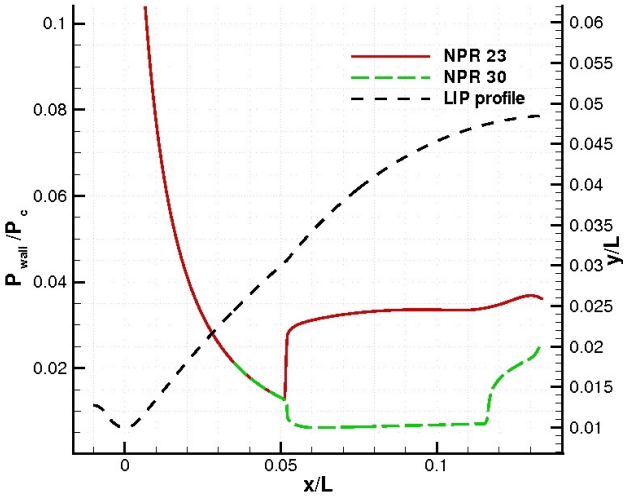


Fig. 7 LIP profile: separation point in the inflection region at NPR=23 and in the exit plane at NPR=30

The following figures describe the Mach number flow field for the LIP profile.
Figure 8 depicts the nozzle in the OMI. The separation point is anchored at the inflection point and the divergent part is filled with the back flow from the ambient. In the recirculation zone, where the flow is separated, the mixing layer between the main supersonic flow and the back flow can be noticed.

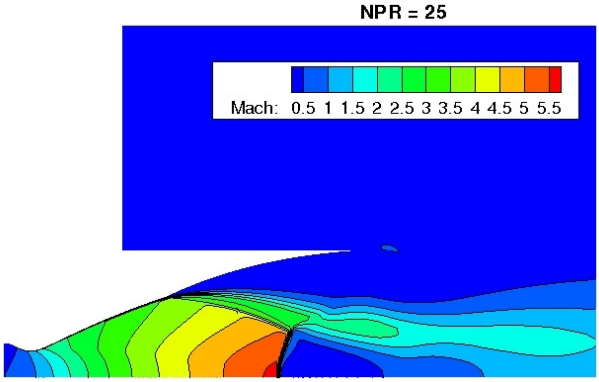


Fig. 8 LIP Mach number flow field at the transition NPR= 25

For a higher NPR the flow is attached in the whole divergent part, as can be seen in Fig. 9. The separation point and the shock pattern are located near the exit plane of the nozzle.

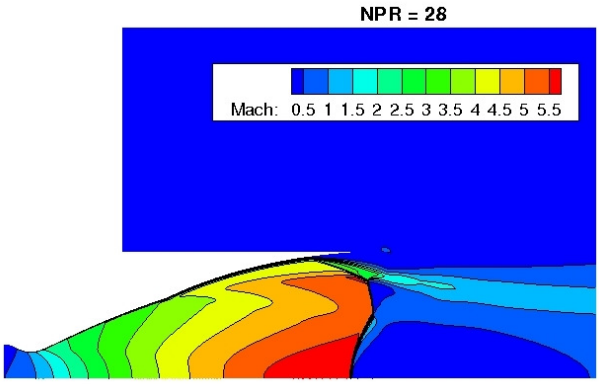


Fig. 9 LIP Mach number flow field at the transition NPR= 28

Figure 10 illustrates the sequence of steady-state simulations carried out for the LIP configuration. The sequence of simulations considers the NPR values from NPR=23 to NPR=30. The wall pressure behavior is that typical of a free shock separation: at the separation point there is the sudden increase of pressure due to the compression shock. Then the pressure continues to increase in the recirculation zone, up to the value of the ambient pressure. From numerical results it can be observed that with the positive gradient profile the separation point takes place on the inflection point at NPR= 25. At NPR= 27 the separation point is located in the inflection region while at NPR= 28 the transition is occurred, as reported in Fig. 10.

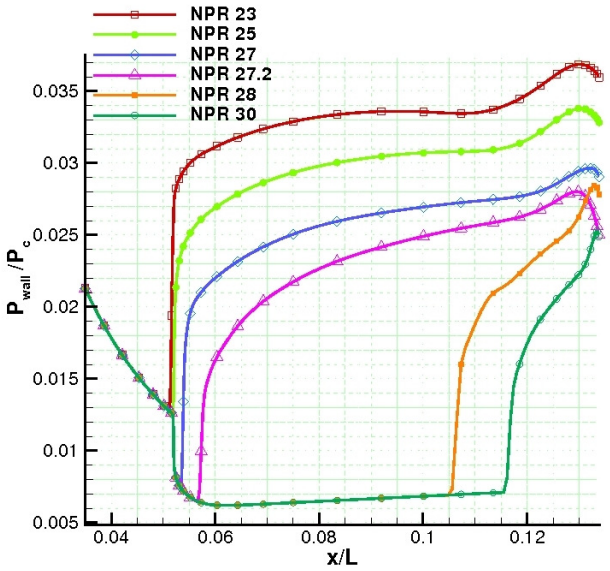


Fig. 10 LIP profile: Steady-state analyses varying NPR

5.2 Film cooling effect on the transition NPR

Four different sets of steady-state computations have been carried out depending on different Mach number film (M_{film}) value adopted [1.1, 1.4, 1.8 and 2.2]. All of the computations have been conducted with similar grids. A different value of mass flow rate ratio $\chi = \frac{\dot{m}_f}{\dot{m}_{main}}$ is associated with each test, as reported in Table 4. The value of the $\dot{m}_{main}$ = 2.02278 kg/s.

Mfilm	χ
1.1	0.0325
1.4	0.0419
1.8	0.0539
2.2	0.0655

Table 4 Secondary flow conditions

The mass flow rate is a function of the p_{film}, T_{film} and M_{film} as reported in eq. (1). The slot geometrical data are constant. Since the secondary jet has the static properties of the main flow, the mass flow rate depends only by the M_{film} value.

$$\dot{m}_f = \rho u A = \frac{p}{RT} a M_{film} A = \sqrt{\gamma} \frac{p}{\sqrt{RT}} M_{film} A \quad (1)$$

Thus, the change of the M_{film} has the effect to vary the mass flow rate injected into the secondary duct. In Fig. 11 the total temperature field has depicted for the case of M_{film} = 1.1. In particular an enlargement of the secondary jet has reported. It is visible the onset of the mixing layer when the film cooling meets the main flow. The different flow velocities and, thus, mass flow rates, yield different wall temperature behaviours along the film cooled wall. In particular, just near the inflection point the wall temperature is that one of the secondary flow.

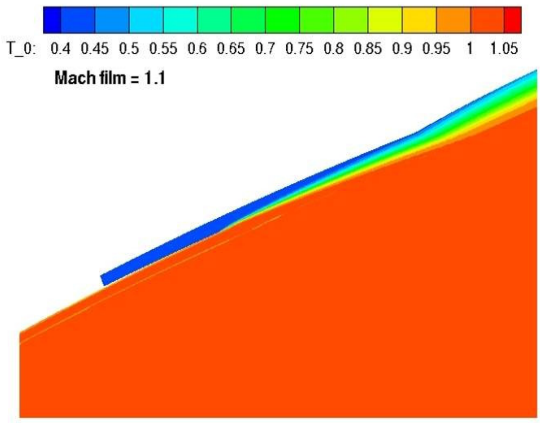


Fig. 11 Total temperature field for M_{film}=1.1: onset of the mixing layer

Then the mixing layer occurs, see also Fig. 11. The wall temperature is maintained nearly constant until it is influenced by the total temperature value of the main flow, see Fig. 12. Just downstream the inflection point, the wall temperature reduces because of the expansion occurring at the inflection point and then it starts increasing due to mixing. It can be observed that the secondary flows at higher M_{film} are less influenced by the main flow.

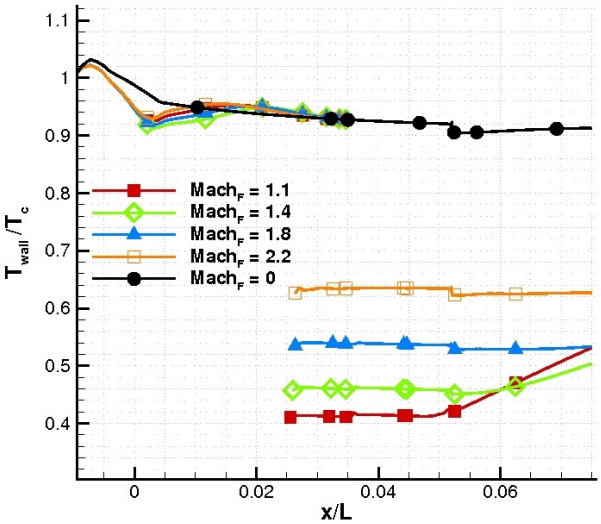


Fig. 12 Wall Temperature profiles at varying M_{film}

For the case of M_{film}= 2.2 Fig. 13 shows the wall pressure distributions for the increasing pressure ratio, from NPR= 22 to NPR= 29.

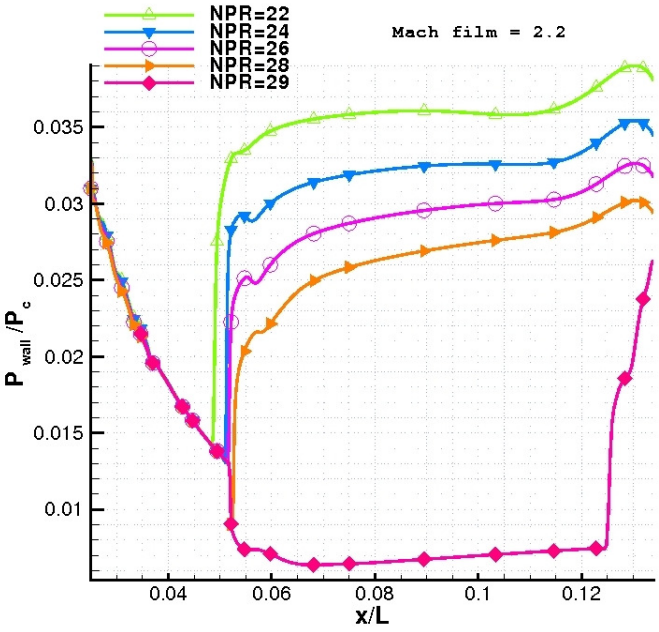


Fig. 13 Wall pressure distribution for varying NPR at M_{film}= 2.2

Figure 14 depicts the results achieved through the simulations, including the case without the film cooling.

The Table 5 gives the different transition NPR for each M_{film} case. Figure 15 shows the transition pressure ratio as a function of M_{film}: in the clean case the PR is 28, while in the film-cooled cases the PR increases up to 30 for M_{film}=1.1, to 31 for M_{film}= 1.4 and it remains the same for M_{film}= 1.8 and then it falls to 29 for M_{film}=2.2.

The transition NPR in the film-cooled configuration increases by the 9% compared to the clean case.

M_{film}	NPR
0	28
1.1	30
1.4	31
1.8	31
2.2	29

Table 5 Different transition NPR at M_{film} varying

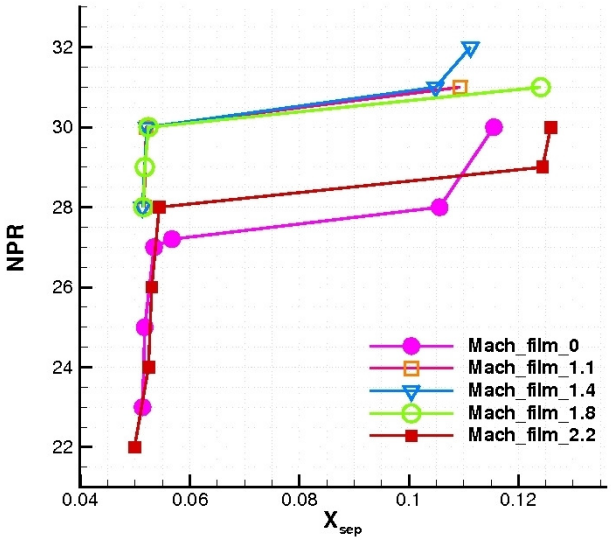


Fig. 14 Position of the separation point abscissas for increasing pressure ratio NPR at varying M_{film}.

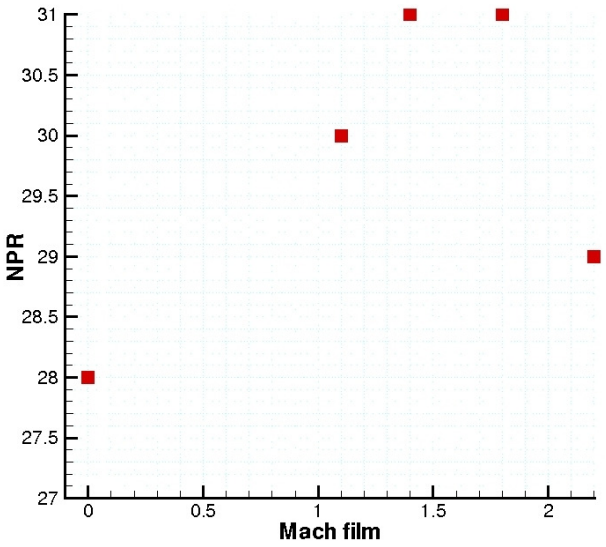


Fig. 15 Different transition NPR at varying $Mach_{film}$

About the inflection region, from Table 6, it can be noticed that the increasing M_{film} does not make evident changes on the inflection length.

M_{film}	l_i [cm]	l_i/l_e
0	1.015	12.22 %
1.1	1.624	19.56 %
1.4	1.660	20.00 %
1.8	1.646	19.83%
2.2	1.645	19.82%

Table 6 Inflection region length at $Mach_{film}$ varying

However, the presence of the secondary flow inlet makes the inflection region length larger than that one of the clean case, as depicted in Fig. 16.
As a consequence the separation point can find stable positions, in the OMI, in a wider zone, increasing the negative effect of the side- loads.

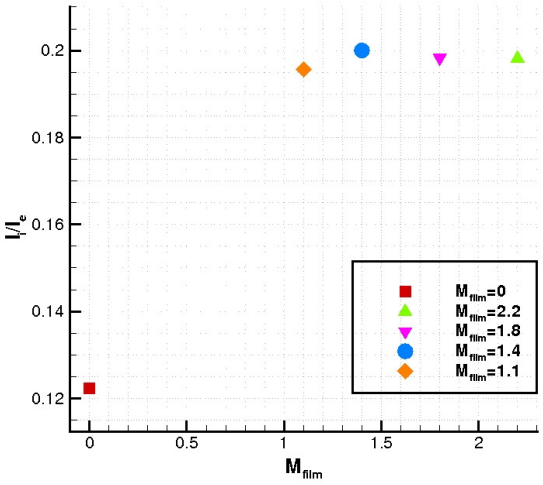


Fig. 16 Length of the inflection region vs Mach$_{film}$

Conclusion

This work has been focused on the effect of the secondary gas jet injection on the transition NPR and on the inflection region length considering steady state simulations. Numerical results have demonstrated that, including the secondary gas injection, the transition NPR increases. Instead the M_{film} variations slightly influence the transition pressure ratio. The presence of the film cooling causes the enlargement of the inflection region, increasing the onset of the side-loads. It has been demonstrated that the rise of the cooling mass flow rate, i.e. the enhancement of M_{film} among the different values adopted, does not affect anymore the inflection region length. Further parametric tests are foreseen to better understand the effect of the secondary gas on the flow transition by means of time accurate simulations.

References

[1] Hagemann G., Immich H., Nguyen T.V. and Dumnov G.E., "Advanced Rocket Nozzles", *Journal of Propulsion and Power*, Vol. 14, No. 5, Sept-Oct 1998, pp. 620-634.

[2] Martelli E., Nasuti F. and Onofri M., "Numerical Analysis of Film Cooling in Advanced Rocket Nozzles", *AIAA Journal,* Vol. 47, No. 11, 2009, pp.2558-2604.

[3] Onofri M., Nasuti F. and Martelli E., "Role of Wall Shape on Transition in Axisymmetric Dual-Bell Nozzles", *Journal of Propulsion and Power*, Vol. 21, No. 2, 2005, pp.243-250.

[4] Martelli E., Nasuti F. and Onofri M., "Numerical Parametric Analysis of Dual-Bell Nozzle Flows", *AIAA Journal,* Vol. 45, No. 3 , 2007, pp.640-650.

[5] Martelli E., Nasuti F. and Onofri M., "Film Cooling Effect on Dual-Bell Nozzle Flows", JPC, 2009

[6] Nasuti F. and Onofri M., "Analysis of Unsteady Supersonic Viscous Flows by a Shock Fitting Technique", *AIAA Journal,* Vol. 34, No. 7, Jul. 1996, pp.1428-1434.

[7] Moretti G., "A Technique for Integrating Two-Dimensional Euler Equations", *Computer and Fluids,* Vol. 15, No. 1, 1987, pp. 59-75.

[8] Spalart P.R. and Allmaras S.R., "A One-Equation Turbulence Model for Aerodynamic Flows", *La Recherche Aerospatiale*, No. 1, 1994, pp. 5-21.

[9] Nürnberger-Genin C. and Stark R., "Flow Transition in Dual Bell Nozzles", *Shock Waves*, Vol. 19, No. 3, 2008, pp. 265-270.

[10] Nasuti F., "A Multi-Block Shock-Fitting Technique to Solve Steady and Unsteady Compressible Flows, "*Computational Fluid Dynamics"* 2002, edited by S. Arnfield, P. Morgan and K. Srinivas, Springer-Verlag, Berlin, 2003, pp.217-222.

[11] Muss J., Nguyen T., "Evaluation of altitude Compensating Nozzle Concepts for RLV", *C33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, AIAA Paper, 1997.

[12] Paciorri R., "Sviluppo di una metodologia numerica in forma quasi-lineare e analisi di modelli termochimici per lo studio di flussi viscosi ipersonici", PhD. Dissertation, Department of Mechanics and Aerospace, La Sapienza University of Rome, Italy, 1995.

[13] Ostlund J., "Flow Process In Rocket Engine Nozzles With Focus On Flow Separation And Side-Loads" Ph.D. Dissertation, Department of Mechanics, Royal Institute of Technology, Sweden, 2002.

Numerical Studies of Nonlinear Aeroelastic Characteristics of Aircraft in Transonic Flow

Svetlana I. Kuzmina, Fanil Z. Ishmuratov, and Oleg V. Karas
Central Aerohydrodynamic Institute (TsAGI), Russia

Keywords: *flutter, shock wave, aeroelasticity, transonic flow*

Abstract

This paper describes the methods developed in TsAGI for aeroelasticity analysis taking into account specific responses of an elastic airplane in transonic flow.
For the analysis of strength, static and dynamic aeroelasticity characteristics at transonic velocities the mathematical model of an elastic aircraft is created on the basis of the ARGON software. Two approaches based on CFD codes for aeroelastic application are described. One of them is the Godunov method where linear flutter investigations in the frequency domain are supplemented with an iterative procedure. The second is effective algorithm of aerodynamic influence coefficients computation with taking into account a viscosity model. It is available for rather complex airplane configurations.

1 Introduction

Industrial prediction of aeroelastic characteristics in transonic flow taking into account a shock boundary layer interaction is really the actual for flying vehicles of different types. For modern transport and passenger airplanes the transonic flight regime is critical in view of the necessity to ensure the safety from such aeroelasticity phenomena as flutter, limit cycle oscillations, aeroservoelastic instability, static stability and controllability, aeroelastic divergence of lifting surfaces and controls reversal. Therefore it is important to perform an airplane aeroelastic analysis for complete flight envelope including critical transonic regime.

Till now the majority of industrial practical aeroelasticity analysis is conducted mainly with the use of linear aerodynamic methods. As a rule the classical Doublet-Lattice Method (DLM) is used. The necessity of an aerodynamic characteristic linearization has also another reason: application of effective methods in a stability analysis of complex linear dynamic systems.

The traditional approach allows investigating the majority of flutter types and modes that can appear in practice. Numerous experiments have shown that dips occur at transonic Mach numbers in the flutter boundaries. This phenomenon is caused by motion of shock waves on wings [1-3]. In order to describe the physics of such moving shock wave it is required to use computational methods for solving nonlinear partial differential equations in regions with mixed subsonic and supersonic flow. The solution of the fluid dynamics problem is usually accompanied by the solution of the aeroelastic problem to obtain the structural deflections, stability limits and flutter speeds. The application of the nonlinear aerodynamic theory has led to the creation of new transonic flutter analysis approaches [4-6]. One of such approaches in which the aerodynamic loads are determined using a

Godunov finite-difference method has been in use at TsAGI for many years [7-9]. A mathematical model for the elastic deformation of structures in transonic flow was developed both in the frequency and time domain. Unsteady aerodynamic coefficients are determined by solving the nonlinear Euler equations for the specified distribution of the vibration amplitudes. Conventional linear flutter investigations in the frequency domain are supplemented with an iterative procedure. As an initial approximation of the frequency and deformation shape the results of linear flutter problem or experimental data are used.

Another more recent method was developed on a basis of a finite difference solution of the small disturbance unsteady Euler equations with a viscosity model. This numerical method models both flow physical and geometrical complexities in more detail than the standard DLM approach. It belongs to the class of so called time linearized frequency domain Computational Fluid Dynamics (CFD) approaches, mainly for unsteady Euler and RANS equations. The aim was to introduce small-disturbance Euler equations that are well-fitted for aeroelastic analyses in the industrial processes and to compare its result with standard potential theory methods like the DLM. In the current paper the examples of aeroelastic applications of developed in TsAGI approaches based on CFD codes.

2 Airfoil oscillations in transonic flow

The stability of the airfoil response in transonic inviscid flow has been investigated with use of a Godunov finite difference method. Pure pitch harmonically oscillations about axis of rotation x_0 were studied for symmetrical airfoil: $\alpha = \alpha_0 Sin\omega_\alpha t$. The airfoil shape is defined by formula: $y = \pm 2c\,x(1-x)$, where *c*-relative thickness of the profile. The airfoil chord *b*=1.

For nonlinear dependence of aerodynamic forces on time an aerodynamic work done by the flow on the structure over one period of oscillation can servers as a stability criteria of the vibration process.

$$\Delta\bar{A}(t) = \frac{\int_0^b \int_t^{t+T} p(t,x)(x-x_0)\,\dot{\alpha}\,dx\,dt}{\pi q \alpha_0^2 \omega_\alpha}$$

The coefficient of aerodynamic damping was defined by the formula: $m_{z_1}^{\dot{\alpha}} = \dfrac{\Delta A}{\pi q \alpha_0^2 \omega_\alpha}$.

In order to determine an influence of different airfoil parts on stability character it is convenient to consider the value of the aerodynamic work on the unit area in point *x* of the chord: $A(x) = \int_0^T p(t,x)(x-x_0)\,\dot{\alpha}\,dt$, and distribution of damping coefficient on the chord: $m_{z_1}^{\dot{\alpha}}(x) = \dfrac{A(x)}{\pi q \alpha_0^2 \omega_\alpha}$.

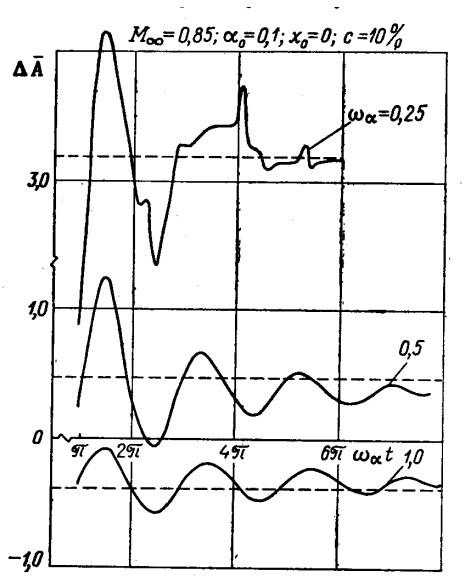


Fig. 1 Time process of convergence of non-dimensional aerodynamic work for three oscillation frequencies

Periodical numerically converged solutions have been obtained for a determination of unsteady aerodynamic coefficients. Process of the convergence depends on oscillation frequencies, and the average time is three or four periods of oscillations. The aerodynamic work over one period is constant for the converged solution. Figure 1 shows the process of the non-dimensional work convergence for three values of oscillations frequency.

Shock waves displacements on upper and lower surfaces of the airfoil are typical for transonic flow over oscillations at specific Mach numbers M_∞ and frequencies ω_α. These displacements cause significant pressure changes in the corresponding airfoil points. Computed non-dimensional pressure in three points on the lower surface (solid line) and on the upper surface (dotted line) as time functions are presented in Fig. 2.

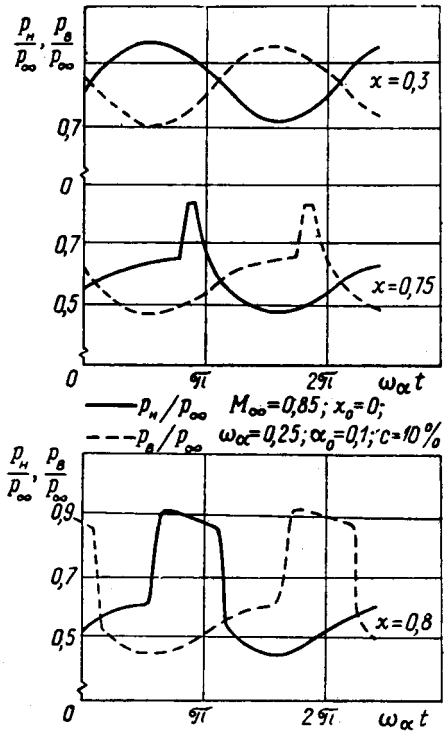


Fig. 2 Computed pressure coefficients in three points (*x*=0.3, 0.75 and 0.8) on the upper ($\frac{p_B}{p_\infty}$) and lower ($\frac{p_H}{p_\infty}$) surfaces of airfoil over one period of oscillations

Numerical results show that all flow parameters, namely the pressure, in points far from shock motion are smooth harmonically functions, and therefore may be specified by an amplitude and phase of oscillations. To determine a role of shock waves motion in stability process it is convenient to introduce the following schematic presentation of the pressure change in different points of the surface: pressure functions may be presented as a sum of two functions – smooth harmonically and stepwise:

$$p_l(t,x) = -p_0(x)Sin[\omega_\alpha t + \varphi(x)] + \Delta p_l(t,x)$$

$$p_u(t,x) = p_0(x)Sin[\omega_\alpha t + \varphi(x)] + \Delta p_u(t,x),$$

where $\Delta p_l(t,x)$ and $\Delta p_u(t,x)$ - piecewise-constant functions which may be equal to zero or a pressure jump across the shock.

Such representation describing the computed pressure change in surface points over the period of oscillations is equivalent of physical values $\Delta p_l(t,x)$ and $\Delta p_u(t,x)$. Analogous suggestion about a character of pressure function in the surface points over oscillations was done in Reference 10.

Figure 3 shows computed pressure jump in points of the lower surface over one period of oscillations.

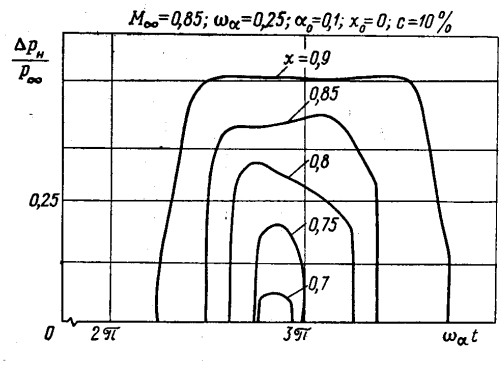


Fig. 3 Computed pressure jumps $\frac{\Delta p_H}{p_\infty}$ across the shock in points of the lower surface over one period of oscillations

It was suggested that the shock movement is described by harmonically function: $x(t) = \bar{x}_0 + \delta\, Sin(\omega t + \psi)$, where δ-shock wave amplitude, ψ- phase of shock motion with respect to airfoil oscillation, $\bar{x}_0$- mean shock location, ω- shock frequency. Computed results show that the amplitude, phase and mean shock location are functions of Mach number and frequency of the airfoil oscillations (Fig. 4).

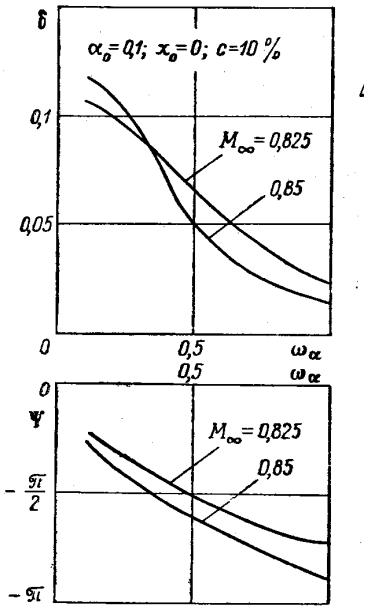


Fig. 4 Computed values of amplitude and phase of shock wave as a function of Mach number and frequency of airfoil oscillations

Function of aerodynamic work may be represented as the sum of two parts:

$$A(x) = A_0(x) + \Delta A(x),$$

where $A_0(x) = 2x\int_0^T p_0(x) Sin[\omega_\alpha t + \varphi(x)]\dot{\alpha}\, dt$,

$\Delta A(x)$ - the aerodynamic work due to the shock movements:

$$\Delta A(x) = x\int_0^T [\Delta p_u(t,x) - \Delta p_l(t,x)]\dot{\alpha}\, dt.$$

The total value of aerodynamic work over one period of oscillations is expressed by the formula:

$$\Delta A = \int_0^b A_0(x)\, dx - 4\alpha_0\, Sin\psi \int_{\bar{x}_0-\delta}^{\bar{x}_0+\delta} x\, \Delta p_x \sqrt{1-(\frac{\bar{x}_0 - x}{\delta})^2}\, dx$$

$$\Delta p_x = \Delta p_l(t,x) - \Delta p_u(t,x)$$

At negative values of the phase ψ (see computational results in Fig.4) the work ΔA is positive. It means that such dynamics of the shock result in negative aerodynamic damping $m_{z_1}^{\dot{\alpha}}$. The range of positive meanings of the damping coefficient $m_{z_1}^{\dot{\alpha}}(x)$ is proportional to the shock amplitude δ.

Figure 5 shows the computed damping coefficient as the function of Mach number and frequency of oscillations. Negative aerodynamic damping corresponds to the narrow range of Mach numbers and to low frequencies.

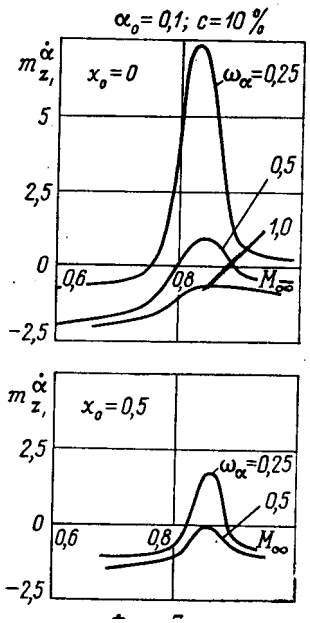


Fig. 5 Aerodynamic damping as function of Mach number and oscillations frequency

The function of the damping coefficient $m_{z_1}^{\dot{\alpha}}(x)$ shows the input that has been contributed by the shock motion in different points of the airfoil surface (see Fig. 6).

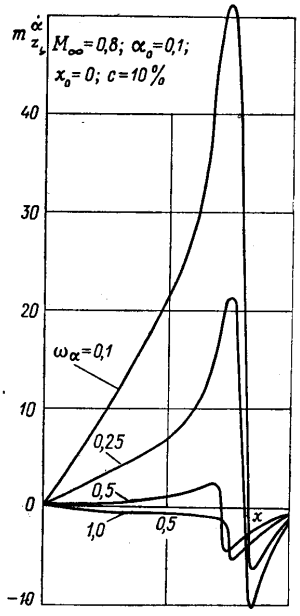


Fig. 6 Computed damping coefficient along the airfoil chord

3 Application of the small disturbance Euler equations for aeroelastic analysis

3.1 Small disturbance Euler method of transonic flow computation

The developed algorithm of time-harmonic characteristics calculation has been constructed on the basis of the small disturbance Euler equations technique used in the Blwf100 solver. The code Blwf100 was created for analysis of transonic flow over complex aerodynamic configuration taking into consideration the viscous effects on the wings including thin separation zones. The possible configurations are: body + arbitrary number of wings attached to a body + arbitrary number of nacelles. An interference of the elements is also modeled (Fig.7).

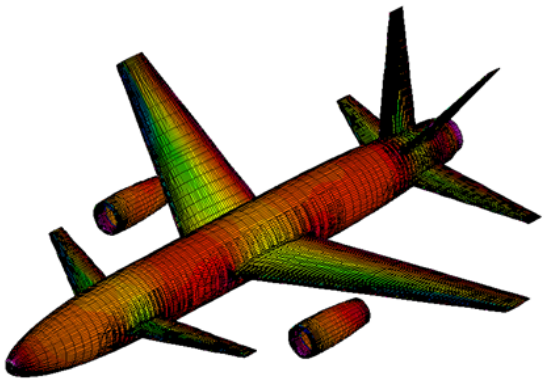


Fig. 7 Airplane configuration

A viscous-inviscid interaction procedure is in the frame of the boundary layer theory. The calculation of the external inviscid steady flow is performed by a numerical integration of a conservative system of the Euler equations by fast implicit method and using the Chimera grid-embedding technique, which simplifies the problem of computational grid generation over complex configuration. The second order finite volume cell centered differentiable Osher type flux difference scheme is used. Effective Newton like implicit solver based on approximate LU decomposition and GMRES algorithm provides very fast convergence (residual diminishing per iteration for the interacted grids is illustrated in Fig. 8).

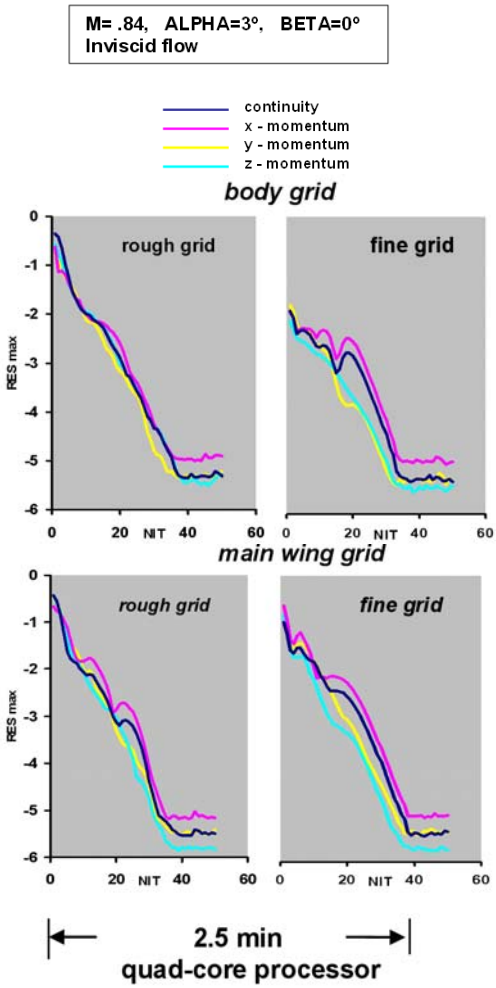


Fig. 8 Residual diminishing for body and wing grids

The laminar and turbulent compressible three-dimensional boundary layers are computed by a finite difference method using a predictor-corrector scheme applied to the Keller formulation. The equilibrium algebraic turbulence model is used. The boundary layer in the separated regions is determined by the inverse procedure. The viscous wing wakes are calculated approximately by the two-dimensional Green's integral method. It allows one to take into account the expected boundary layer response to the chordwise velocity variation and provides an effective and rapid computation of viscid-inviscid interaction. Comparison of computed steady pressure distribution with experimental data for LANN wing (AGARD configuration) is shown in Fig.9. Results are in a good agreement.

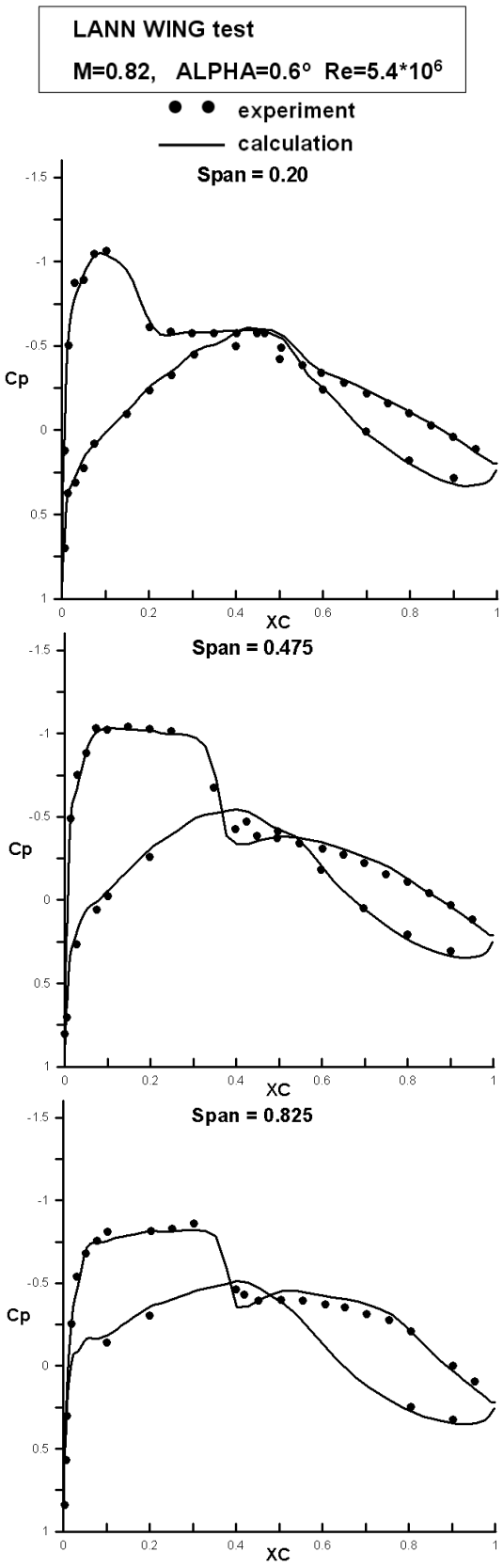


Fig. 9 Comparison of steady pressure distributions for LANN wing

The transonic time-harmonic calculation code (TTHC) has the same structure as BLWF100, uses the same algorithmic elements and can be applied for analyzing the similar complex configuration. The spatial disturbed fluxes are calculated on the basis of Euler fluxes linearization in the assumption that the local entropy does not change. The resulting system is conservative and linear one. The prescribed harmonic oscillations of an airplane are simulated by the corresponding flow transpiration on the undisturbed surface. An additional transpiration models the boundary layer response to the flow disturbance. As the resulting system of finite-difference equations of time-harmonic problem is linear, the solver provides the fast convergence of the iteration process. Unsteady pressure distribution results of computations and experimental data for LANN wing are compared in Fig.10.

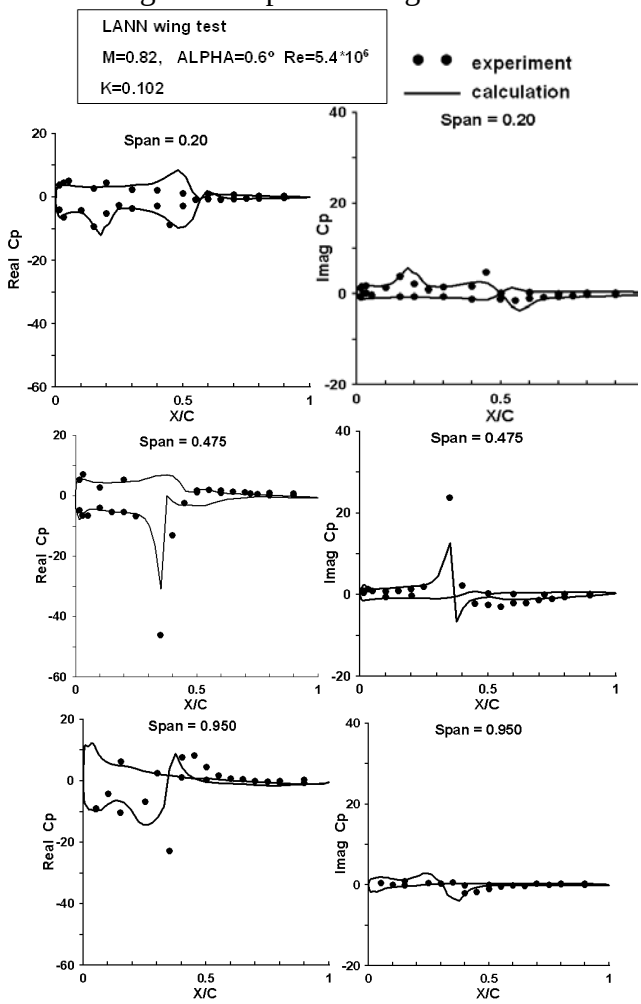


Fig. 10 Comparison of unsteady pressure distributions for LANN wing

Fig.10 shows acceptable correspondence for the case of the wing oscillations in transonic flow with partly separated boundary layer.

One of the main advantages of the solver in the TTHC is its high speed at acceptable accuracy. The computation time for real airplane configuration is approximately 2min (modern quad-core processor PC)

3.2 Computation of modal generalized aerodynamic forces

Equations of elastic airplane oscillations in the frequency domain are presented in the following form:

$$\left(-\omega^2 C + j\omega D_0 + G\right)q = Q \qquad (1)$$

where q is a vector of modal generalized coordinates,

C, D_0, G are matrices of inertia, structural damping and stiffness,

Q is vector of modal generalized forces.

In the case of linear aerodynamics a vector of modal generalized forces can be presented as:

$$Q = -\rho V j\,\omega D - \rho V^2 B \qquad (2)$$

where ρ, V are a density and velocity of flow, D, B are matrices of aerodynamic damping and stiffness, which depend on Mach number and reduced frequency.

It is supposed that modal analysis hase been performed, flutter with linear aerodynamics has been studied, and a set of natural modes, which determine flutter characteristics has been chosen (15-25 modes for symmetrical or antisymmetrical cases).

A set of reduced frequencies is also determined on the basis of linear flutter analysis (in general 5-7 values are sufficient).

Natural mode shapes are input to TTHC as nodal displacements f_x, f_y, f_z of DLM aerodynamic grid. Modal shape is determined by the displacements of 4 corners of each panel for the lifting surfaces (wing, horizontal tail, vertical tail), and by the displacements of the nodes of central line for bodies (fuselage, nacelle) as beam deformations. Torsion of fuselage and nacelle does not result in aerodynamic forces.

Finite-difference solution of linearized unsteady Euler equations is conducted for each mode and each reduced frequency and the complex amplitudes of flow parameter oscillations are determined. Obtained pressure distribution is transformed to the same grid in which modal shapes were specified in order to aeroelasticity analysis could be conducted with the use of the same methods and computational procedure as for linear aerodynamics. The following data are obtained as a result of TTHC execution:

- A distribution of real and imaginary parts of nondimensional pressure difference in panel nodes – for lifting surfaces
$$\Delta c_{P\,m} = \Delta c^R_{P\,m}(x,z) + j\,\Delta c^I_{P\,m}(x,z)$$
- A pressure distribution for body is integrated upon the surface of unitary length and is represented as two complex (vertical and horizontal) components
$$\Delta c_{Py\,m} = \Delta c^R_{Py\,m}(x) + j\,\Delta c^I_{Py\,m}(x)$$
$$\Delta c_{Pz\,m} = \Delta c^R_{Pz\,m}(x) + j\,\Delta c^I_{Pz\,m}(x)$$

Such approach is connected with the fact that in general a beam deformation is sufficient for airplane aeroelasticity problems, and local shell deformation are studied in other tasks.

Here m is mode number,

x,z are coordinates in lifting surface plane.

For aeroelastic applications modal generalized aerodynamic forces are determined using computed pressure distributions. They are defined by the formula

$$Q_m = \frac{\rho V^2}{2}\iint f_m\ \Delta c_P(x,z)\,dxdz$$

where f_m is the modal component normal to the lifting surface.

Pressure difference Δc_P is represented as a superposition of computed above modal pressure distributions $\Delta c_{P\,n}$:

$$\Delta c_P = \sum_{n=1}^{N} \Delta c_{P\,n}\, q_n\ .$$

The integration is then replaced by sum of mean values of pressure and displacement. The

summation is conducted by different ways for lifting surfaces and bodies:

$$Q_m = \frac{\rho V^2}{2}\sum_{n=1}^{N}\left(\sum_{l=1}^{L_{LS}} f_{ml}^{LS}\Delta c_{Pnl} S_l + \sum_{l=1}^{L_B} f_{myl}^{B}\Delta c_{Pynl}\Delta x_l + \sum_{l=1}^{L_B} f_{mzl}^{B}\Delta c_{Pznl}\Delta x_l\right) q_n$$

Here S_l is an area of panel of lifting surface ($l = 1,\ldots,L_{LS}$),

f_{ml}^{LS} is a displacement component (normal to lifting surface) for *m*-mode and *l*-panel,

Δx_l is a distance between nodes along central line of body ($l = 1,\ldots,L_B$)

f_{myl}^{B}, f_{mzl}^{B} are vertical and horizontal components of displacement of central line of body for *m*-mode and *l*-part of body.

The right part of Eq. (1) in transonic case in linear form as Eq. (2) is obtained by extracting real and imaginary parts of generalized force Qm. Then standard aeroelasticity analysis procedure has been applied as for linear analysis. An algorithm similar to known PK-method is used for flutter calculation.

4 Conclusion

The nonlinear Euler equations provide a good flow model for many transonic calculations, if viscous effects are not significant. The shock locations and nonlinear shock dynamics are modeled accurately [1,2].

One of the main advantages of the developed solver in the transonic time-harmonic code is its high speed at acceptable accuracy. The computation time for real airplane configuration is approximately 2min (modern quad-core processor PC)

The developed approach may be used for transonic aeroelasticity parametric study in airplane design

Acknowledgement

This work was partially sponsored by the ISTC Project No 4035.

References

[1] Voss R., Tichy L., Thormann R., “A ROM based flutter prediction process and its validation with a new reference model”, *IFASD*-2011-036,Paris, 2011.

[2] Bendiksen O., “Review of unsteady transonic aerodynamics: Theory and application”, *Journal of Aerospace sciences*, N7, 2010.

[3] Dowell, E., Edwards, J. & Strganac, T. “Nonlinear Aeroelasticity”, *Journal of Aircraft*, 40(5), 2003, pp. 857-874.

[4] Vidy C., Iatrou M., Weishäupl C., and al. “Development of Aerodynamic Influence Coefficients based on CFD Codes for Aeroelastic Applications”, *IFASD*-2011-065, Paris, 2011.

[5] Palacios, R., Climent, H., Karlsson, A., and Winzell, B., “Assessment of strategies for correcting linear unsteady aerodynamics using CFD or experimental results”, *IFASD*-2001–074, Madrid, 2001.

[6] Lepage A., J-M. David J-M., and Girodroux-Lavigne P., “An overall strategy for the assessment of fluid-structure coupling tools for the prediction of control surface efficiency”, *IFASD*-2011-156, Paris, 2011.

[7] Kuzmina S., Ishmuratov F., and Mosounov V., “Theoretical study of transonic flutter / buzz in the frequency and time domain”, *ICAS*-98-4, 9,2, Melbourne, 1998.

[8] Kuzmina S., Ishmuratov F., and Mosounov V., “Static aeroelasticity analysis in transonic flow”, *ICAS*-2000, Harrogate, 2000.

[9] Kuzmina S., Ishmuratov F., Zichenkov M., and Chedrik V., “Nonlinear Aeroelastic Analysis of Adaptive Passenger Aircraft” *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Material Conference*, the USA, 2007.

[10] Tijdeman H., Zwaan R., “On the prediction of aerodynamic loads on oscillating wings in transonic flow”, *AGARD Report* N 612, 1973.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Using behavior based algorithms for missions of multiple UAVs providing pictures for a three-dimensional visualization

J. Holsten, P. Hartmann and D. Moormann
Institute of Flight System Dynamics RWTH Aachen, Germany

Keywords: *Reinforcement learning, UAV*

Abstract

Even though Unmanned Aerial Vehicles have gained importance during the last decade autonomous planning of complex missions remains a challenge. Within the research project AVIGLE different autonomous mission scenarios for multiple UAVs are investigated. In this contribution behavior based algorithms are used to control UAVs in one of the missions. The algorithm is adapted to account for mission specific control problems and to account for delays between control input and air vehicle movement towards a target position. Simulations are used to show the learning capabilities of one UAV and the learned behaviors are verified on multiple UAVs in simulation.

1 Introduction

Current research in UAVs deals e.g. with aspects of autonomous behaviors, trajectory planning and teaming of multiple UAVs. The research project AVIGLE, funded by the EU and the local government of North-Rhine-Westphalia [1], focuses on the development of a team of UAVs to complete different missions. One mission considered is gathering data for a virtual reality application and another is building dynamical on demand radio networks to provide additional mobile communication capabilities.

In this contribution the implementation of behavior based algorithms for the mission providing pictures for a three-dimensional visualization is described. Within the project a hybrid approach of deterministic planning coupled with autonomous swarming behavior was selected. Basically the guidance structure for the described missions can be split into global planning for all UAVs and local teaming behavior. For global planning different deterministic approaches were analyzed, e.g. by FRAZOLLI ET AL. [2] and LAVALLE [3]. For the individual steering of the UAV various approaches such as formation flight or swarming [4] were investigated. A first implementation covering these topics is given in [5]. The behavior based algorithm introduced by MAES ET AL. [6] offers advantages for unknown tasks due to its adaptability. Generally, MAES algorithm can be classified as reinforcement learning, where behaviors are learned by interacting with the environment. One widely known approach in reinforcement learning is *SARSA – State Action Reward State Action* which is detailed by SUTTON ET AL. [7]. Besides classic control problems to which reinforcement learning can be applied the selected mission requires the algorithm to account for delays. SCHUITEMA ET AL. [8] introduce a *memoryless* learning method to deal with delays. In contrast to this KATSIKOPOLOUS ET AL. [9] work with an extended state space of the learning agent to account for delays. This state space includes actions from the delay interval.

Another method accounting for a constant, beforehand known delay is introduced by WALSH ET AL. [10]. In their method the learning algorithm learns based on the underlying undelayed process and then defines additional control actions accounting for the delay. In this contribution we adapted the algorithm by MAES ET AL. [6] to deal with constant and averaged delay for the special case described.

In addition to the adaption of the behavior based algorithm all mission planning requirements within AVIGLE are detailed briefly. Out of these missions the requirements and necessary presets for the algorithm can be derived. For validation the investigated strategies are analyzed applying a 3-degrees of freedom simulation.

2 Mission specification

Within the AVIGLE project an aerial platform in tiltwing configuration is developed, carrying either a RGB-camera for image acquisition or a LTE- base station for cellular network coverage. The data between the ground station and between the aerial platforms is transmitted via LTE or WLAN. Additional to a flight control board, an embedded PC is installed on board allowing mission planning and mission control.

The mission this contribution focuses on is a three-dimensional reconstruction of a building, part of a city or a general structure. A number of UAVs can be selected for this mission and a volume to be observed has to be specified. Furthermore the desired scanning distance to the surface area to obtain the images can be set.

The pictures taken during the mission have to meet requirements regarding overlap and perspective. At least two pictures from different angular perspectives for each part of the target and an overlap between all pictures are required. Ideally, the pictures are directly transferred to the ground station, only a finite number of pictures is planned to be stored on board. It is also possible to transfer pictures via other UAVs of the swarm, for this purpose serving as relay aerial vehicles. For transmission a minimal distance has to be ensured.

Other missions considered within the project are observing a given area of interest. Possible parameters are image resolution and the interval between the images. Further missions are the supply of additional mobile capabilities and a standard mission to validate the basic functionality of the system. An overview of the missions with possible parameters is given in Table 1.

Table 1: Mission parameters.

ID	Location	Mission	Parameter
# UAV	$(x,y,z)_n$ volume	3D Visualization	- distance - level of detail
# UAV	$(x,y)_n$ area	observe	- resolution - interval - altitude
# UAV	$(x,y)_n$ area	network coverage	- max. coverage? - max throughput? - LTE/WLAN?
# UAV	$(x,y,z)_n$ point	manual path	- speed - loop? - in order? - etc.

3 Adaption of behavior based algorithm

To investigate the behavior of UAVs during the three-dimensional observation the behavior based algorithm by MAES ET AL. [6] was adapted. In the following the basic principles of the algorithm are detailed in Section 3.1. In Section 3.2 the adaption and in Section 3.3 the simulated cases are described.

3.1 Behavior based algorithm

In this contribution only a brief overview on the algorithm is given. Further information can be found in [6].

In the algorithm desired behaviors and perceivable conditions are defined. From positive and negative feedback the system learns when to activate its behaviors and how the perceived conditions influence the relevance and reliability of the behavior. Simplifying assumptions made include the existence of a conjunction of preconditions ensuring a probability of positive feedback between boundaries for all behaviors, as well as immediate feedback.

The algorithm is distributed in a way that each behavior figures out independently whether it should activate itself or not. The probability of activation is based on two parameters, *relevance* and *reliability*. Relevance is determined by monitoring positive and negative feedback over the active states of a behavior. The reliability is a measure for the consistency of achieving feedback and serves as a measure of how good the algorithm has learned. It can be determined using the following relation [6]

$$\min\left(\max\left(\frac{j_P}{j_P+l_P},\frac{l_P}{j_P+l_P}\right),\max\left(\frac{j_N}{j_N+l_N},\frac{l_N}{j_N+l_N}\right)\right) \quad (1)$$

in which index P stands for positive feedback and N for negative feedback, j and l are the number of times positive j, or no positive l feedback happened when the behavior was active. The reliability ranges between zero and one, where one represents a high reliability.

Additionally to the feedback and active states, all condition states are monitored with the received feedback when the behavior was active to define a preset of conditions influencing the outcome of the activated behavior. For each condition monitored by the behavior a set of data is recorded for positive feedback, illustrated in Table 2, and analogue for negative feedback [6].

Table 2: Condition monitoring [6].

	Condition on	Condition off
Positive feedback	n	o
No positive feedback	p	q

The correlation between the condition state and positive or negative feedback is calculated in the following way

$$Corr(P,on)=\frac{n*q-p*o}{\sqrt{(p+q)*(q+o)*(n+o)*(n+p)}} \quad (2)$$

Out of this correlation a set of predefined conditions for each behavior is calculated which influence the probability that a behavior becomes active. To ensure adaptiveness the states monitored are decayed with each learning step.

At each time step the behaviors that are activated are selected probabilistically. The probability that a behavior is activated is composed of its relative relevance, its reliability and the status of the learning process so far. Global parameters influencing the algorithm are

- The strength of the correlation for condition to become precondition,
- Number of time steps a condition is monitored before it either becomes a precondition or is dropped,
- Aimed threshold for reliability,
- Adaptiveness of the behaviors.

These global parameters have to be adjusted for each task manually.

3.2 Adaption of the algorithm

The basic principles of the behavior based algorithm described can be used for the mission of three-dimensional visualization if the conditions, behaviors and feedback are selected appropriately. For the first tests the relevance of the behaviors was excluded for the probability calculation since it was assumed, that all implemented and possible behaviors were relevant for positive feedback.

In contrast to the behaviors being relevant for the feedback it was also assumed that not all conditions are relevant for all behaviors. Therefore different types of conditions were introduced. For each behavior it was asserted whether a condition is a hard preset (has to be on or off for the behavior to become active), should be learned or is irrelevant for the behavior and therefore should not be monitored and learned. This distinction also allowed introducing multiple conditions without increasing computational time significantly.

Another important change necessary is the consideration of delays in feedback. Especially for path changing behaviors the effect of the behaviors can only be measured several time steps later. To include this, a delay between the feedback and the conditions monitored was modeled. First a fixed delay of i time steps was con-

sidered linking the feedback from time step t_k to the conditions in time step t_{k-i}. In a later step, to account for changes in the delay, for example due to flight velocity and distance to an area where positive feedback arises, the delay shall be calculated variably, linking it to relevant factors.

Another way to deal with small fluctuations of delay was investigated by averaging the conditions. Assuming that the correct delay between an action and its outcome lies within a and b time steps, where $b>a$, the correlation between feedback, active behavior and condition is calculated using a mean value of the condition states when the behavior was active within the interval $[a,b]$using the relation given in Eq. (3).

$$n_k = n_{k-1} \cdot d^{\frac{\#A}{m}} + FB_k^+ \cdot \frac{1}{m} \sum_{i=a}^{b} (A_{k-i} \wedge C_{k-i}) \quad (3)$$

where d is the decay factor, $\#A$ the number of active states within the interval $[a,b]$, FB^+, A and C logical indicators for positive feedback, whether the behavior was active and whether the condition was on at a certain time step t_k.

3.3 Simulated cases

For an initial analysis a circular flight path around a known area or building was assumed. The simulation and further assumptions are described in more detail in the following section.

Different behavior, feedback and condition sets were selected to investigate the practicability of the described algorithm. Behaviors investigated include a change in radius to increase or decrease the distance towards the target, and a change in height. Additional behaviors such as taking a picture and transferring it will be considered in a following step. Different condition sets were investigated including the distance towards the target and whether a not yet recorded area above, below or straight ahead dominates.

Two different feedback methods were implemented. The first gave positive feedback when the distance towards the target improved. This feedback almost represents a normal control loop since it is directly linked to the observed conditions and possible behaviors. In a second feedback method positive feedback is given when not yet recorded areas are observed. Negative feedback is given when no new areas are detected for the first time after a period of positive feedback. In the case that for a long time no new areas are detected no feedback is given. This avoids giving negative feedback when the UAV actually moves towards not yet recorded areas, but hasn't reached them yet.

4 Simulation and Validation

In this section the simulation and then three validation scenarios are described.

4.1 Simulation

The simulation used models basic flight mechanics of the UAV, its payload and the surrounding area. The flight mechanics modeled include a maximum forward velocity, a correlation between attitude and acceleration and a maximum climb rate. As a payload of the UAV a camera is simulated. The camera position and attitude within the UAV is fixed. Out of the UAV position and attitude as well as the cameras field of view a field of vision can be calculated. The object to be observed is modeled as a hemisphere. Future research will include different object geometries as well. A distance and angle within which suitable pictures can be taken is set within the simulation. During runtime, if the camera observes unseen areas of the object these areas are marked for visualization and stored globally. In a simplified visualization the position of the UAV, the object, field of vision of the camera and observed/not observed areas are shown. Figure 1 illustrates a setup with one UAV, its attitude and position is marked by the cross, its field of vision and a nearly completely unobserved object.

4.2 Validation using distance controlling

For a first validation only one UAV and its distance towards the object was controlled by the learning algorithm.

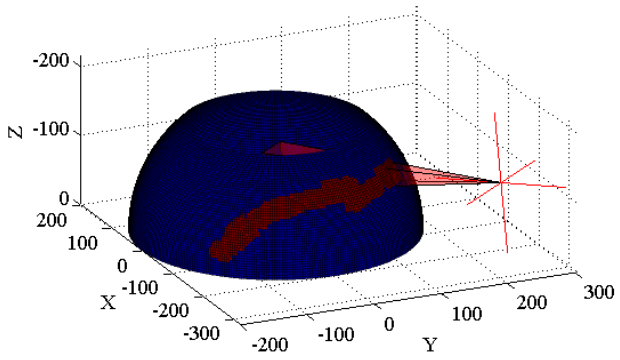


Fig. 1 Visualization of the used simulation.

The height of the UAV was increased continually to ensure a complete observation of the target. Two behaviors were implemented, one to increase the distance towards the target and one to decrease the distance. To ensure that the behaviors are not carried out simultaneously two conditions representing the state of each behavior were included. These conditions were set as hard presets for the other behavior. Conditions monitored by the behaviors are the following:

- Too far away
- Too close
- Undetected area straight ahead
- Undetected area above outweighs below
- Undetected area below outweighs above

Depending on the validation scenario, either the first two or the last three conditions were included in the learning algorithm. All conditions were included in the cases were a combination of the two feedback methods was analyzed.

In this first validation only the distance feedback and the first two conditions with the respective behavior was analyzed to validate the basic usability of the algorithm for the described use case. This scenario was also used to investigate the influence of the in Section 3.1 described global parameters. Figure 2 illustrates the distance for a robust setting of the parameters. The learn steps indicate an update of the statistic data collected and are a multiple of the time step with which the simulation runs. Starting from a distance close to the optimum distance, the UAV immediately learns to decrease its distance when it's too far and increase when it is too near. The fluctuation seen in the distance is caused by the UAV to *try* behaviors even though the conditions do not match. In this case a robust set of parameters is depicted, which leads to small variations in contrast to the case depicted in Figure 5.

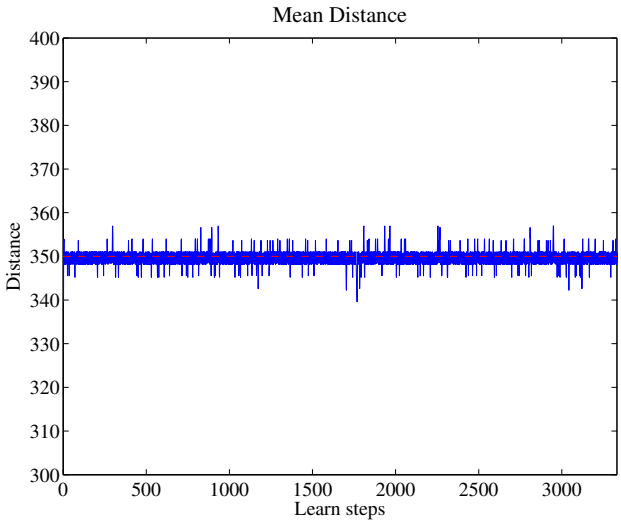


Fig. 2 Controlled distance for a robust setting.

In addition to the first very robust set of parameters, a more adaptive set was implemented and tested by switching the behaviors and observing whether the UAV relearned the right condition correlation. The learning process is in the beginning similar to the process of the robust case. It is visualized in Figures 3 and 4 by the relation between active behavior and the respective conditions.

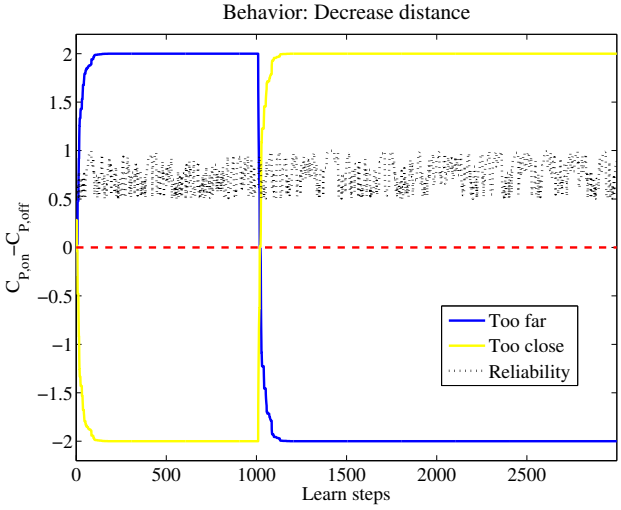


Fig. 3 Correlation of the decreasing distance behavior with the conditions and adaption.

On the y-axis the difference between positive and negative correlation for each condition is depicted. It has to be noted that the data is only

updated at times when the behavior is active. The actual behavior of the UAV in contrast to the learned behavior was switched after 1000 learn steps. It can be seen that before this switch, the UAV learned to decrease its distance when the condition too far was true. After switching the behavior the algorithm adapted rapidly to the change. To ensure this the decay was increased in contrast to the first case and the probability for the UAV to *try* a behavior even though the learned conditions do not correlate with that behavior was increased as well. This led to a higher fluctuation around the optimum distance, as depicted in Figure 5.

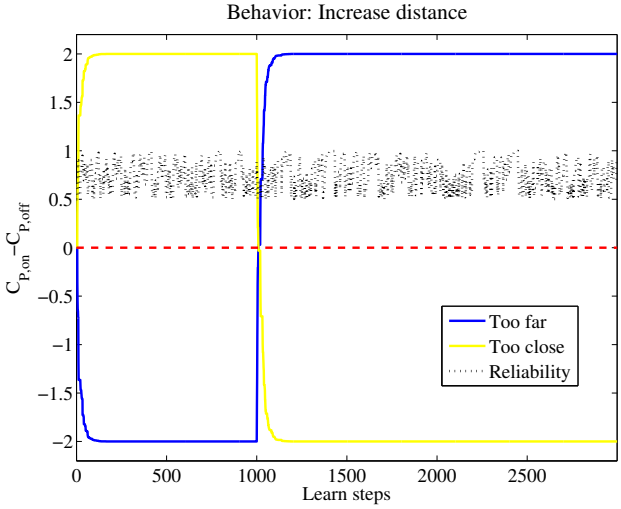


Fig. 4 Correlation of the increasing distance behavior with the conditions and adaption.

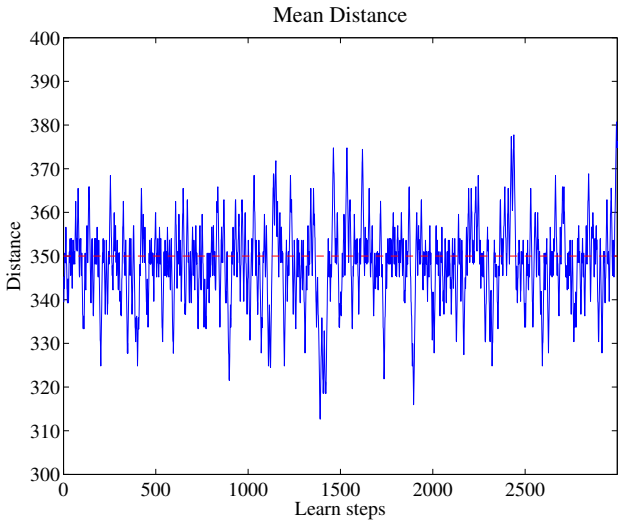


Fig. 5 Controlled distance for an adaptive setting.

4.3 Validation using height control

A second validation was completed using the height control as learning behavior. The beforehand learned behavior for the distance was activated as well, using the learned conditions. As feedback the second method was used. For a first learning large parts of the target were marked as already seen, to achieve a height margin for the UAV to learn.

Figure 6 depicts the flight path of the UAV during the learning process. The learned correlation between the conditions and activating behavior are depicted in Figures 7 and 8.

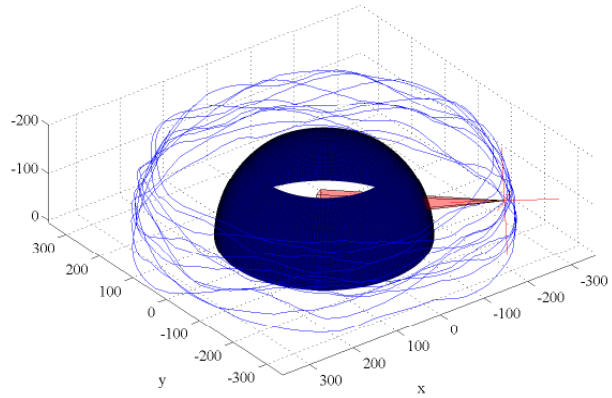


Fig. 6 Flight path for the height learning process.

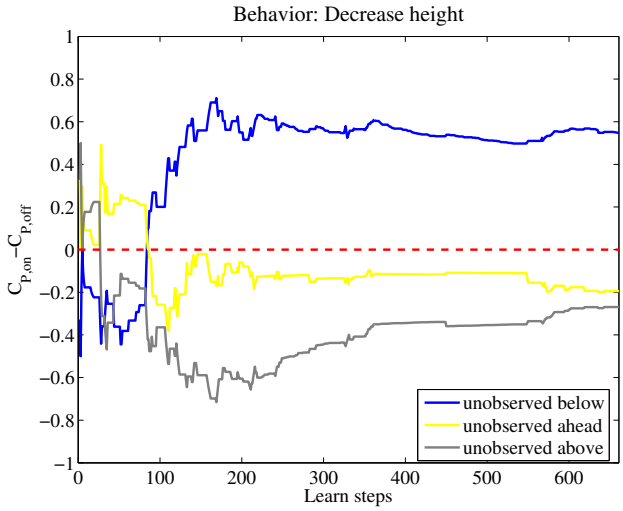


Fig. 7 Correlation of the decreasing height behavior with the respective conditions.

As can be seen both behaviors, increase height and decrease height, have a negative correlation with the condition *unobserved area ahead*. This indicates that the UAV has learned to fly straight ahead as long as there are unobserved areas. Furthermore each behavior correlates positively with its respective condition.

This indicates that the learning algorithm has well adapted, since in this validation scenario the feedback is not directly related to the behaviors and conditions.

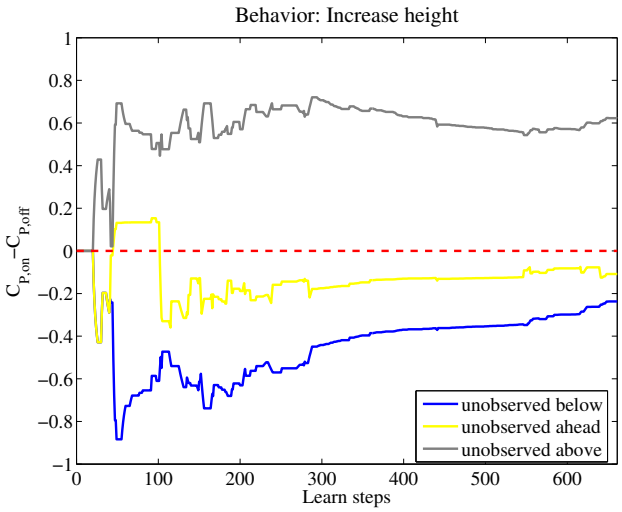


Fig. 8 Correlation of the increasing height behavior with the respective conditions.

4.4 Applying learned behavior for multiple UAVs

As last scenario described in this contribution a case where the learned height control and radius control are given to two UAVs recording a new obstacle is presented. Two cases were investigated. In the first case one UAV starts at the top, the other at the bottom. In the second case both start at the bottom only at different positions. The observed areas are stored globally with the target, which represents the planned real mission scenario.

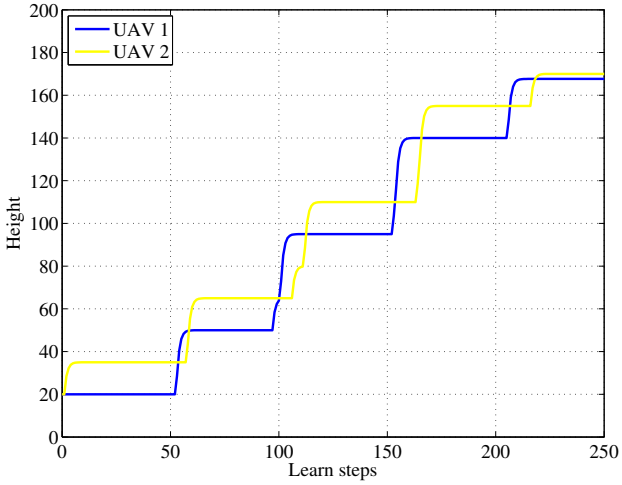


Fig. 9 Height of the two UAVs along their flight path around the object.

Figure 9 illustrates the height of the two UAVs. It can be seen that each UAV observes its own height and recognizes the areas already observed/heights already covered by the other UAV. Figure 10 shows the position and field of view of the two UAVs applying the learned behaviors. The gaps in the area are covered by the two UAVs after reaching the top of the object. The size of the gaps and the climb rate of the UAVs can be tuned by tuning the parameters of the learning process.

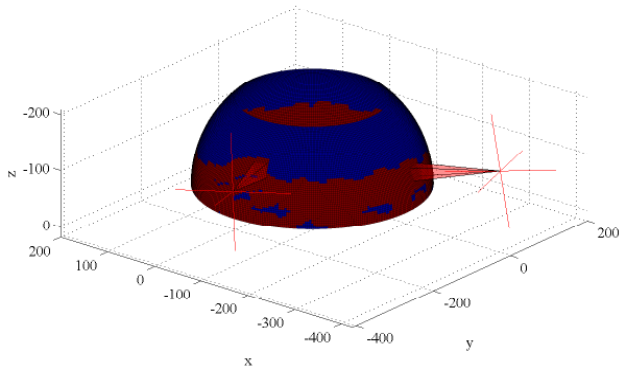


Fig. 10 Two UAVs applying learned behavior.

5 Conclusions and future work

The behavior based algorithm was successfully adapted for a three-dimensional visualization mission. The simulated results show, that the parameters of the learning algorithm have to be carefully adjusted to each scenario. Without further adaption the risk of failure in unknown environment is too high to use this algorithm for inflight learning. For analysis of the correlation between the behavior and the state of the UAV, the target and completion of the mission, the algorithm is well suited. Resulting from these correlations predefined conditions for behaviors can be implemented and thus lead to a deterministic flight path planning. Within a simulation the predefined conditions can be applied to multiple UAVs and the consequent behavior can be determined.

The results of this study will be implemented in the flight guidance system of AVIGLE. Further investigations in extending the learned conditions and behaviors are planned.

Acknowledgement

Our work has been conducted within the AVIGLE project, which is part of the Hightech.NRW research program funded by the German Ministry for Innovation, Science and Research in North Rhine-Westphalia and the EU. AVIGLE is conducted in cooperation with several industrial and academic partners. We thank all project partners for their work and contributions to the AVIGLE project.

Abbreviations and Nomenclature

AVIGLE	EU and NRW funded research project
UAV	Unmanned Aerial Vehicle
SARSA	State Action Reward State Action
RGB	Red Green Blue
LTE	Long Term Evolution
WLAN	Wireless Local Area Network
#A	Number of active states
A	Logical Indicator for active state
C	Logical Indicator for Condition on
Corr	Correlation
FB^+	Logical indicator for positive feedback
N	Negative
P	Positive
a	Interval boundary
b	Interval boundary
d	Decay factor
j	Decayed counter for positive feedback with active behavior
l	Decayed counter for negative feedback with active behavior
n	Decayed counter for condition on and positive feedback
o	Decayed counter for condition off and positive feedback
p	Decayed counter for condition on and no positive feedback
q	Decayed counter for condition off and no positive feedback
t_k	Time step

References

[1] Rohde, S., Goddemeier, N., Wietfeld, C., Steinicke, F., Hinrichs, K., Ostermann, T., Holsten, J. and Moormann, D., "AVIGLE: A System of Systems Concept for an Avionic Digital Service Platform Based on Micro Unmanned Aerial Vehicles", In *2010 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Istanbul, Turkey, IEEE, Oct 2010

[2] Frazolli, E., Dahleh, M. A. and Feron,E., „Real-time motion planning for agile autonomous vehicles," *AIAA J. Guid., Control, Dynam.*, Vol 25, No. 1, 2002, pp. 116-129

[3] LaValle, S. M., "Rapidly-exploring random trees: A new tool for path planning." Technical Report 98-11.Iowa State University, Ames, IA, 1998

[4] Reynolds, C., "Flocks, herds, and schools: A distributed behavioural model," 1987

[5] Holsten, J., Loechelt, S., Ostermann, T., Moormann, D., "Approach to swarming behavior and deterministic planning considering motion models", In *International Micro Air Vehicle Conference,* Braunschweig, Germany, 2010

[6] Maes, P. and Brooks, R. A., "Learning to Coordinate Behaviors," In *Proceedings Eighth National Conference on Artificial Intelligence*, 1990, pp. 796-802

[7] Sutton, R.S., Barto, A.G., *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, 1998

[8] E. Schuitema, L. Busoniu, R. Babuska, P. Jonker, "Control Delay in Reinforcement Learning for Real-Time Dynamic Systems: A Memoryless Approach.", In *Proceedings 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-10)*, Taipei, Taiwan, 18–22 October 2010, pages 3226–3231

[9] Katsikopoulos, K.V. and Engelbrecht, S.E. "Markov Decision Processes with delays and asynchronous cost collection", *IEEE Transactions on Automatic Control,* Vol. 48, No. 4, 2003, pp 568-574

[10] Walsh, T.J., Nouri, A., Li, L. and Littman, M.L., "Learning and Planning in environments with delayed feedback", *Autonomous Agents and Multi-Agent Systems,* Vol. 18, No. 1, 2005, pp. 83-105

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Prediction of complete helicopter aerodynamics at low- and high-speed conditions

M. Biava
AgustaWestland-Politecnico Advanced Rotorcraft Center, Milano, Italy,
e-mail: biava.awparc@agustawestland.com

L. Vigevano
Dipartimento di Ingegneria Aerospaziale - Politecnico di Milano, Milano, Italy,
e-mail: luigi.vigevano@polimi.it

Abstract

Complete helicopter simulations gathered with the *ROSITA* RANS solver are be presented and compared to experimental measurements for two significant flight conditions. The first condition refers to high-speed, level flight, in which compressibility effects are important. The second condition corresponds to low-speed, descent flight, in which interactional aerodynamics plays a crucial role. Discussion of results allows to underline the flow physics and to better understand the measurements. The agreement between computations and experiments is rather good for the main rotor and the front part of the fuselage, less satisfactory in the turbulent wake region.

1 Introduction

The flow field past a complete helicopter configuration is three-dimensional, highly unsteady and features a variety of complex physical phenomena, like compressible flow with shocks at the advancing blade, separated flow in the hub region, complex vortex wakes due to the main and tail rotors. It represents a true challenge to prediction methods, although the capabilities of CFD tools to predict these complex features is being demonstrated in recent years [1],[2],[3], even considering fluid-structure coupling [4],[5]. The numerical methods still however need to be thoroughly validated for this class of flows and this calls for the availability of a suitable experimental database. One such important experimental database has been generated during the EU project GOAHEAD [6], [7], in which a complete helicopter model was tested at the German-Dutch low speed wind tunnel (DNW) in spring 2008 and which will be used as a benchmark for CFD methods validation [8],[9]. In addition to the global aerodynamic forces and moments, the GOAHEAD test campaign included detailed steady and unsteady surface pressure measurements on the helicopter and tunnel walls, transition locations, inflow velocity and turbulence kinetic energy profiles and PIV measurements for a wide array of flight and rotor loading conditions. As a contribution to the validation of unsteady, Reynolds Averaged Navier–Stokes (RANS) prediction methods, in this paper numerical computations are presented and compared to experimental measurements for two significant flight conditions. The first condition refers to high-speed, level flight, in which compressibility effects are expected to be important. The second condition corresponds to low-speed, descent flight, in which interactional aerodynamics plays

a crucial role. The numerical results were obtained via time-accurate solution of RANS equations in three dimensions using the finite volume, block-structured, parallel solver *ROSITA*, developed at Politecnico di Milano [10],[11], which uses the Chimera overlapping grid approach to introduce the motion of the blades.

The paper is organized as follows: section 2 summarizes the main characteristic of the CFD solver, while section 3 describes the examined configuration and flow conditions and gives some details on the grids used. Numerical results are discussed and compared with experimental data in section 4 and the conclusions of the study are reported in the last section.

2 The flow solver *ROSITA*

The *ROSITA* flow solver [12] numerically integrates the RANS equations, coupled with the one-equation turbulence model of Spalart–Allmaras [13], in systems of moving, overset, multi-block grids (figure 1). The equations are discretized in space by means of a cell-centred finite-volume implementation of the Roe's scheme [14]. Second order accuracy is obtained through the use of MUSCL extrapolation supplemented with a modified version of the Van Albada limiter introduced by Venkatakrishnan [15]. The viscous terms are computed by the application of the Gauss theorem and using a cell-centred discretization scheme. Time advancement is carried out with a dual-time formulation [16], employing a 2nd order backward differentiation formula to approximate the time derivative and a fully unfactored implicit scheme in pseudo-time. The generalized conjugate gradient (GCG), in conjunction with a block incomplete lower-upper preconditioner, is used to solve the resulting linear system.

The connectivity between the (possibly moving) component grids is computed by means of the Chimera technique. The approach adopted in *ROSITA* is derived from that originally proposed by Chesshire and Henshaw [17], with modifications to further improve robustness and performance. The domain boundaries with solid wall conditions are firstly identified and all points in overlapping grids that fall close to these boundaries are marked as holes (seed points). Then, an iterative algorithm identifies the donor and fringe points and lets the hole points grow from the seeds until they fill entirely the regions outside the computational domain. To speed up the search of donor points, oct-tree and ADT (alternating digital tree) data structures are employed. For integration of the aerodynamic forces on overlapping surface grids (figure 2), the so-called "zipper-grid" technique of Chan and Buning [18] is used. It consists in eliminating the overlapped surface cells and in filling the created gap with triangles; the integration is then performed on the resulting hybrid mesh.

The *ROSITA* solver is fully capable of running in parallel on computing clusters. The parallel algorithm is based on the message passing programming paradigm and the parallelization strategy consists in distributing the grid blocks among the available processors. Each grid block can be automatically subdivided into smaller blocks by the CFD solver to attain an optimal load balancing.

3 Numerical grids and flow conditions

The computational model refers to the tested GOAHEAD scaled model, featuring a 4.1 m NH90 fuselage, ONERA 7AD main rotor, reduced scale BO105 tail rotor, rotor hub and model support, and includes the 8×6 m test section of the DNW low-speed wind tunnel (fig. 3, where also the location of pressure sensors is reported for later reference). Several multi-block grids were generated around the different components of the configuration (fuselage, rotor blades, etc.) to form a complete Chimera system. The same grids were used for all test cases, since the Chimera method allow for any arbitrary orientation and relative motion of the individual grids within the system, provided sufficient overlap is assured. The dimensions of the grids are given in table 1. The total number of cells is about 27 million.

Two test cases were selected from the GOAHEAD database: TC2, which refers to a low-speed, pitch-up condition, and TC6, which features the largest speed among the measured conditions. The corresponding angle of attack of the fuselage and Mach numbers are reported in ta-

Grid	No. Blocks	No. of Nodes ($\times 10^6$)
Fuselage	79	17.4
Main rotor blade ($\times 4$)	4	1.0
Tail rotor blade ($\times 2$)	6	0.5
Rotor hub	9	2.0
Strut	4	1.3
Wind tunnel	6	1.3
Total	126	27.0

Table 1: Grid dimensions

	Fuselage incidence	Wind tunnel Mach N.	Main rotor tip Mach N.	Tail rotor tip Mach N
TC2	1.9	0.059	0.617	0.563
TC6	-1.0	0.280	0.617	0.563

Table 2: Flow conditions for the two cases considered

ble 2. Both the main and tail rotor blades have been considered rigid, with prescribed kinematics derived from measured data. The main rotor is rotating clockwise when observed from above. The upper vertical tail rotor blade is advancing and the lower vertical blade is retreating.

The wind tunnel walls have been modelled with inviscid boundary conditions, whereas the fuselage, the strut, the main rotor and the tail rotor where modelled through no-slip boundary conditions. The time-accurate simulation was started from an undisturbed free stream condition (*i.e.* impulsive start) and four main rotor revolutions have been performed using a constant azimuthal step size $\Delta\Psi = 1°$. The number of pseudo-iterations for each azimuthal step was equal to 50, with CFL=4.0 for the mean flow and CFL=1.0 for the Spalart–Allmaras equation. Calculations were run on the Lagrange cluster at CILEA, using 240 processors; CPU time per revolution was 47 hours.

4 Discussion of results

The complexity of the flow field around the complete helicopter can be appreciated from a qualitative visualization, using the Q-criterion [19], of the rotor and hub wakes as shown in figure 4. In the low-speed case TC2 the main rotor wake stays close to the disk plane and merges to form a pair of large trailing vortices. Also in the TC6 case, given the low negative incidence of the fuselage, the rotor wake is close to the blades, the main differences being observed in the retreating blade region, where the blade near wake is blown by the forward speed on the blade itself. It has to be noted that the turbulent hub wake extensively interacts with the main rotor and the fuselage fins.

4.1 Rotor loads

We start the analysis of the local rotor loads from the advancing blade range. Figures 5 and 6 show the blade chordwise pressure distributions, corrected for the rotation velocity, at two outboard spanwise stations, respectively at $r/R = 0.915$ and $r/R = 0.975$, where compressibility effects are to be expected. Quite obviously, compressibility effects are limited for the TC2 case, being the local tip Mach number (the free-stream Mach number due to rotation and forward motion) equal to 0.676. For case TC6 the local tip Mach number is as high as 0.897, but the parabolic leading edge sweep at the blade tip reduces the influence of compressibility. Some indication of the formation of a shock wave can be observed for TC6 at $\Psi = 90°$ and $\Psi = 120°$. It is however not easy to detect a clear trace of a shock from the surface pressure distributions shown in figure 7.

The flow on the retreating blade range is dominated by large angles of attack. Since the thrust level of the considered cases is not very high, it

does not seems from the pressure distribution of figure 8 that flow separation occurs. Some experimental data are not reliable in figure 8(b). Overall, the agreement between computed and experimental distributions is fairly good along a complete blade revolution for both test cases, suggesting that the numerical methods are adequate for rotor loads calculations.

4.2 Fuselage loads

A snapshot of the pressure distribution on the fuselage symmetry axis at $\Psi = 180°$, when two of the blades are aligned with the centerline, is shown in figure 9. A fairly good agreement between computations and experiments is observed in the front part of the fuselage (the origin of the X axis of the figure is at the fuselage nose), but large discrepancies are found in the aft part of the centerline, over the tail boom and on the vertical fin leading edge, where the turbulent hub wake impinges the fuselage. Time histories of pressure in the front part of the fuselage (figures 10 and 11) confirm the quality of the simulations in this region. There is a strong influence of the blade passage for both test cases, with a dominating 4/rev frequency. In addition, the averaged values of pressure are higher at sensor A12 on the advancing blade side than at sensor A13 (figure 11). From experiments in the TC2 case, the sensor A06, located closer to the fuselage nose, features high frequency oscillations that suggest the effect of flow separation or vibration of the model; the latter is not reproduced by the computations. Unfortunately the scale of the plots are different for the two test cases; it has to be noticed that pressure oscillations are quite larger for TC2 case than for TC6, showing a stronger rotor wake interaction and unsteadiness of the flow.

Finally we present in figure 12 the pressure distribution on an outboard section of the horizontal fin (section V3 in figure 3). The distributions are quite different in the two flow conditions, due to the different aircraft attitude. The agreement between experiments and computations is rather good in both cases, notwithstanding that for TC6 the numerical results somewhat fail to reproduce the correct variation of the peak pressure with azimuth.

5 Conclusions

Simulations of two test cases from the GOAHEAD database, the low-speed TC2 and the high-speed TC6, are presented and compared with experimental data. The analysis of the achieved results, although still limited, yield a better understanding of the physics of the flow field around a complete helicopter in different operating conditions and allows for a deeper interpretation of the measurements. The agreement between computations and experiments is rather good for the main rotor and the front part of the fuselage, less satisfactory in the turbulent wake region. Taking into account that the present simulations cannot be considered fully resolved in space, the level of agreement that has been reached can be judged satisfactory for engineering purposes.

Acknowledgements The present work has been partially supported by Regione Lombardia and CILEA under the project LISA.

References

[1] T. Renaud, C. Benoit, J.-C. Boniface, and P. Gardarein. Navier–Stokes computations of a complete helicopter configuration accounting for main and tail rotor effects. In *29th European Rotorcraft Forum, Friedrichshafen, Germany*, 2003.

[2] W. Khier, T. Schwarz, and J. Raddatz. Time accurate simulation of the flow around the complete BO105 wind tunnel model. In *31st European Rotorcraft Forum, Florence, Italy*, 2005.

[3] W. Khier, M. Biava, and L. Vigevano. Cfd prediction of air flow past a full helicopter configuration. In *36st European Rotorcraft Forum, Paris, France*, 2010.

[4] W. Khier, M. Dietz, T. Schwarz, and S. Wagner. Trimmed CFD simulation of a complete helicopter configuration. In *33rd European Rotorcraft Forum, Kazan, Russia*, 2007.

[5] M. Dietz, W. Khier, S. Wagner, and E. Krämer. Numerical simulation of a full

helicopter configuration using weak fluid-structure coupling. In *46th AIAA Aerospace Science Meeting and Exhibit, Reno, Nevada. AIAA 2008-401*, 2008.

[6] K. Pahlke. The goahead project. In *33rd European Rotorcraft Forum, Kazan, Russia*, 2007.

[7] M. Raffel, F. de Gregorio, W. Sheng, G. Gibertini, A. Seraudie, K. de Groot, and B. van der Wall. Generation of an advanced helicopter experimental aerodynamic database - the wind tunnel test of the EU-Project GOAHEAD. In *35th European Rotorcraft Forum, Hamburg, Germany*, 2009.

[8] O. Boelens, G. Barakos, M. Biava, A. Brocklehurst, M. Costes, A. D'Alascio, M. Dietz, D. Drikakis, J. Ekaterinaris, I. Humby, W. Khier, B. Knutzen, F. Le Chiuton, K. Pahlke, T. Renaud, T. Schwarz, R. Steijl, L. Sudre, L. Vigevano, and B. Zhong. The blind-test activity of the GOAHEAD project. In *33rd European Rotorcraft Forum, Kazan, Russia*, 2007.

[9] W. Khier. Numerical simulation of air flow past a full helicopter configuration. In *35th European Rotorcraft Forum, Hamburg, Germany*, 2009.

[10] M. Biava, J.-C. Boniface, and L. Vigevano. Influence of wind-tunnel walls in helicopter aerodynamics predictions. In *31st European Rotorcraft Forum, Firenze, Italy*, 2005.

[11] M. Biava. *RANS computation of rotor/fuselage unsteady interactional aerodynamics*. PhD thesis, Politecnico di Milano, 2007.

[12] M. Biava, A. Pisoni, A. Saporiti, and L. Vigevano. Efficient rotor aerodynamics predictions with an Euler method. In *29th European Rotorcraft Forum, Friedrichshafen, Germany*, 2003.

[13] P.R. Spalart and S.R. Allmaras. One equation model for aerodynamic flows. In *30th AIAA Aerospace Science Meeting & Exhibit, Reno, NV. AIAA 92-0439*, 1992.

[14] P.L. Roe. Approximate riemann solvers, parameter vectors and difference schemes. *J.Comput. Phys*, 43:357–372, 1981.

[15] V. Venkatakrishnan. On the accuracy of limiters and convergence to steady state solutions. In *31st AIAA Aerospace Science Meeting and Exhibit, Reno, Nevada. AIAA 1993-880*, 1993.

[16] A. Jameson. Time dependent calculations using multigrid with applications to unsteady flows past airfoils and wings. In *10th AIAA Computational Fluid Dynamics Conference, Honolulu, HI. AIAA 91-1596*, 1991.

[17] G. Chesshire and W. D. Henshaw. Composite overlapping meshes for the solution of partial differential equations. *J.Comput. Phys*, 90:1–64, 1990.

[18] W. M. Chan and P. G. Buning. Zipper grids for force and moment computation on overset grids. In *12th AIAA Computational Fluid Dynamics Conference, San Diego, CA. AIAA 1995-1681*, 1995.

[19] J.C.R. Hunt, A. Wray, and P. Moin. Eddies, stream, and convergence zones in turbulent flows. Center for Turbulence Research Report CTR-S88, 1988.

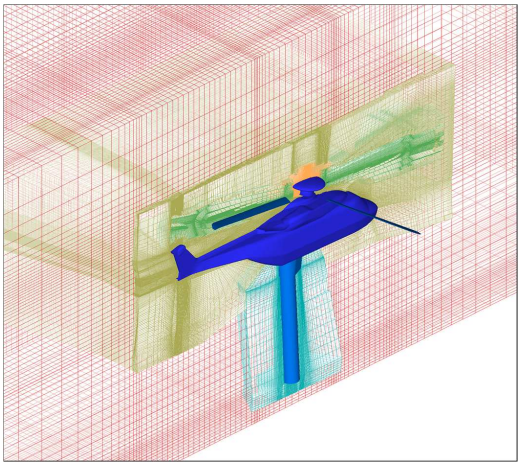


Figure 1: Chimera overset grid system

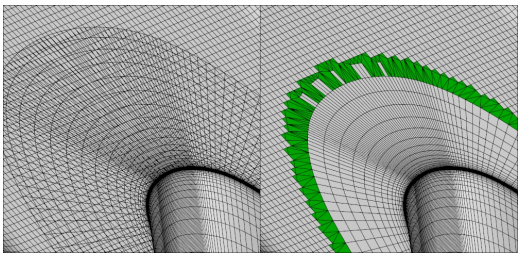
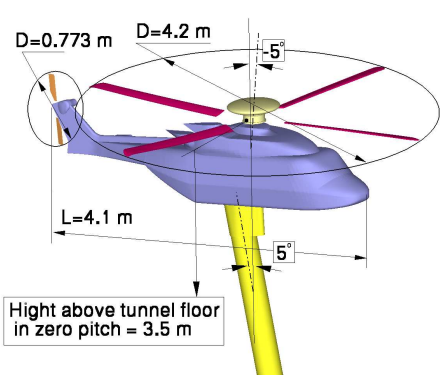
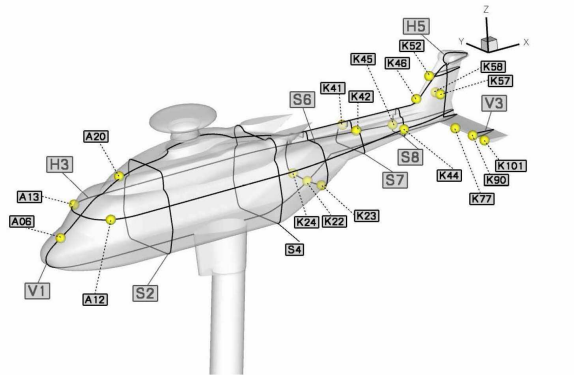


Figure 2: Overlapped surface grids (left) and hybrid surface mesh with zipper grid (right).

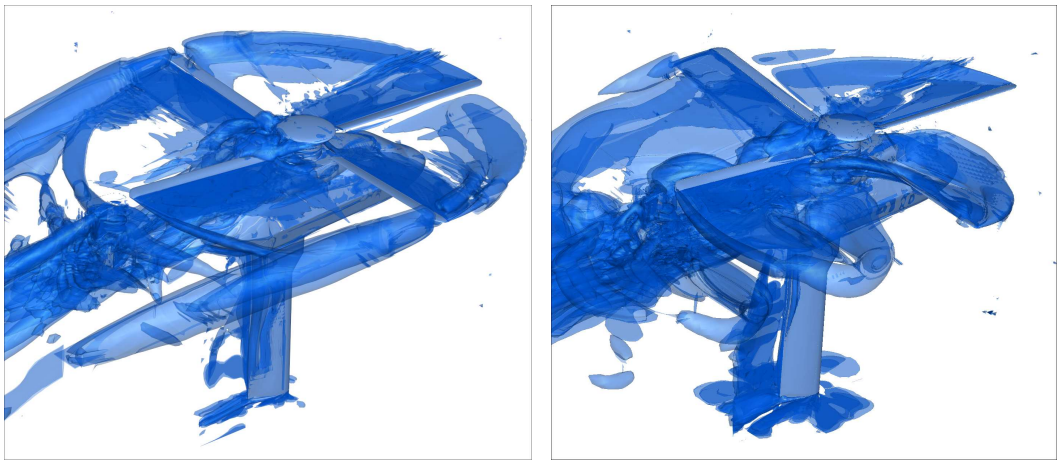


(a) Model dimensions



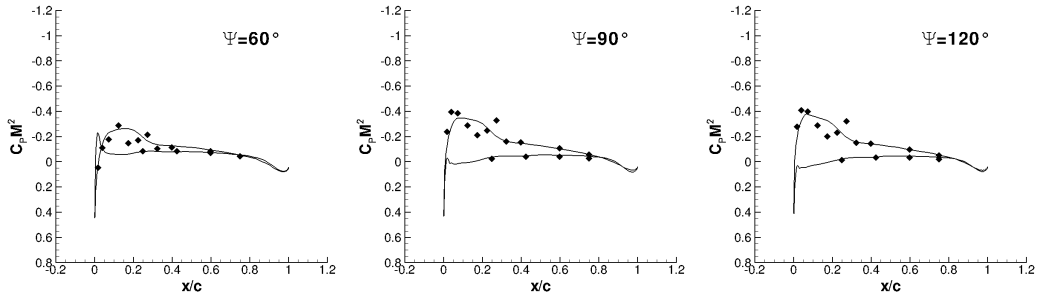
(b) Sensor locations

Figure 3: Complete helicopter model dimensions and location of some selected pressure sensors.

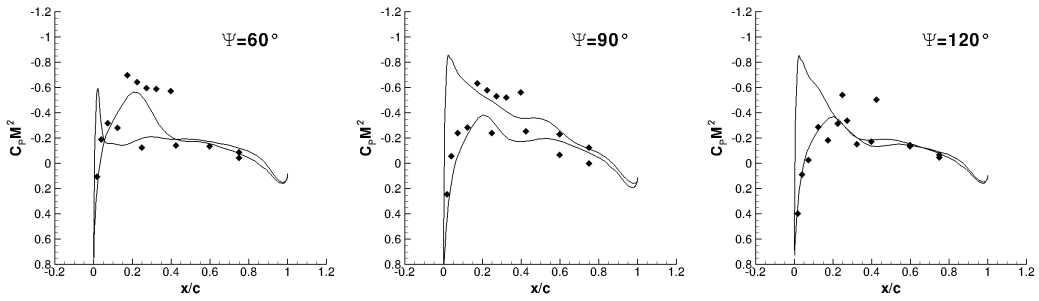


(a) TC2 (b) TC6

Figure 4: Visualization of the flow field with Q-criterion for the two cases.

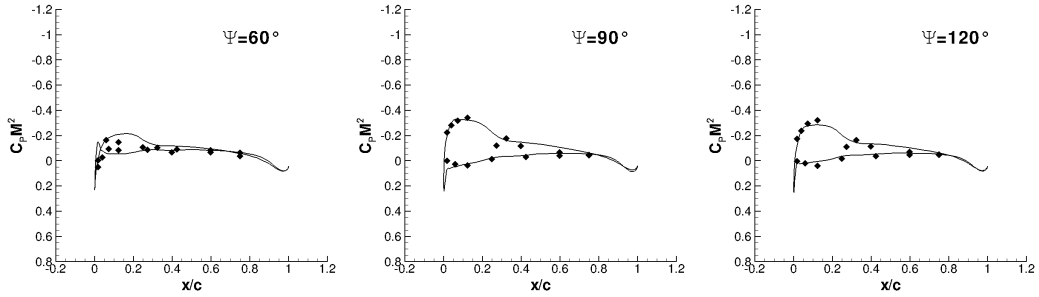


(a) TC2

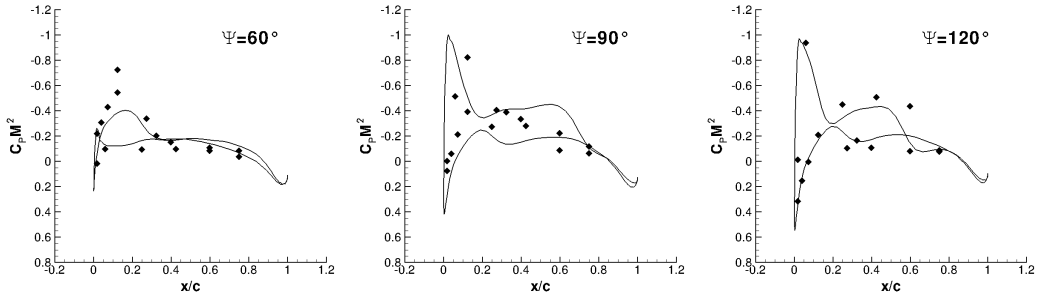


(b) TC6

Figure 5: Advancing blade chordwise pressure distributions at $r/R = 0.915$.

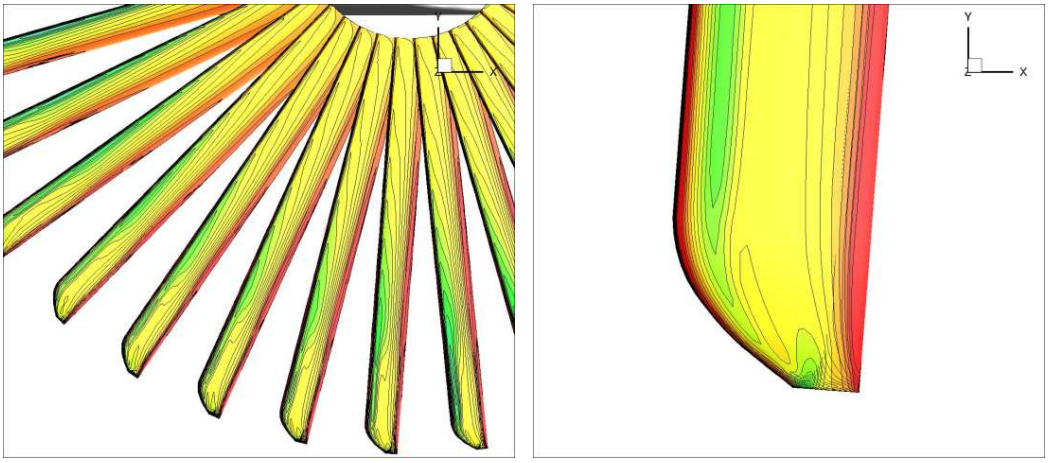


(a) TC2



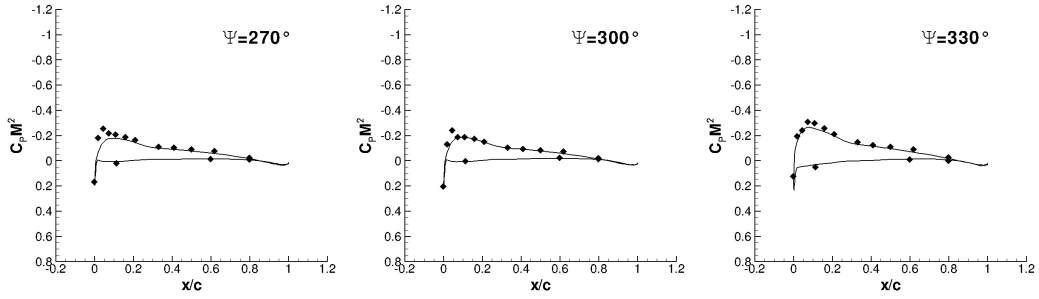
(b) TC6

Figure 6: Advancing blade chordwise pressure distributions at $r/R = 0.975$.

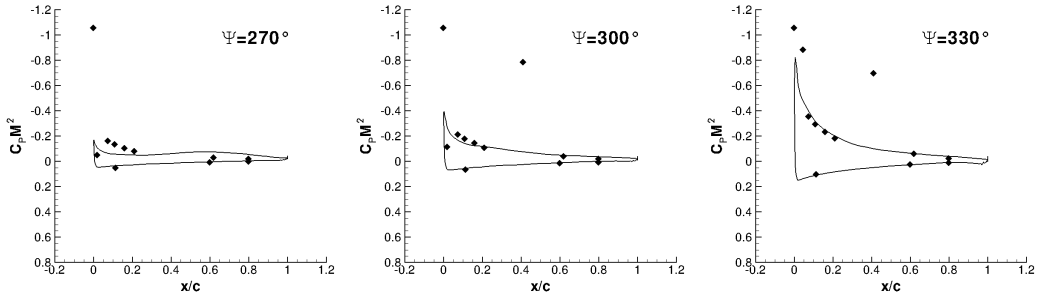


(a) Advancing blade range (b) $\Psi = 90°$

Figure 7: Pressure distribution on the blade upper surface for TC6 case.

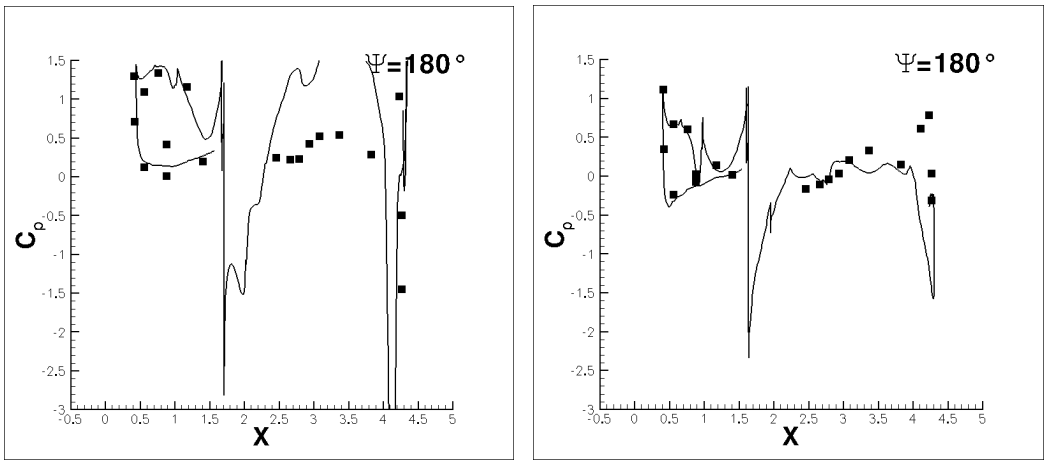


(a) TC2



(b) TC6

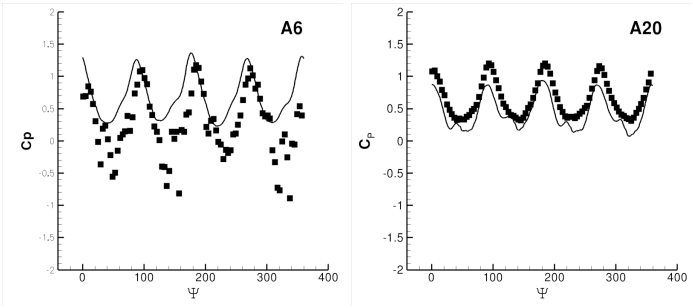
Figure 8: Retreating blade chordwise pressure distributions at $r/R = 0.825$.



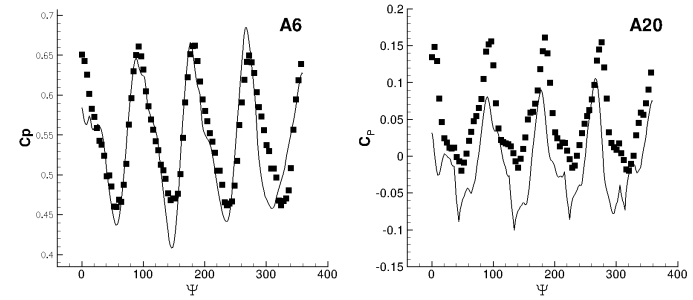
(a) TC2

(b) TC6

Figure 9: Pressure distribution on fuselage symmetry plane at $\Psi = 180°$.

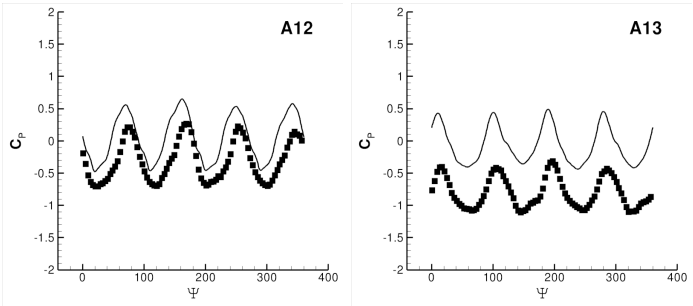


(a) TC2

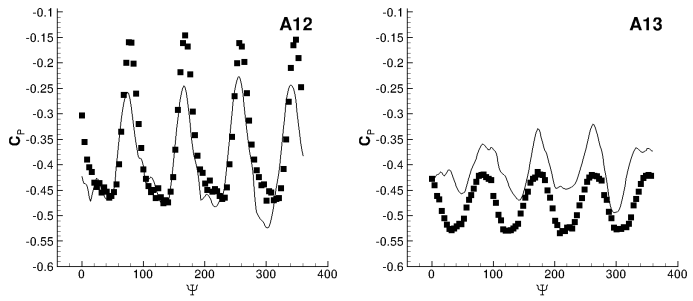


(b) TC6

Figure 10: Unsteady pressure signals on fuselage symmetry plane.

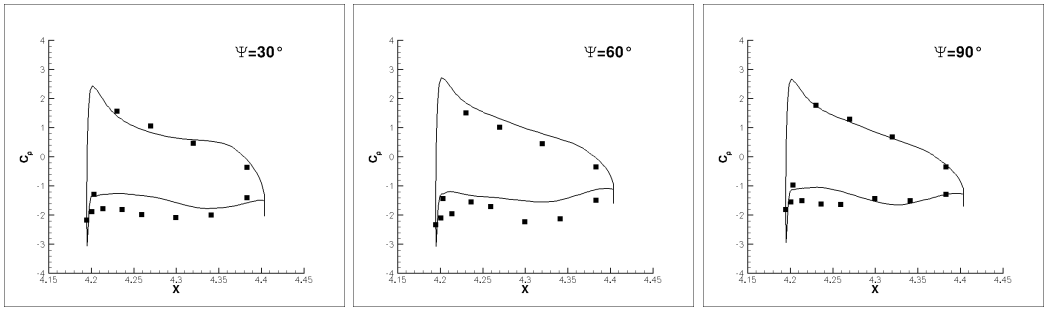


(a) TC2

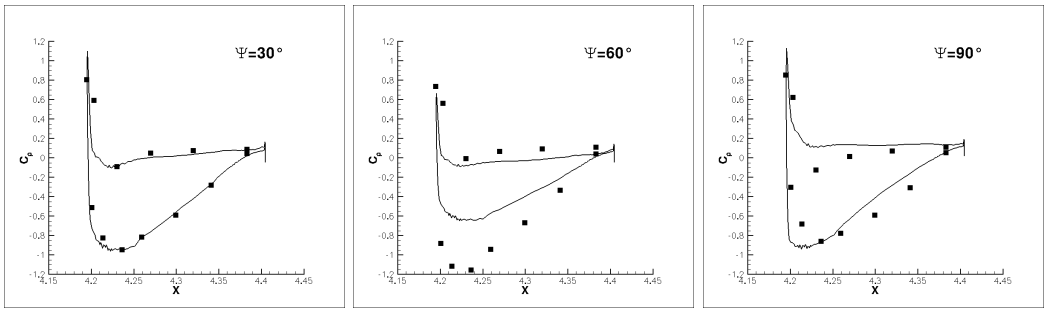


(b) TC6

Figure 11: Unsteady pressure signals on front part of the fuselage.



(a) TC2



(b) TC6

Figure 12: Pressure distribution on the horizontal tail plane.

Stress Analysis of the RASAT Satellite Structure Considering Vibration Induced Loads during Launch

Suat ONTAÇ
TÜBİTAK Space Technologies Research Institute, TURKEY

Assoc.Prof.Dr. Serkan DAĞ, Prof.Dr. Mustafa İlhan GÖKLER
Mechanical Engineering Department, Middle East Technical University, TURKEY

Keywords: *RASAT, micro-satellite, finite element analysis, vibration*

Abstract

The RASAT satellite is the first Earth observation micro-satellite designed and manufactured in TURKEY, which had being developed by TÜBİTAK Space Technologies Research Institute.

The main aim this study was to build a finite element model of the RASAT satellite, make fundamental dynamic analyses like modal analysis, and then by using the results obtained from modal analysis, carrying out a stress analysis on the satellite structure.

By the structural stress analysis, the margins of safety of the satellite structure and components were determined in order to see whether the structural components of the RASAT satellite can withstand the vibration induced launch loads or not. All the critical results were to be tabulated.

The results will show whether the structural design verification regarding the reliability of the structure for the desired mission has been successfully completed or not.

1 Introduction

Micro satellites have a great importance in the satellite industry and several developing countries deal with micro satellite design and production. Turkey is one of these countries by conducting new satellite projects. RASAT project is the first, completed micro-satellite project, which had being developed and manufactured by TÜBİTAK Space Technologies Research Institute. It was successfully launched on August 2011 and is operating in orbit.

Satellites are composed of many sub-systems like structural, thermal, power, altitude determination and control, etc. Structural sub-system must be considered as one of the most important sub-systems because it is simply the skeleton of the whole satellite. During the launch phase, the satellite, especially and normally its structural subsystem will be exposed to extremely high dynamic conditions. In order to see whether the satellite will survive and keep its functionality, satellite structures are being analyzed in design stage and ground tested after production stage.

When designing the satellite structures, couple of analyses is carried out on the finite element model and the dynamic behaviors and

corresponding structural responses of the satellite are obtained. Finite element stress analysis is one of these design-time analyses and used in conclusion on whether the satellite structure design is satisfactory or not.

This study covers the stress analysis of the RASAT satellite structure on which several types of vibration loads are going to be applied from the beginning of the launching action until the beginning of the first launch separation during the orbit landing operation.

First of all, after setting up the finite element model, a modal analysis was conducted for the RASAT structure to identify the natural frequencies of the structure while it is attached to the launch vehicle adapter during launch. After taking the results of modal analysis, they were used to perform structural stress analysis for the RASAT with some special calculations.

For structural stress analysis, quasi-static load factors and random vibration load factors were used. Then they were used to calculate load cases. The load cases were applied to the finite element model of the RASAT in a static case. After applying each load cases, the results of loads were identified and most important ones were used to see the margin of safety of the satellite structure. By the structural stress analysis, the margins of safety of the structure were determined in order to see whether the structural components of the RASAT can withstand the vibration induced launch loads or not.

This study had been performed by both TÜBİTAK Space Technologies Research Institute and Mechanical Engineering Department of Middle East Technical University, and supported by RASAT Satellite Development Project, financed by DPT (State Planning Organization of Turkey) under grant contract no 2007K120030.

2 Structural Specifications of the RASAT Satellite

The RASAT satellite has a dimension of 730 *mm* x 730 *mm* x 554 *mm* with a cubic shape. The main structural design of the RASAT is based on a module-stack frame type. There is a main module-stack which is composed of several micro trays. On both sides of the stack there are stiffened facing plates. The one that looks through the earth on orbit is called as "Earth Facing Facet (EFF)" and the other one that looks through the outer space is called as "Space Facing Facet (SFF)". The EFF is connected to the stack from the bottom on which there is a stiffened module (called as BCR). The SFF is connected to the main stack by means of some stiffened panels, which are called as "Shear Panels". Lastly the solar panels are mounted between SFF and EFF. All the primary structure of the RASAT is formed as above. In Fig. 1, the main module-stack, the space facing facet, the earth facing facet, the shear panels and the solar panels are shown as an exploded view.

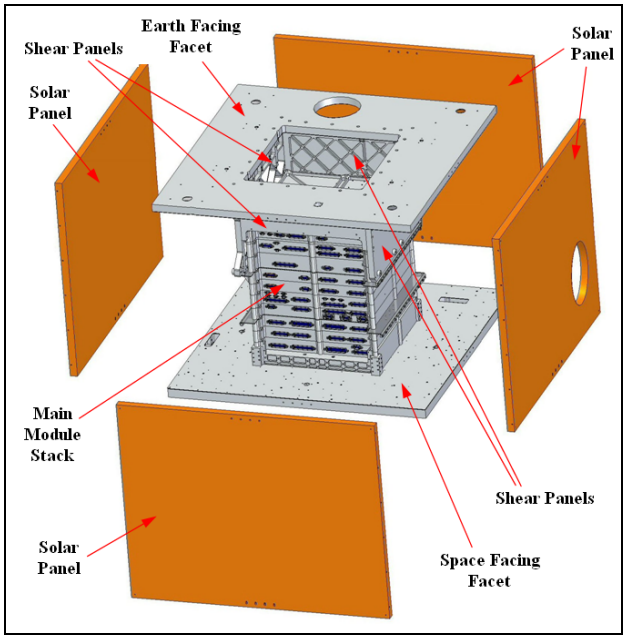


Fig. 1. Exploded View of the RASAT Structure

Therefore, by mounting all the primary structural elements (SFF, EFF, shear panels), the main module-stack, the nano stacks, their modules, solar panels, and the other electronic components like antennas, the overall design of the RASAT is completed as shown in Fig. 2.

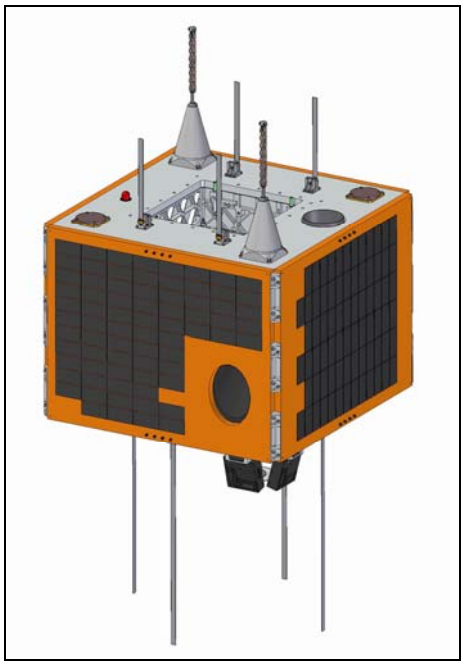


Fig. 2. General 3-D view of the RASAT Satellite

3 Finite Element Model

The finite element model of the RASAT was formed by transferring the CAD solid model from CAD software to the finite element software. Almost all of the parts which have structural importance were included in the finite element model. After transferring the part models, they were prepared to make the required analyses.

First of all, all the parts were assembled in the FEA software as in the CAD assembly. All the locations, placements and constraints were done according to the real case. The next step was to mesh the assembly. While meshing, the assembly was not meshed as a whole body since the parts forming the assembly do not have the same structure and similar geometry. For instance some parts were in the solid structure and some were in the shell structure. So, all the parts were meshed independently. While meshing, firstly, lower number of elements was selected according to the size of the component being meshed. Then, the number of elements was increased in order to obtain the results so that two successive results should have a relative error of 10%. Finally, the RASAT assembly model was meshed to perform analyses as shown in Fig. 3 and Fig. 5. The coordinate axes were chosen consistent with those used in mass calculations.

After meshing all the parts and assigning all the corresponding material properties, the next step was defining the boundary conditions in order to carry out the analysis. Since, the satellite is attached to the launch vehicle from its separation adapter, normally; the boundary condition is going to be the four connection holes of the separation adapter on the launch vehicle side. Boundary conditions consisted of nodes on the bottom flange of the separation system were being constrained in all degrees of freedom. These constraints were applied at four equally spaced angular locations around the separation system, corresponding to the actual bolted connections to the launch vehicle, as shown in Fig. 4.

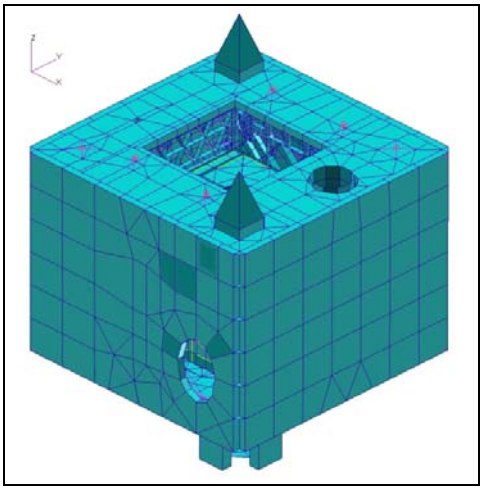


Fig. 3. Finite Element Model of the RASAT

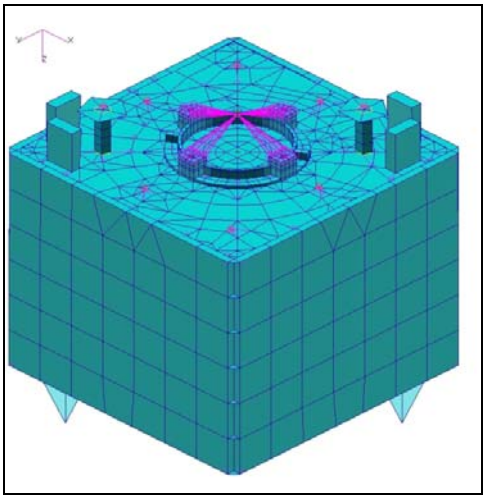


Fig. 4. Boundary Conditions of Finite Element Model

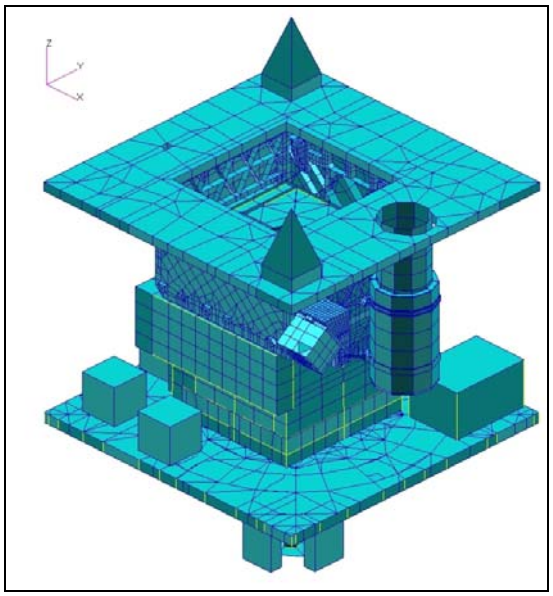


Fig. 5. Inner View of Finite Element Model

The parameters of the finite element model are as follow:

- The total number of elements: 213577
- The total number of nodes: 97268

Various materials were used in the structural parts of the whole satellite. Generally aluminum alloys were selected due to their easy provision, easy machining, good strength/weight ratio and good thermal properties. Most of the structural elements like SFF, EFF, shear panels, module boxes, and brackets were made of aluminum alloys. In addition to aluminum alloys, titanium was selected for tie-bars due to its good strength and mass properties. The isotropic materials used in the RASAT were listed in Table 1.

Other than the isotropic materials, there are some materials having orthotropic properties. For instance, the structure used for the core of the solar panels is honeycomb. The honeycomb core material has a 3-D orthotropic property as listed in Table 2. Similarly, the skin material for the solar panels is Carbon Fiber Reinforced Plastic (CFRP). Since these skins were modeled as shell elements, a 2-D orthotropic material property was used as listed in Table 3.

Table 1. Isotropic materials used in the model with ref [1]

Material	MoE [GPa]	Poisson's Ratio	Density [kg/m^3]	UTS [Mpa]
Al 5050	68.9	0.33	2.69	172
Al 6061-T6	68.9	0.33	2.70	310
Al 7075-T6	71.7	0.33	2.81	572
SS 416	200	0.33	7.80	760
Ti6Al4V	113.8	0.342	4.43	950
PEEK 450G	3.5	0.40	1.30	97

Table 2. Honeycomb core material properties (3-D orthotropic) with ref [1]

Property		Value
Moduli of Elasticity [MPa]	E_{11}	0.264
	E_{22}	0.264
	E_{33}	1034
Poisson's Ratios	ν_{12}	0.40
	ν_{23}	-
	ν_{31}	-
Shear Moduli [MPa]	G_{12}	1
	G_{23}	214
	G_{31}	214
Density [kg/m^3]	0.0190	

Table 3. CFRP-M18/45%/193P/AS-4-3K material properties (2-D orthotropic) with ref [2]

Property	Value
E_{11} [GPa]	101.92
E_{22} [GPa]	101.92
ν_{12}	0.30
G_{12} [GPa]	3
Density [kg/m^3]	1.4349

4 Finite Element Stress Analysis

4.1 Design Load Factors

The most important and effective load conditions for performing stress analysis are launching loads. Launching loads differ as relative to the launching system. In a few minutes, when the engines of the launching system run, the satellite is exposed to extremely high dynamic loads.

The launch environment is generally the main satellite structural design driver. While designing and analyzing, a set of loads are defined for the structure in order to be sustained. These are called design load factors and generally expressed in terms of equivalent accelerations, in other words in g's.

There are two different types of design load factors:

- Quasi-static load factor (QSLF)
- Random vibration load factor (RVLF)

The relative timing of the quasi-static and random vibration environments is unique for each launch vehicle, but simultaneous occurrence of maximum low frequency quasi-static and maximum random vibration load is not so probable. Therefore, a root-sum-square (RSS) approach is suitable for combining the maximum low frequency and maximum random vibration loads for the liftoff flight event. When the low frequency quasi-static and random vibration environments are specified in a time correlated manner, a time consistent approach is also acceptable for combining the low frequency quasi-static loads and the random vibration loads.

When an RSS approach to defining component loads is utilized, the maximum low frequency and maximum random vibration load factor are root sum squared. The root sum squared values are then applied in all axes simultaneously. When the load combination is directly applied to member loads, the maximum random vibration and low frequency member forces are root sum squared.

Because quasi-static and transient events have positive and negative load factors, all possible combinations must be considered when combined with respective random vibration loads. Table 4 describes how the random vibration loads are superimposed, one axis at a time, with quasi-static and transient loads. The impacts of other loads, such as acoustic, thermal, and misalignment, shall also be considered where applicable. Resulting design load factors for each axis can be calculated and applied simultaneously in a stress analysis.

Table 4. Combination of Load Factors

Load Case	X-Axis	Y-Axis	Z-Axis
LC 1 to 8	$\pm ((QSLF_x)^2 + (RVLF_x)^2)^{1/2}$	$\pm QSLF_y$	$\pm QSLF_z$
LC 9 to 16	$\pm QSLF_x$	$\pm ((QSLF_y)^2 + (RVLF_y)^2)^{1/2}$	$\pm QSLF_z$
LC 17 to 24	$\pm QSLF_x$	$\pm QSLF_y$	$\pm ((QSLF_z)^2 + (RVLF_z)^2)^{1/2}$

4.2 Random Vibration Load Factor

The Random Vibration Load Factors (RVLF) acting on a spacecraft structure results from the resonant structural response of the equipment to random vibration environments (input from rocket engines mechanically and acoustically induced vibrations) during launch. Therefore, random vibration load factors shall be combined with low frequency load factors and other loads which apply during the launch phase of flight operations. Also, the RVLF's do not apply to large mass items since they will not respond to random vibrations (at frequencies significantly above those included in the low frequency load factor range).

Specific RVLF's for satellite structures may be calculated using the following procedure:

A. By calculation or from the modal analysis, the first system natural frequency, (f_n), in each axis is determined.

B. The resonant amplification factor, (Q), (equal to 1/2x, where x is the structural damping coefficient) is determined. This factor is taken from test data, if available. In the other hand, 10 is used if no data is available, which is general amplification factor for spacecraft applications.

C. Using the natural frequency, f_n calculated in step A, the applicable acceleration Power Spectral Density (PSD_n in g^2/Hz) corresponding to the natural frequency f_n can be determined. If f_n falls on a sloped portion of the curve, PSD_n can be interpolated using the following relationship:

$$PSD_n = PSD_1 (f_n / f_1)^{0.3322s} \quad (1)$$

Where:
PSD_n: PSD value which corresponds to f_n
PSD_1: PSD value corresponding to f_1
s: The slope of the PSD function (dB per octave)
f_1: Lowest frequency given for the portion of the PSD table being interpolated
f_n: Natural frequency of system in the direction in which the load factor is being calculated.

D. The peak RVLF in each axis is determined using the Miles' relationship:

$$RVLF = 3\sqrt{(\pi/2) Q f_n PSD_n} \quad (2)$$

The factor of three in each of the above equations is a statistical factor applied because the load factors calculated from the power spectral density curves are standard deviation amplitude for random vibrations, and design shall be based upon three standard deviations. In other words, the results obtained are the RMS stress values. But these values cannot be used in stress analysis due to they just give 1 sigma, or 68.3% certainty. In order to be on the safe side, we have to increase this certainty value and find the effective stress values.

The simplest and most effective way of converting these RMS stress values to the effective stress values is to multiply the RMS values by a greater sigma value. This sigma value is selected according the application type on which the analyzed components are being used. Generally for spacecrafts, sigma value is chosen as 3, in order to estimate with 99.737% certainty [3].

For the given PSD spectrum, the RVLF is calculated for each resonant frequency. In this case, in order to find the overall RVLF for the whole spectrum, each RVLF is multiplied by the effective mass of the corresponding frequency and the multiplication is divided by the overall mass. So, by this procedure, the mass weighted RVLF (i.e. RVLFMW) for each frequency is determined with the following formula:

$$RVLF_{MW} = \frac{RVLF_i \cdot EFFM_i}{M} \quad (3)$$

Finally, for every direction the RVLF is calculated with the relation of RSS:

$$RVLF_{(x,y,z)} = \sqrt{\sum_1^n (RVLF_{MW})^2} \quad (4)$$

Table 5. Spectral Density of Vibro-accelerations at SC/LV Interface with ref [4]

Frequency sub-band [Hz]	Power Spectral Density [g^2/Hz]
20 to 40	0.007
40 to 80	0.007
80 to 160	0.007-0.022
160 to 320	0.022-0.035
320 to 640	0.035
640 to 1280	0.035-0.017
1280 to 2000	0.017-0.005
Root Mean Square Value, σ [g]	6.5
Duration [sec]	35

For RASAT, the random vibration levels were given in Table 5 in order to calculate the RVLF's. After making the calculations using the above relations, the RVLF's of the RASAT for all three axes are presented in Table 6.

Table 6. Random Vibration Load Factors

Random Vibration Load Factor	Value [g]
$RVLF_x$	6.17
$RVLF_y$	6.29
$RVLF_z$	14.24

All the RVLF's found were combined with the quasi-static load factors as shown in red in Table 7 and the load combination table was determined. In Table 8, all these 24 load cases are presented.

Table 7. Maximum Quasi-static and Dynamic Accelerations at SC/LV Interface with ref [4]

	Acceleration [g]	Acceleration [g]
Load Source	**Longitudinal (X)**	**Lateral (Y, Z)**
LV movement inside TLC	2.5±0.7	±0.3
After LV exit from TLC	±1.0	±0.8
1st stage burn		
Maximum dynamic head	3.0±0.5	**0.5±0.5**
Maximum longitudinal acceleration	7.5±0.5	0.1±0.5
2nd stage burn – maximum longitudinal acceleration	**7.8±0.5**	0.2
3rd stage burn	-0.3...-0.5	0.25

4.3 Load Cases and Stress Analyses

In order to calculate the equivalent stresses occurred on the structure during launching, all these load combinations were applied on the satellite model and the most effective one is going to be selected for the strength case. In this case there were 24 load cases. All these 24 load cases were applied on the finite element model of the RASAT respectively. In other words all the margins of safety for each case were determined and the smallest margin of safety for the satellite was obtained.

The formulas used in calculating yield and ultimate margin of safety were given below as Formula (5) and (6) [5].

$$MOS_y = \left(\frac{\sigma_{yield}}{SF_{yield} \cdot Stress_{\max}^{VonMises}} - 1\right) \quad (5)$$

$$MOS_u = \left(\frac{\sigma_{yield}}{SF_{yield} \cdot Stress_{\max}^{VonMises}} - 1\right) \quad (6)$$

Table 8. Load Cases of the RASAT Stress Analysis

LOAD CASE	X [g]	Y [g]	Z [g]
LC 1	10,342	1	1
LC 2	10,342	1	-1
LC 3	10,342	-1	1
LC 4	-10,342	1	1
LC 5	10,342	-1	-1
LC 6	-10,342	1	-1
LC 7	-10,342	-1	1
LC 8	-10,342	-1	-1
LC 9	8,3	6,369	1
LC 10	8,3	6,369	-1
LC 11	8,3	-6,369	1
LC 12	-8,3	6,369	1
LC 13	8,3	-6,369	-1
LC 14	-8,3	6,369	-1
LC 15	-8,3	-6,369	1
LC 16	-8,3	-6,369	-1
LC 17	8,3	1	14,275
LC 18	8,3	1	-14,275
LC 19	8,3	-1	14,275
LC 20	-8,3	1	14,275
LC 21	8,3	-1	-14,275
LC 22	-8,3	1	-14,275
LC 23	-8,3	-1	14,275
LC 24	-8,3	-1	-14,275

In calculations, the yield safety factor was chosen as 1.25 and the ultimate safety factor was chosen as 1.4. These values were taken for metallic structures from Table-1 of ref [6].

It is known that to pass for the stress analysis, the required safety margins are:

$$MoS_y > 0 \text{ and } MoS_u > 0 \qquad (7)$$

It was observed after the analyses that all the yield and ultimate margins of safety were all positive. The most critical three load cases were tabulated in Table 9. So, the stress analysis has been finished with satisfactory results.

Table 9. Chosen critical Results of Stress Analysis

Load Case	Component	Max. Von Mises Stress Value [Mpa]	Material	Allowable Stress Value		Minimum Margin of Safety (MoS)	
				σ_{ty} [Mpa]	σ_{tu} [Mpa]	MoS_y	MoS_u
1	Camera Left Flexure	149.1	Ti6Al4V	890	950	3.78	**3.55**
9	Camera Left Flexure	147.4	Ti6Al4V	890	950	3.83	3.60
17	Camera Left Flexure	144.2	Ti6Al4V	890	950	3.94	3.71

From Fig. 6 to Fig. 8, the Von Mises stress contour plots were represented in order to show the maximum stress nodes and locations. In some cases, the maximum stress locations are the same. The maximum stress locations are generally on the rear left e-flexure of the main camera optics structure. But when the corresponding material properties are considered, all these components left on the safe side.

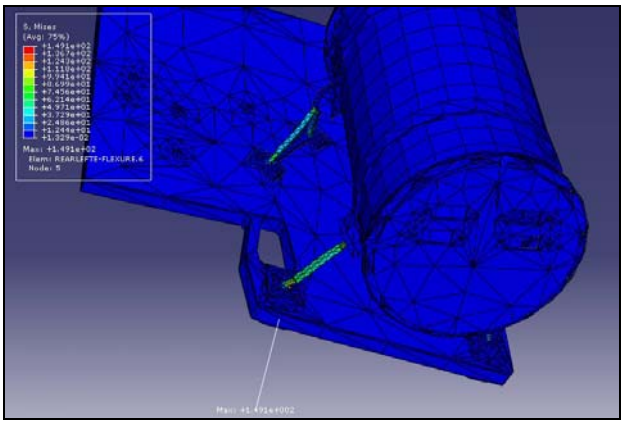


Fig. 6. Von Mises Stress contours of LC 1

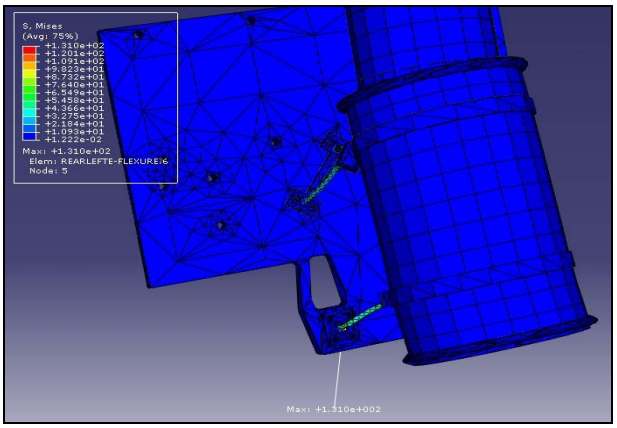


Fig. 7. Von Mises Stress contours of LC 9

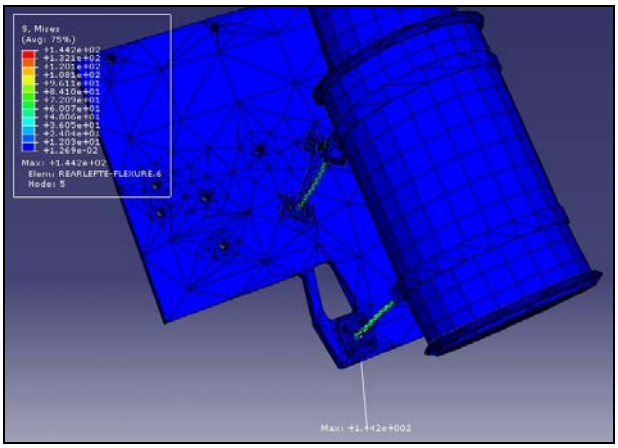


Fig. 8. Von Mises Stress contours of LC 17

5 Conclusion

It was determined that all the 24 load cases have positive safety margins. This means that after the stress analysis the resulting stress values were on the safe side. So, the satellite and its components satisfy the requirements on stress analysis.

But, although all the margins of safety obtained from the stress analysis were positive, their values were a little greater than the optimum ones. Actually, safety margins were accepted as optimum if they are between 0 and 0.5, and not efficient if they are greater than 1.5 [7]. So, the stress analysis of RASAT can be considered so that the RASAT structural design was not satisfactorily efficient since the minimum margin of safety found was 3.55, which is much greater than the optimum interval. Therefore, some design improvements should have been made in order to get a more efficient structural design.

References

[1] Material Property Data Website, "http://www.matweb.com", Last Access date: April 2011

[2] GAM 215/07 Volume I, "User Manual", Galileo Avionica Milano, 2008

[3] NASA-STD-5002, "Load Analyses of Spacecraft and Payloads", June 1996, Chap. 5.3.3

[4] "DNEPR User's Guide", Issue 2, ISC Kosmotras, November 2001

[5] NASA-SSP5200B, "Payload Flight Equipment Requirements and Guidelines for Safety Critical Structures", International Space Station Program, December 1998

[6] NASA-STD-(I)-5001A, "Structural Design and Test Factors of Safety for Spaceflight Hardware", December 2006

[7] Fraser D., Kleespies H., Vasicek C., "Spacecraft Structures", 1991, Chap. 9.5

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

COMPUTATIONAL EVALUATION AND EXPLORATION OF COMBUSTION PERFORMANCE FOR LIQUID JET FUELS DERIVED FROM BIOMASS

Nurul Musfirah Mazlan[1], Mark Savill[2], Timos Kipouros[3]
[1] PhD student
n.mazlan@cranfield.ac.uk
[2] Professor
mark.savill@cranfield.ac.uk
[3] Lecturer
t.kipouros@cranfield.ac.uk
Department of Power and Propulsion, School of Engineering, Cranfield University
Beds, MK43 0AL, United Kingdom

Abstract

This study is designed to estimate droplet evaporation rate and the spray penetration of bio-fuels in an attempt to predict combustion completeness using equations related to droplet motion, energy, and boiling mass transfer obtained from FLUENT. Variation of fuel properties such as density, vapour pressure, specific heat capacity, diffusivity, and latent heat of vaporisation with temperature, are necessary in accommodating these equations. Fuels considered in this study are Ethanol, Rapeseed Methyl Ester (RME), and Jatropha Bio-Synthetic Paraffinic Kerosine (JSPK). Additionally, a blended fuel comprising 50% JSPK and 50% Jet-A (B50), as considered in some recent flight tests, is also examined. RME is observed to evaporate slowest owing to its smallest vapour pressure resulting in the longest length of spray penetration, whilst ethanol evaporates fastest due to its high volatility. Investigation of B50 has found to increase the evaporation rate and spray penetration length compared with pure JSPK.

1 Introduction

Major challenges society is facing today include increases in pollutant emissions and rising oil demand, which consequently increases fuel prices. An improvement in energy efficiency has been suggested to help minimise both of these problems. Utilising alternative fuels in aircraft engine is considered to be one of the options available for improving energy efficiency instead of improving aircraft design and changing aircraft operations [1, 2]. Although alternative fuels are able to reduce the quantities of emissions polluting the atmosphere, utilising such fuels in aviation is not easy owing to their poor properties, which are not fully attuned to the combustion conditions of gas turbine jet engines. In an attempt to establish an ideal candidate for supplement or even replacement to conventional fuel, a number of studies and several flight tests had been conducted, which have shown that fuels derived from biomass have the potential to be utilised in aircraft engines, although they can only be used as blended fuel [3]. The capability of bio-fuels in reducing emissions has been considered through a number of studies which have shown that the quantity of nitrogen oxide

(NO_x), carbon monoxide (CO) and soot was reduced.

Despite the fact that bio-fuels potentially can reduce the emissions of pollutants into the atmosphere, there is little knowledge concerning the effect of these fuels during the atomisation process, which consequently affects combustion performance and the formation of such pollutants. With this in mind, this study is conducted with the main objective to evaluate the effect of alternative fuels specifically, bio-fuels and synthetic fuel on combustion performance, focusing on the evaporation rate and spray penetration. The combustion performance of bio-fuels and synthetic fuel has been investigated numerically using equations obtained from FLUENT with established essential of fuel properties. For comparison purposes, combustion performance of conventional jet fuel (kerosine) has also been evaluated.

2 Bio-fuels challenges

Rapid progression in alternative fuels studies has been encouraged by a desire to reduce dependency on petroleum based fuel such as Jet-A and Jet-A1, in civil aviation, both of which have been known to have high energy content, stable properties, and to offer relatively high volumetric and gravimetric energy [4].

Although, bio-fuels could promise reductions in carbon footprint and provide great savings in carbon dioxide emissions, there nevertheless appear to be problems due to their sustainability and properties that could make bio-fuels inappropriate for aircraft engines. Besides low energy content, the tendency of the fuels to freeze at normal cruise temperature would require significant modification to the engine design. For instance, alcohol-based fuels, such as ethanol and methanol, cannot be used in the case of commercial aircrafts because of their poor mass and volumetric heat of combustions; this will result in the aircraft requiring a larger wing to carry the fuel, causing increases in aircraft weight [3].

Vegetable oil is another example of bio-fuels with the potential to be used in gas turbine engines; however, owing to the high viscosity, serious problems in engines can be experienced, such as poor atomisation, incomplete combustion, and coking of fuel injectors.

Biodiesel, on the other hand, has been found to have an improvement in viscosity, but due to poor freezing point (-20°C), thermal stability [3] and lower energy content in biodiesel (Low heating value, LHV = 36 – 39 MJ/kg) compared with conventional jet fuel (typical value is 42 MJ/kg) [4], biodiesel is only allowed to be used as a blended fuel at a maximum of 20% [3].

Other options available for aviation purposes come from the process of the hydrotreating/hydroprocessing of fatty acids or fatty acid esters to produce synthetic paraffinic kerosine (SPK) type fuel (Bio-SPK). SPK fuel can also be produced via gasification from the Fischer-Tropsch (FT) synthesis process using coal, gas or biomass as feedstock. Both of these methods (Bio-SPK and FT-SPK) have a similar process, although the difference relies on the n-paraffins product form at the end of the process. Notably, the Bio-SPK's n-paraffins product usually has a much narrower carbon distribution than FT-SPK [5].

The capability of Bio-SPK type of fuels derived from Jatropha, Camelina and Algae has been investigated, and several test flights have been performed to evaluate both engine performance and emissions [6]. The flight tests conducted have shown that Bio-SPK fuel is able to reduce aircraft fuel consumption and emissions, especially NO_x and smoke. Unfortunately, however, the low density of this fuel (e.g. 749 kgm^{-3} for JSPK) is a constraint to meeting specific requirements for gas turbine jet fuel (density allowed is between 775–840 kgm^{-3}); therefore, these fuels have to be blended with conventional jet engine fuel.

3 General behaviour of vaporisation process and evaporation rate prediction

The behaviour of a droplet during the vaporisation process and the estimation of evaporation rate generally can be described from d^2 law, as shown in Fig. (1) and has been discussed in detail by Ghasemmi *et al* [7].

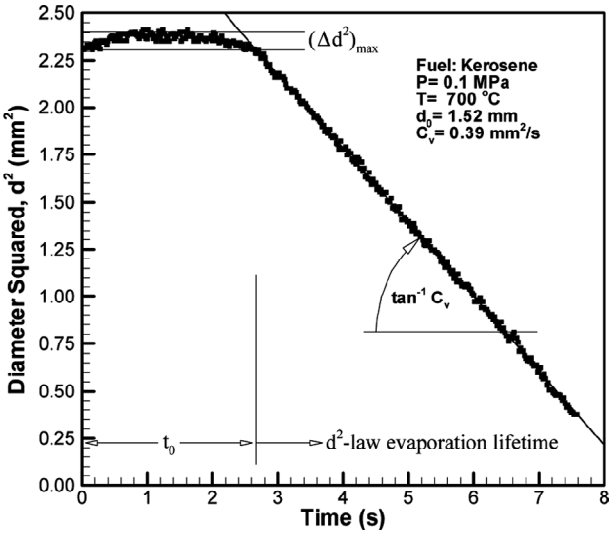


Fig. 1: Evaporation history of kerosine droplet [7]

The process of evaporation comprises two regions. The first region represents non-linear behaviour and indicates the heat-up process (t_0). In this region, droplets are heated-up, and thermally conducted and expanded from gas, resulting in increases in droplet diameter. As the droplet reaches its boiling temperature, the evaporation process starts and will control the changes in droplet size. This phenomenon is presented by the second region, where d^2 law is valid. For the small size of the droplet, the heat-up process can be neglected and is less important. Therefore, d^2 law (as expressed in eq. (1)) becomes significant to predict evaporation rate. However, for large liquid fuel droplets, t_0 region is essential. The evaporation rate for small droplets is predicted from eq. (1) below.

$$d^2 = {d_0}^2 - C_v t \qquad (1)$$

where d_0 is initial droplet diameter and C_v is the evaporation coefficient. C_v is calculated from time derivative of droplet squared diameter (eq. 2) which can be extracted from the linear part of the evaporation history curve using least linear regression. The same method was also applied by Chin and Levebvre [13] who predict the evaporation rate from the correlation between $d/{d_0}^2$ and $t/{d_0}^2$.

$$C_v = \frac{-d(d^2)}{dt} \qquad (2)$$

4 Materials and methods

This work focuses on the effects of bio-fuels properties in the evaporation process and spray penetration as part of combustion performance. Bio-fuels, indicated in Table (1), are chosen and the effect of blend Bio-SPK fuel with kerosine is also investigated.

Table 1: Molecular formula and composition of fuels chosen in this study

Case	Molecular Formula	Fuel Composition
Kerosine	$C_{12}H_{23}$	100%
RME	$C_{19}H_{32}O_2$	100%
Ethanol	C_2H_5OH	100%
JSPK	$C_{12}H_{26}$ *	100%
B50	Kerosine	50%
	JSPK	50%

* By assuming JSPK only comprises of carbon and hydrogen, molecular formula of JSPK is estimated from ratio of carbon to hydrogen in conjunction with average carbon number in Jet-A which in this study is considered as 12.

In assessing the effect of bio-fuels on evaporation time, equations related to mass, momentum and energy exchange between droplet and air from FLUENT are used. Most of the content in this section are based on, and further detailed explanations can be found in the FLUENT User's Guide [8].

FLUENT predicts droplet trajectory by integrating drag force, gravitational force, and additional forces applied on a droplet. Additional forces will usually include those forces which are necessary to accelerating the fluid surrounding the particle, but are not very important. Accordingly, this study focuses only on the particle's drag; therefore, the second and third parts of right-hand side of the eq. (22.2-1) in FLUENT is neglected.

Thus, the new equation for predicting the motion of the particle is as below:

$$\frac{dV_p}{dt} = \frac{18\mu}{\rho_p {d_p}^2} \frac{C_D \,\mathrm{Re}_p}{24} \left(V_\infty - V_p\right) \qquad (3)$$

where μ is molecular viscosity of air ($kgm^{-1}s^{-1}$), ρ_p is density of the particle (kgm^{-3}), d_p is particle diameter (μm), C_D is drag

coefficient, V_∞ and V_p are velocity of air and velocity of particle respectively (ms^{-1}). Reynolds number in eq. (3), Re_p is defined as

$$\mathrm{Re}_p = \frac{\rho d_p \left| V_p - V_\infty \right|}{\mu} \qquad (4)$$

where ρ is density of air (kgm^{-3}). Drag coefficient, C_D in eq. (3) was taken from Morsi and Alexander, [9] as

$$C_D = a_1 + \frac{a_2}{\mathrm{Re}_p} + \frac{a_3}{\mathrm{Re}_p^{\,2}} \qquad (5)$$

in which a_1, a_2, and a_3 are constants and varies with Reynolds number.

For the condition where the temperature of the fuel's droplet is less than the boiling temperature, a reduction in mass particle $\frac{dm_p}{dt}$ was calculated according to equation (6).

$$\frac{dm_p}{dt} = -N_i A_p MW_p \qquad (6)$$

where A_p is particle area (m^2), MW_p is molecular weight of the particle (kg). N_i, is the molar flux of droplet's vapour, which corresponds to the difference in vapour concentration between droplet surface and bulk gas.

$$N_i = k_c \left(C_{i,s} - C_{i,\infty} \right) \qquad (7)$$

k_c is mass transfer coefficient (ms^{-1}), $C_{i,s}$ and $C_{i,\infty}$ are vapour concentration at droplet surface ($kgmolm^{-3}$) and vapour concentration of vapour gas ($kgmolm^{-3}$) respectively.

In order to calculate $C_{i,s}$, FLUENT assumes that the partial pressure of vapour at the interface is equal to the saturated vapour pressure at the particle droplet temperature.

$$C_{i,s} = \frac{p_{sat}\left(T_p\right)}{RT_p} \qquad (8)$$

p_{sat} is saturated vapour pressure of the particle (Pa) at a particle temperature, T_p (K), R is universal gas constant (287 $Jkg^{-1}K^{-1}$). $C_{i,\infty}$ in eq. (7) is calculated from the transport equation of the gas.

$$C_{i,\infty} = X_i \frac{p}{RT} \qquad (9)$$

where X_i is local bulk fraction of species i, p and T are pressure (Pa) and temperature (K) of the gas respectively. Mass transfer coefficient, k_c in eq. (7), is calculated from Sherwood number correlation

$$Sh_{AB} = \frac{k_c d_p}{D_{i,m}} = 2.0 + 0.6\,\mathrm{Re}_p^{\frac{1}{2}} Sc^{\frac{1}{3}} \qquad (10)$$

$D_{i,m}$ is diffusion coefficient of vapour pressure in the bulk (m^2s^{-1}) and Schimdt number, Sc is calculated from equation below

$$Sc = \frac{\mu}{\rho\, D_{i,m}} \qquad (11)$$

Finally, the changes in droplet temperature $\frac{dT_p}{dt}$ are calculated from the heat balance, which is associated with the changes in sensible heat in the droplet and latent heat transfer between the droplet and the continuous phase. By assuming no radiation heat transfer takes place, and rearranging the eq. (22.9-25) from FLUENT, the droplet temperature is calculated using the equation below

$$\frac{dT_p}{dt} = \frac{hA_p\left(T - T_p\right) + \frac{dm_p}{dt} h_{fg}}{m_p C_p} \qquad (12)$$

h is convective heat transfer coefficient ($Wm^{-2}K^{-1}$), h_{fg} is latent heat of vaporisation (Jkg^{-1}) and C_p is heat capacity of particle ($Jkg^{-1}K^{-1}$).

All equations above are used in conjunction with established fuel properties obtained from literature. This numerical calculation has been conducted in pure air, at temperature of 1000K, and at a pressure of 1 atm. The droplet size produced from fuel injector is assumed to be 20µm. The temperature of particle is also assumed to be the same for all the fuels which 300K. The droplet generated by the fuel injector is assumed to be a sphere owing to its

surface tension; however, this is not valid for diesel engine, as the fuel is injected at high pressure causes the droplet to have long ligament. This calculation is done in a time step of 10μs and stopped when droplet's velocity is approaching zero.

5 Results and discussion

In this work, the effect of bio-fuels properties on combustor performance particularly evaporation rate and spray penetration are investigated numerically using ethanol, rapeseed methyl ester (RME), and bio-synthetic paraffinic kerosine type fuels derived from jatropha curcas (JSPK). Additionally, a blend of 50% JSPK with 50% kerosine (B50) is also evaluated. Finally, the investigation on combustor performance from bio-fuels is compared with kerosine.

5.1 Vaporisation general behaviour

Evaporation history curve of Jet-A used in this study is presented in Fig. (2). Compared with Fig. (1), the evaporation curve obtained in this study is followed the evaporation curve from experimental work in Ghasemmi *et al* [7] apart from no increases in droplet diameter during heating-up region. This is consistent with the study by Frohn and Roth [10] (cited from [7]), which mentions that t_0 region is neglected for the case of small size droplet.

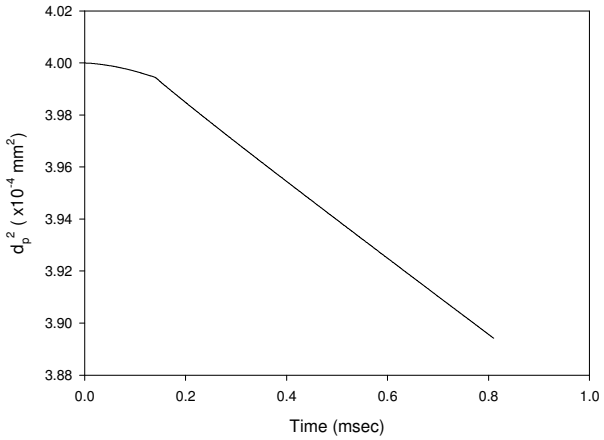


Fig. 2: Evaporation curve of Jet-A

The vaporisation curves of fuels chosen in this study are shown in Fig. (3) and Fig. (4) presented by the variation of normalised squared diameter with the normalised time. Fig. (4) shows the enlarged vaporisation curve for Jet-A, JSPK and B50.

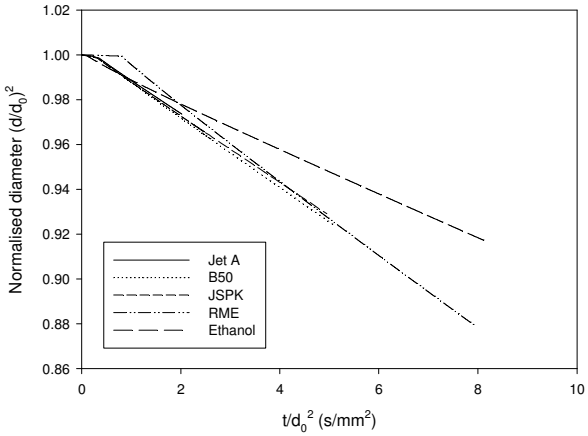


Fig. 3: Comparison of evaporation curve for chosen fuels

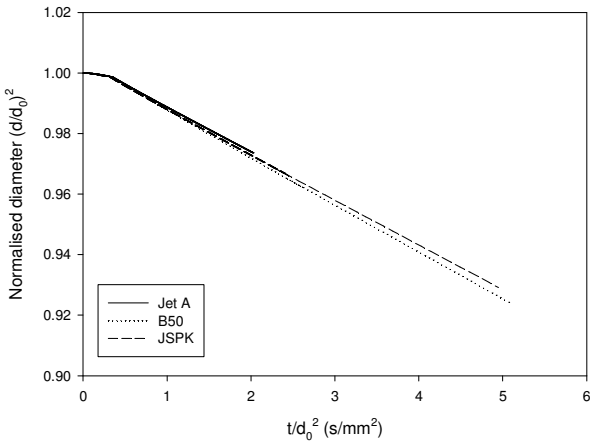


Fig. 4: Comparison of evaporation curve for Jet-A, B50 and JSPK (enlarged)

It can be seen from the Fig. (3) and Fig. (4), that the behaviour of all the fuels during the vaporisation process follows d^2 law, with no increases in droplet diameter observed.

5.2 Comparison in Fuels properties

Comparisons between the properties of the selected fuels, which are required to accommodate the equations above, are presented in Fig. (5) - Fig. (8). Properties of Jet-A, JSPK and B50 were obtained from Kinder [5]; properties of ethanol were obtained from Khasanshin [11], whilst properties of RME were obtained from Rochaya [12]. Properties of fuels

indicated in Table (1) are summarised in Table (2).

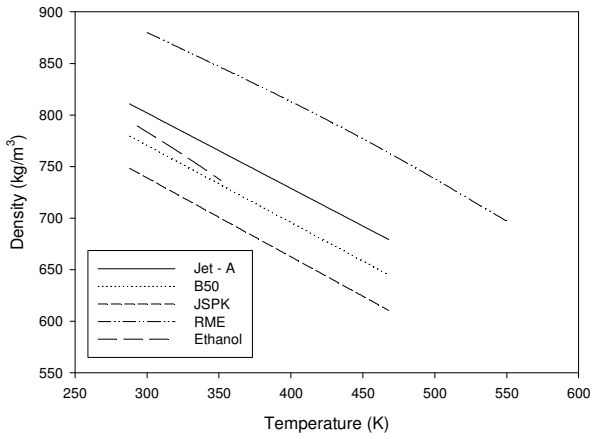


Fig. 5: Density of the Chosen Fuels

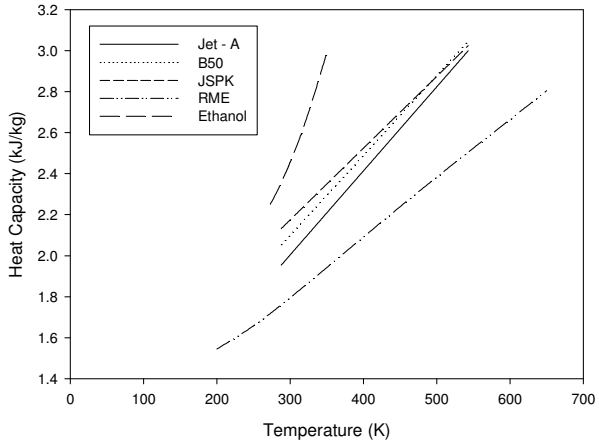


Fig. 6: Heat capacity of the Chosen Fuels

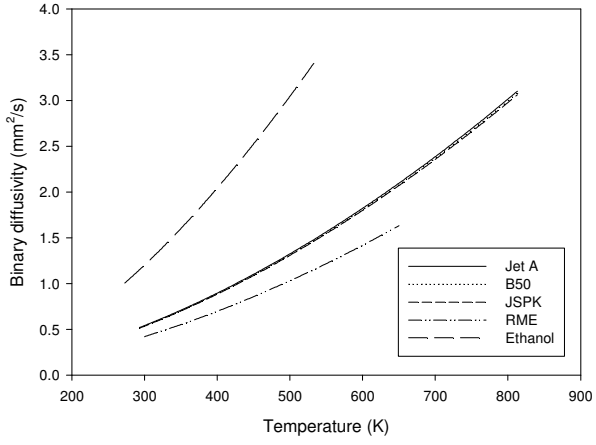


Fig. 7: Binary Diffusion Coefficient of Selected Fuels*

*Binary diffusivity of fuels is estimated using an equation suggested in [18], which requires information of fuel's molecular formulae to calculate fuel's weight and fuel's diffusion volume. Although binary diffusivity coefficient is using in this work, the effect of this property in evaporation rate and spray penetration however is not been investigated.

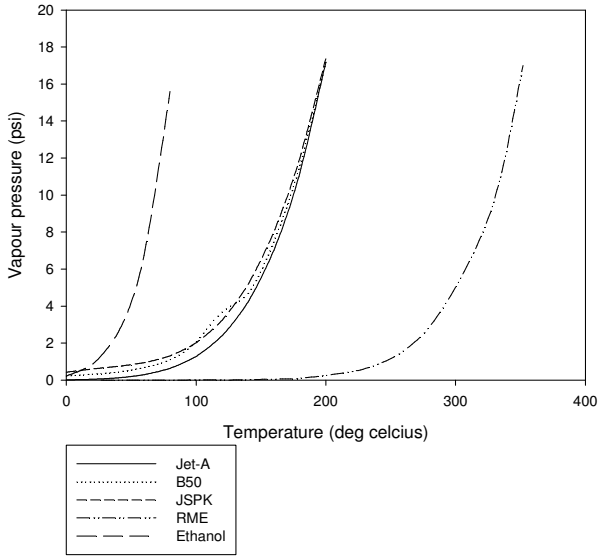


Fig. 8: Vapour pressure of the Selected Fuels

Table 2: Summary of fuel's properties

Fuels	Density at 15°C (kg/m^3)	Boiling temperature (°C)	Enthalpy of vaporization (J/kg)
Kerosine	775 - 840	191.7*	251000
RME	877.6	350	105000
Ethanol	790	78	846000
JSPK	749	189.87*	106842
B50	779	190.75*	205921

* The value is obtained from vapour pressure curve. For JSPK and B50, the value is consistent with value for 50% distillation temperature. However, at 50% distillation temperature, no value is reported for Jet A.

5.3 Evaporation rate

The evaporation rate of the fuels is deduced from the relationship between d/d_0^2 and t/d_0^2 (Fig. 3) by applying the method developed by Chin and Levebvre [13] and is presented in Table (3).

Table 3: Evaporation rate

Fuels	Evaporation rate (mm^2/s)
Jet-A	0.0141
B50	0.0155
JSPK	0.0148
RME	0.0164
Ethanol	0.0101

The evaporation process of ethanol is the fastest followed in sequence with Jet-A, B50,

JSPK and RME. This is due to the properties of ethanol as will be discussed later in the following section.

5.4 Effect of fuel properties in evaporation process

Effect of boiling temperature

The effect of boiling temperature on evaporation can be seen from the changes of the droplet's temperature throughout the process (Fig. 9). During the heating-up process, temperature of the droplet increases until it reaches boiling point. In this process, increases in temperature are needed to change the liquid droplet into a vapour phase. Subsequently, changes in the droplet's temperature are very small and almost constant. As ethanol has the lowest boiling temperature amongst the others, the time required for ethanol to change from liquid into vapour is much faster compared with other fuels. This is important as the heat-up region (t_0) has to be kept as minimum as possible to allow combustion to take place. Minimising heat-up region is not only done by having a small size of droplet, but also by choosing a fuel able to vaporise quickly. RME, on the other hand, has high boiling temperature which consequently takes a longer time (approximately 0.3 sec) to vaporise the liquid droplet into vapour phase with the heat-up region unfortunately seeming significant (Fig. 3). However, for Jet-A, JSPK and B50, time required for these fuels during heat-up process are almost equal between each other, considering that there is not much different in their boiling temperature.

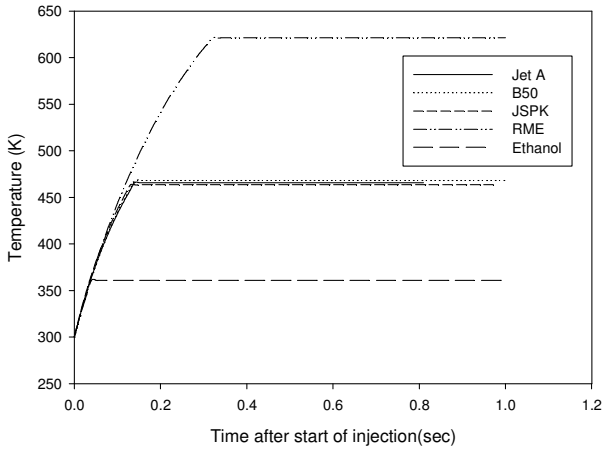


Fig. 9: Temperature distribution through evaporation process

Effect of heat capacity

Increasing droplet temperature also corresponds with internal energy, which depends on the fuel's heat capacity (Fig. 6) [14]. The heat capacity of ethanol is much higher than other fuels; in other words, ethanol requires more energy to heat the droplet up to a certain temperature, but requires a considerably shorter time to change into the vapour phase (Fig. 9).

Effect of vapour pressure

The contribution of the boiling temperature in the exaggerate vaporisation process is determined by the vapour pressure of the fuels. The comparison in the vapour pressure of fuels chosen in this study is shown in Fig. (8). The boiling temperature itself is determined by the vapour curve, which can be measured when the vapour pressure is equal to the atmospheric pressure. By definition, vapour pressure is the tendency of the molecules to escape from a liquid (or solid). At a normal temperature, fuels with high vapour pressure tend to vaporise faster than fuels with low vapour pressure. This is referred to as volatility. Volatility is inverse proportional to boiling temperature. As ethanol has a low boiling temperature and comparatively high vapour pressure, and is hence more volatile, the tendency of ethanol to vaporise is much faster than other fuels. The effect of vapour pressure and fuel volatility on evaporation rate is shown in Table (3).

Effect of enthalpy of vaporization (ΔH_{vap})

Despite the other properties discussed above, the evaporation rate of fuels is also influenced by their difference in enthalpy of vaporisation, which is defined as the amount of energy required by particles in a liquid phase to vaporise and change into gaseous phase. Fuel which has a low ΔH_{vap} will require more energy, and therefore requires more time to completely vaporise. As shown in Table (2) and Table (3), RME with comparatively low ΔH_{vap} is found to evaporate slower than other fuels, as opposed to ethanol which, comparatively has high ΔH_{vap}, and therefore requires a shorter time to evaporate.

5.5 Effect of fuel properties in spray penetration

Spray penetration is defined as a propagation distance of a droplet in the combustion chamber. Spray penetration is an important parameter that should be taken into account during the design of a combustor as it will affect the combustion performance. The penetration of a droplet has to be matched with the combustor size and geometry. If the droplet penetrates too short or inadequate, fuel-air mixing is unsatisfactory, the flame will appear in the fuel injector resulting in fuel coke and producing great deal of soot. While, as penetration is too long, fuel will impinge on walls surface and combustors [17].

Thus, just enough penetration is required to avoid soot and whilst, at the same time ensuring all the space in the combustion primary zone is used appropriately.

In order to estimate spray penetration, s equation recommended by Sazhin *et al* [15] which was used.

$$s = \frac{\sqrt{V_{in} D_0 t}}{(1-\alpha_d)^{\frac{1}{4}} \tilde{\rho}^{\frac{1}{4}} \sqrt{\tan\theta}} \times \left(1 - \frac{\sqrt{D_0}}{4\sqrt{V_{in}}(1-\alpha_d)^{\frac{1}{4}} \tilde{\rho}^{\frac{1}{4}} \sqrt{\tan\theta}\sqrt{t}} \right) \quad (13)$$

where

$$\tilde{\rho} = \frac{\rho}{\rho_d} \quad (14)$$

V_{in} is initial velocity of the particle (ms^{-1}), t is time (sec), α_d is volume fraction of droplets in spray, θ is half angle of spray cone, D_0 is nozzle diameter and $\tilde{\rho}$ is ratio of air density ρ and droplet's density ρ_p. In order to calculate spray penetration of the fuels, D_0 is taken as 1-mm and θ is 34.89° as used in [16] were implemented.

In the literature, spray penetration was found to be proportional with evaporation rate [14]. Contradictory in this work, spray penetration is not known to be affected by evaporation rate, but is much more significantly influenced by the density of the fuel's drop. It can be explained from eq. (13) where spray penetration is conversely proportional with ratio between gas density and droplet density (eq. 14). In other words, spray penetration is proportional to the density of the droplet. Since the density of JSPK is the lowest compared with other fuels, the, propagation of JSPK fuel's droplet is therefore the shortest amongst the others whilst RME is the longest owing to the high density of this fuel (Fig. 10).

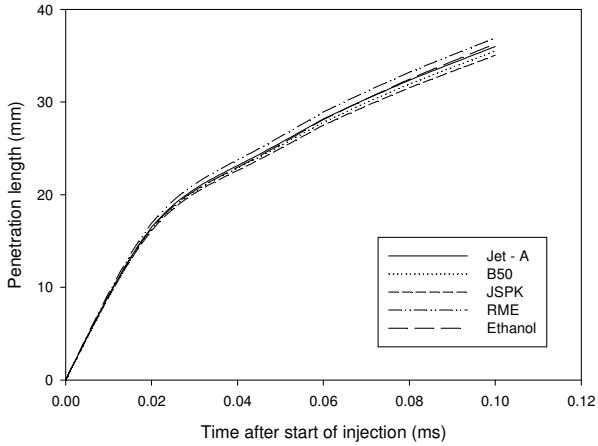


Fig. 10: Fuel's penetration length

5.6 Effect of blend fuel

Evaluation of evaporation rate and spray penetration of bio-fuels chosen in this study do not fully accommodate combustor specification in current aircraft. Although ethanol showed an improvement in evaporation and spray

penetration compared with Jet-A, it cannot however be used in the conventional combustor without any modification performed. In regard to bad evaporation consequently penetrating longest than the others, RME is also not accommodating of current combustor specification. To accommodate current combustor specification, bio-fuels need to have certain properties which are within the range of Jet-A properties.

The production of JSPK from the hydro-treating process must not only have those properties which are much more improved than RME, but also carbon and hydrogen distribution within Jet-A range which, at the first place, is a good candidate for replacing the use of Jet-A in aircraft engine. Nevertheless, the density of this fuel is rather lower, which consequently affects the propagation distance and yet still requires modification to the conventional combustor. For this reason, blend 50% of JSPK with Jet-A (B50) is found to be a good option as it is not only increasing density to be within Jet-A density, but also increases in enthalpy of vaporisation compared with pure JSPK, which is important for the evaporation process. The boiling temperature of B50 is almost similar with Jet-A consequently requires similar time to vaporise. Blend 50% JSPK with 50% Jet-A however increases evaporation rate compared with 100% JSPK owing to reduction in heat capacity. Additionally, increases in evaporation rate (increases 9.9% relative to Jet-A) also due to relatively low enthalpy of vaporization, which therefore evaporates slower than Jet-A.

Blend fuel is also found to penetrate shorter than kerosine (Fig. 11). As discussed, penetration was affected by density of the fuels. In comparison to JSPK, blend 50% of JSPK with kerosine increases its density consequently increasing penetration length.

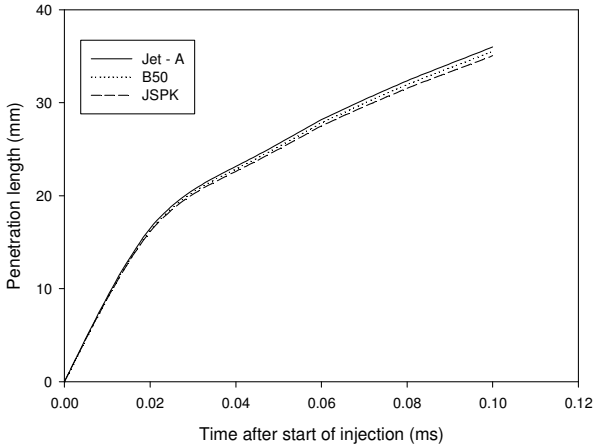


Fig. 11: Penetration length of blend fuel compared with Jet-A and JSPK

6 Conclusion

This study investigates the effects of bio-fuel properties on the evaporation rate and spray penetration using simple calculations obtained from FLUENT. The variation of properties, such as density, heat capacity, enthalpy of vaporisation, vapour pressure and binary diffusivity with temperature, are used. The effect of fuel volatility is found to influence evaporation rate whilst density is found to affect the propagation distance of a droplet in combustion chamber. Blend bio-fuel with kerosine becomes a solution to ensuring no modification has to be made within the combustion chamber.

The above analysis represents an initial step towards the identification of critical physical fuel characteristics that can be considered at a later stage as design parameters for the computational optimisation of bio-fuels for existing aircraft engines. For this purpose the development and validation of a performance evaluation computational tool for the combustion performance of bio-fuels as will be presented in this paper is essential.

Acknowledgements

During this PhD work, Nurul Musfirah Mazlan was fully sponsored by Universiti Sains Malaysia, Malaysia under Academic Staff Training Scheme (ASTS).

References

[1] Le Clercq P. and Aigner M., "Impact of Alternative Fuels Physical Properties on Combustor Performance," *11th Triennial International Annual Conference on Liquid Atomisation and Spray Systems (ICLASS)*, Vail, Colorado, USA, 2009, pp. 1-6.

[2] Lee, J. J., "Can We Accelerate the Improvement of Energy Efficiency in Aircraft Systems?", *Energy Conversion and Management*, Vol. 51, No. 1, 2010, pp. 189-196.

[3] Dagget D., Hadaller O., Hendricks R., and Walther R., "Alternative Fuels and Their Potential Impact on Aviation (ICAS – 2006 – 5.8.2)," *25th Congress of the International Congress of the Aeronautical Sciences (ICAS)*, Hamburg, Germany, 2006, pp. 1 - 8.

[4] Urgya-Bugajska I., Pourkashanian M., Borman D., Catalanotti E., and Wilson C. W., "Theoretical Investigation of the Performance of Alternative Aviation Fuels in an Aero-engine Combustion Chamber," *Aero-engine Combustion Chamber, European Air and Space Conference*, Manchester, UK, 2009, pp. 1 - 9.

[5] Kinder J. D., "Evaluation of Bio-derived Synthetic Paraffinic Kerosines (Bio-SPKs) Report Version 5.0," The Boeing Company, UOP United States Air Force Research Laboratory, 2010.

[6] Rahmes T. F., Kinder J. D., Henry T. M., Crenfeldt G., LeDuc G. F., Zombanakis G P., Abe Y., Lambert D M., Lewis C., Juenger J A., Andac M G., Reilly K R., Holmgren J R., McCall M J., Bozzano A G., "Sustainable Bio-derived Jet Fuel Flights and Engine Tests Program Results," *9th AIAA Technology, Integration, Aviation and Operations Conference,* Hilton Head South, Carolina, 2009, pp. 1-19.

[7] Ghassemi H., Baek S W., and Khan Q S., "Experimental Study on Evaporation of Kerosine Droplets at Elevated Pressures and Temperatures", *Combust. Sci. And Tech*, Vol. 78, 2006, pp. 1669 - 1684.

[8] FLUENT 6.3 User's Guide, Fluent. Inc, 2006

[9] Morsi S A., Alexander A J., "An Investigation of Particle Trajectories in Two-Phase Flow Systems", *J. Fluid Mech*, Vol. 55, No. 2, 1972, pp. 193-208.

[10] Frohn A. and Roth N., *Dynamics of Droplets*, 2nd ed., Springer – Verlag Berlin, 2000, Chaps. 7, 14.

[11] Khasanshin T S. and Aleksandrov A A., "Thermodynamic Properties of Ethanol at Atmospheric Pressure", *Journal of Engineering Physics and Thermophysics* Vol. 47, No. 3, 2001, pp. 1046 - 1052.

[12] Rochaya D., "Numerical Simulation of Spray Combustion Using Biomass Derived Liquid Fuels," Ph.D. Dissertation, School of Mechanical Engineering, Cranfield University, UK, 2007.

[13] Chin J S. and Lefebvre A H., "The Role of the Heat-up Period in Fuel Drop Evaluation", *International Journal of Turbo and Jet Engines*, Vol. 2, No. 4, 2001, pp. 315 - 326.

[14] Reddemann M A., Mathieu F., Martin D. and Kneer R., "The Influence of Fuel Properties on Spray Propagation, Atomisation and Evaporation (ILASS)," *23rd Annual Conference on Liquid Atomisation and Spray Systems*, Brno, Czech Republic, 2010, pp. 1-6.

[15] Sazhin S S., Feng G. and Heikal M R., "A Model for Fuel Spray", *Fuel*, Vol. 80, 2001, pp. 2171 - 2180.

[16] Lacava P T., Bastos-Netto D. and Pimenta A P., "Design Procedure and Experimental Evaluation of Pressure Swirl Atomizer," *24th International Congress of the Aeronautical Sciences*, 2009, pp. 1-9.

[17] Dikshit, S., Channiwala, S., and Kulshershtha, D., "Experimental Investigations of Performance Parameters of Pressure Swirl Atomizer for Kerosene Type Fuel," *ASME Turbo Expo 2009: Power for Land, Sea, and Air,* Orlando, Florida, US, 2009, pp 61-78.

[18] Fuller, E. N., Schettler, P. D., and Giddings, J. C., " A New Method for Prediction of Binary Gas-Phase Diffusion Coefficients, *Industrial and Engineering Chemistry,* Vol. 58, No. 5, 1966, pp. 18-27.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Boundary Element Model for Piezoelectric Dynamic Strain Sensing of Cracked Structures

A. Alaimo[1], G. Davì[2]
[1]*University of Enna Kore, Italy,* [2]*University of Palermo, Italy*

A. Milazzo[3], C. Orlando[4]
[3]*University of Palermo, Italy,* [4]*University of Palermo, Italy*

Keywords: *Piezoelectric Patch, Structural Health Monitoring, Fracture Mechanics, Boundary Element Method, Transient Analysis.*

Abstract

A piezoelectric sensor model is here presented for the Structural Health Monitoring (SHM) of damaged structures. A boundary element approach based on the Dual Reciprocity BEM is then used to model and analyze the transient response of a piezoelectric patch bonded on a cracked beam. The BE model is written for the piezoelectric problem employing generalized displacements. The multidomain boundary element technique is implemented to model non-homogeneous and cracked configuration, taking contact conditions into account. Analyses have been carried out for an isotropic beam with a piezoelectric strip attached on it and dynamical results for the undamaged structure have been compared with FE results showing good agreement. Parametric analysis is performed with the aim to characterize the sensitiveness of the piezoelectric patch by varying the crack positioning along the beam span. More particularly, seven electrodes are considered on the piezoelectric patch and the electrical response of the SHM piezoelectric sensor is characterized by means of a damage index in order to locate the crack.

1 Introduction

Piezoelectric media represent a class of smart materials able to convert mechanical energy into electric one and vice versa. More particularly, this kind of material exploits the direct piezoelectric effect to produce an electric field under applied mechanical strain and the converse piezoelectric effect to induce mechanical displacement when an electric voltage is applied [1]. These properties make the piezoelectric materials very suitable in the field of sensing and actuating technology for practical applications in aerospace, mechanical, electrical, civil and biomedical engineering [2, 3]. In the field of aerospace industry, a very appealing applications of piezoelectric sensing and actuating devices are represented by the active vibration control of structures [4, 5] and the Structural Health Monitoring [6-9]. Focusing on the Structural Health Monitoring (SHM), it can effectively be implemented in the framework of the so called closed-loop smart

structures, where components are able to sense, diagnose and actuate in order to perform their functions [10]. In particular, the main issue of the (SHM) is represented by the detection, location and assessment of the extent of a damage in such a way the health state and the remaining life of a structure can be estimated. Errors of type I, i.e. false positives, or type II, i.e. false negatives can occur in SHM systems, thus the analysis of the optimal sensor/actuator location as well as of their electromechanical response is mandatory. As already mentioned, SHM systems take often advantages of piezoelectric materials, among others, to realize the sensing network. The general methodology for detecting damage in structures is to extract meaningful features from the measured data in order to analyze the changes in sensor output due to damage. It is thus important to obtain an accurate and reliable characterization of the effect of stress field intensification generated by the crack on the piezoelectric output voltage. Accordingly, in this paper parametric analyses are performed in order to characterize the sensitiveness of a piezoelectric patch by varying the crack positioning along the damaged beam. A boundary element approach is then used to model and analyze the transient response of a piezoelectric based structural health monitoring system for cracked beam. The BE model is written for the piezoelectric problem employing generalized displacements. The dual reciprocity technique is used to express the mass matrix in terms of boundary unknowns only. The multidomain boundary element technique is implemented to model non-homogeneous and cracked configuration taking unilateral interface conditions into account. A damage index, characterizing the electric response of the electrodes for the damaged configuration with respect to the un-cracked ones, is defined in order to locate the crack and the results obtained evidence the effectiveness of the BE piezoelectric sensor model for SHM applications.

2 BI Representation

A dual reciprocity boundary integral procedure is proposed for piezoelectric media under the hypotheses of two dimensional piezoelectricity and plane strain conditions. Since the electric waves propagate several order faster than the elastic one, the electric field is here considered to be as quasi-static. Introducing electromechanical generalized variables [11, 12], the governing equations of the piezoelectric problem can be written in Navier-like form and one has

$$\boldsymbol{\mathcal{D}}^T \mathbf{R} \boldsymbol{\mathcal{D}} \mathbf{U} = \boldsymbol{\rho} \ddot{\mathbf{U}} \qquad (1)$$

where $\mathbf{U} = [u_1 \quad u_2 \quad u_3 \quad \varphi]^T$ is the generalized displacement vector, collecting the mechanical displacements u_i and the electric potential φ, $\mathbf{R}$ is the generalized stiffness matrix and $\boldsymbol{\mathcal{D}}$ the generalized differential operator defined as in Davì and Milazzo [12]. Assuming quasi-static electric field, the right hand side of Eq. (1) is representative of the inertial force components only. By applying the Betti's reciprocity theorem with the static electro-elastic fundamental solutions, the Somigliana identity for the electromechanical problem is obtained in terms of generalized variables [12],

$$\begin{aligned} &\mathbf{c}^*(P_0)\mathbf{U}(P_0) + \\ &+\int_{\partial\Omega} \left(\mathbf{T}^*(P,P_0)\mathbf{U}(P) - \mathbf{U}^*(P,P_0)\mathbf{T}(P)\right) d\partial\Omega = \\ &= -\int_{\Omega} \mathbf{U}^*(P,P_0)\boldsymbol{\rho}\ddot{\mathbf{U}}(P) d\Omega \end{aligned} \qquad (2)$$

where $\mathbf{T}(P) = [t_1 \quad t_2 \quad t_3 \quad D_n]^T$ is the vector of the generalized tractions acting at the field point P and collecting the mechanical traction t_i and the normal component of the electric

displacement D_n, while the kernels $\mathbf{U}^*(P,P_0)$ and $\mathbf{T}^*(P,P_0)$ are

$$\mathbf{U}^*(P,P_0)=\left[U_{ij}(P,P_0)\right]^T \qquad (3)$$

$$\mathbf{T}^*(P,P_0)=\left[T_{ij}(P,P_0)\right]^T$$

being $U_{ij}(P,P_0)$ and $T_{ij}(P,P_0)$ the ith component of the generalized displacements and tractions of the jth fundamental solution. The domain integral in the right hand side of Eq. (2) represents the inertial term which contains the unknown acceleration $\ddot{\mathbf{U}}(P)$ inside the domain. In order to transform the domain integral into a boundary one the dual reciprocity technique is used [13]. The piezoelectric dynamic problem is then solved numerically by means of the Boundary Element Method in such a way the following equations of motion are obtained

$$\mathbf{M}\ddot{\boldsymbol{\delta}}+\mathbf{H}\boldsymbol{\delta}=\mathbf{GP} \qquad (4)$$

where $\boldsymbol{\delta}=\left[\boldsymbol{\delta}_B \quad \boldsymbol{\delta}_D\right]$ is the vector of generalized nodal displacements eventually including internal nodal displacements $\boldsymbol{\delta}_D$ deriving from a second set of equation obtained by writing the Dual Reciprocity Boundary integral representation at selected internal nodes. The vector $\mathbf{P}$ of Eq. (4) represents the generalized boundary nodal traction and $\mathbf{M}$ is the mass matrix, approximated through the Dual Reciprocity Method. The time integration of the equation of motion, Eq. (4), is performed by means of the Houbolt method [14] which allows to approximate the acceleration $\ddot{\boldsymbol{\delta}}$ at the instant $t+\Delta t$ as

$$\ddot{\boldsymbol{\delta}}^{t+\Delta t}=\frac{1}{\Delta t^2}\left(2\boldsymbol{\delta}^{t+\Delta t}-5\boldsymbol{\delta}^{t}+4\boldsymbol{\delta}^{t-\Delta t}-\boldsymbol{\delta}^{t-2\Delta t}\right) \qquad (5)$$

3 Host Structure-Patch assembling Contact condition modeling.

The overall assembled structure shown in Fig. (1) represents an inhomogeneous domain that needs to be divided into homogeneous sub-regions in order to apply the dual reciprocity boundary element method. The multidomain technique is then used to obtain the numerical model of the problem.

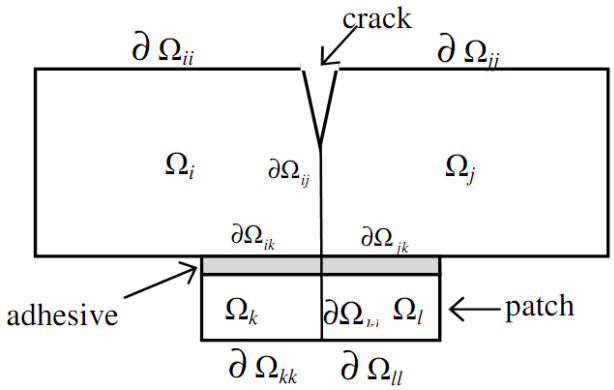


Fig. 1 Multidomain assembling between the host structure and the patch.

More particularly, indicating with M the number of sub-region considered, the equation of motion for each homogeneous domain, by means of the Houbolt approximation scheme, writes

$$\mathbf{M}^{(k)}\ddot{\boldsymbol{\delta}}^{(k)}+\mathbf{H}^{(k)}\boldsymbol{\delta}^{(k)}=\mathbf{G}^{(k)}\mathbf{P}^{(k)} \qquad \mathbf{k}=1,\ 2...,\ M \qquad (5)$$

The global system of equation pertaining the overall assembled structure is then obtained by applying the compatibility and equilibrium conditions along all the sub-regions interfaces

$$\boldsymbol{\delta}^{(i)}_{\partial\Omega_{ij}}-\boldsymbol{\delta}^{(j)}_{\partial\Omega_{ij}}=\Delta\boldsymbol{\delta}^{ij} \qquad i=1,...,N-1$$
$$\mathbf{P}^{(i)}_{\partial\Omega_{ij}}=-\mathbf{P}^{(j)}_{\partial\Omega_{ij}} \qquad j=i+1,...,N \qquad (6)$$

where the subscript $\partial\Omega_{ij}$ indicates quantities associated with the nodes belonging to the

interface between the *ith* and *jth* sub-regions. Eq. (6) allows to assemble heterogeneous sub-regions by considering both rigidly and elastic connection among them. More particularly, the first condition is obtained by forcing the displacement jump $\Delta\boldsymbol{\delta}^{ij}$ to be equal to zero, thus considering no relative displacements between interface coincident nodes belonging two different subregions. Conversely, the elastic interface condition, that allows to model the adhesive layer between the host structure and the piezoelectric patch, is here addressed by means of an interface Spring Model [15]. It allows approximating the adhesive ply as a zero thickness elastic layer characterized by normal and tangential compliance constants, k_N and k_T, with respect to a local interface coordinate systems, as schematically shown in Fig. (2). Thus, it is possible to consider relative displacements between the bi-material contiguous surfaces whose components in normal and tangential directions can be directly linked to the associated interface tractions by means of the compliance constants

$$\begin{aligned} \Delta\boldsymbol{\delta}_N &= k_N \mathrm{P}_N \\ \Delta\boldsymbol{\delta}_T &= k_T \mathrm{P}_T \end{aligned} \qquad (7)$$

The interested reader can find a detailed description on the implementation of the Spring Model in the framework of the multidomain BEM in the work of Alaimo, Milazzo and Orlando [15].

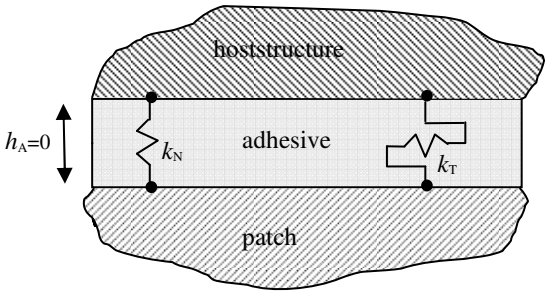


Fig. 2 Elastic zero thickness interface.

Here, the crack modeling strategy is briefly discussed. It is considered as a zero thickness elastic layer by taking the compliance coefficients k_N and k_T as penalty factors [16], whose values allows to characterize the opening or the contact conditions that the crack surfaces can experience. More particularly, the first condition implies no peel and shear tractions at the interface, corresponding to an elastic interface characterized by zero stiffness along both the normal and tangential directions. On the contrary, the frictionless contact involves no interpenetration between the crack surfaces,

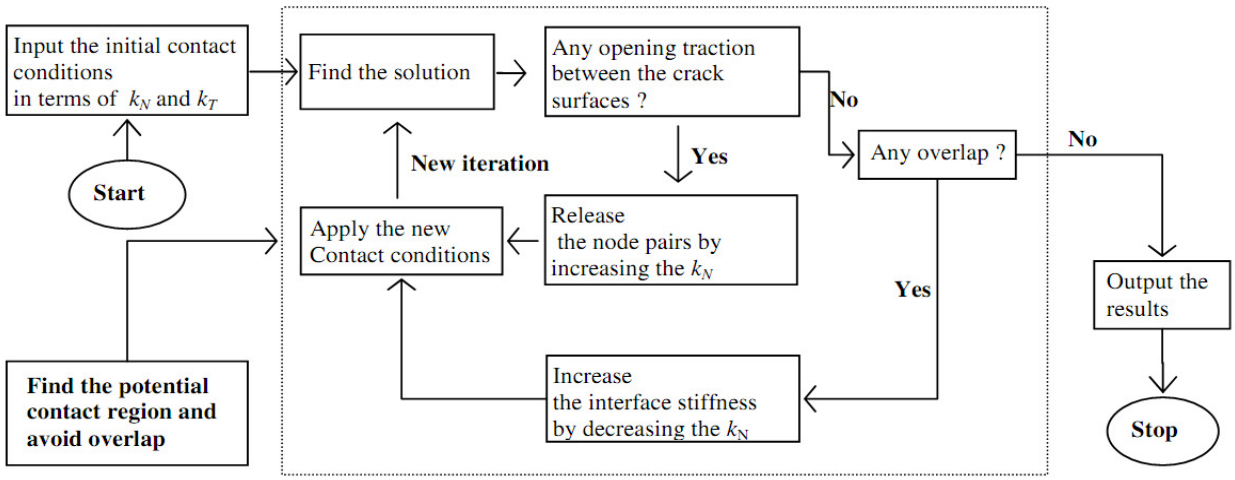


Fig. 3 Iterative solution scheme.

condition that can be achieved through an infinite stiffness elastic interface. From the numerical point of view, the penalty factors are iteratively modified according to the compatibility and equilibrium conditions characterizing both the opening and the contact behaviors of the crack. A schematic of the iterative procedure is given in Fig. (3).

4 Numerical Model and Results

In this section the dynamic boundary element approach for piezoelectric domain previously presented is employed to model and analyze the sensing effectiveness of a piezoelectric based structural health monitoring system for damaged beam-like structures. Firstly, the cracked structure and the piezoelectric sensing circuitry are presented. The time responses of the electrodes mounted on the piezoelectric patch are then analyzed and a damage index is defined. Last, the effect of distinct crack configurations on the sensing capabilities of the proposed structural health monitoring system is investigated in terms of the damage index response. An isotropic cantilever beam undergoing a step input load F and equipped with a piezoelectric patch attached on the bottom surface is analyzed. The beam presents an edge-crack of length a as shown in Fig. (4), where the geometry as well as the boundary conditions are also reported. The host beam has Young's modulus E=3.3 GPa and Poisson's ratio ν=0.35, while the material properties of the PZT4 sensor are those used in reference [15]. The sensing configuration is arranged by using seven electrodes, E_i (i=1,…,7), equally spaced as shown in Fig. (4). Following the work presented in reference [17], the piezoelectric sensors are modeled as charge sources $Q_P(t)$, where t is the time variables, with a shunt capacitor C_p and a resistor R_p connected to a charge amplifier, see Fig. (5).

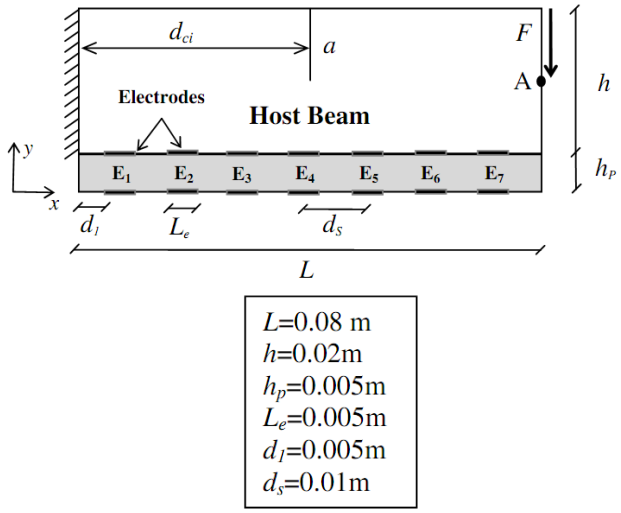


Fig. 4 Damaged beam and monitoring arrangements.

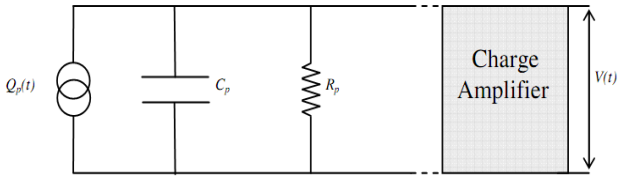


Fig. 5 Piezoelectric sensor circuitry.

By so doing, the output voltage $V(t)$ is written as [18]

$$V(t) = G\frac{dQ_p(t)}{dt} \qquad (8)$$

being G the charge amplifier constant gain, while the piezoelectric generated charge is computed as the integral on the electrodes surface of the electric displacement normal component D_n. The dynamic response of the vertical displacements of point A, corresponding to the point load F, have been firstly computed and compared to those obtained through a finite element analysis and good agreement has been found as evidenced by Fig. (6). The electric current time response of electrode E_3 is shown in Fig. (7) for undamaged and cracked beam. The presence of the crack is

only evidenced, in the time domain, by a change in the natural frequency of oscillation of the structures, as the crack induces an overall stiffness reduction. It is here stressed that the time response of all the electrodes are analogues to the one shown in Fig. (7) and thus no information on crack location are straightforwardly available from the time responses of the electrodes.

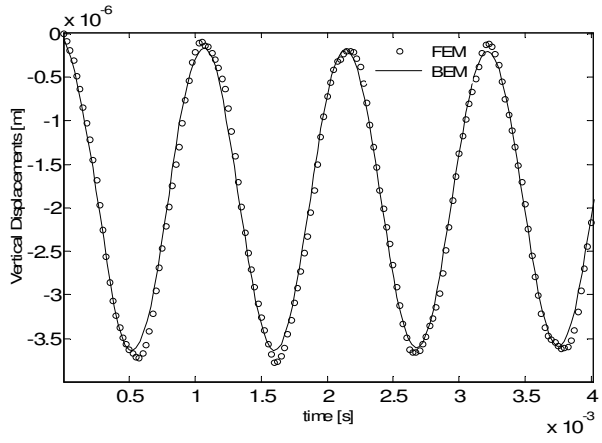


Fig. 6 BE and FE vertical displacement time history.

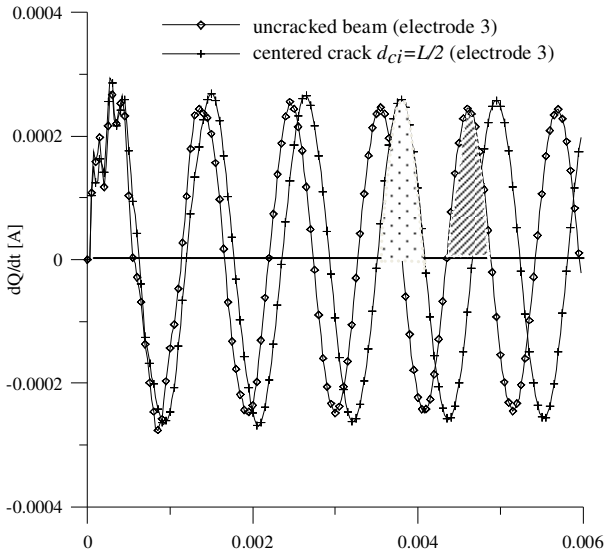


Fig. 7 Electric current generated by electrode E_3.

It is expected that the crack causes a local strain change with respect to the undamaged case that originates from the stress field intensification near the crack tip or it is simply due to a local discontinuity in the beam slope crossing the cracked section. Moreover, from the piezoelectric constitutive relationships it stems that the free charge D_n is proportional to the local strain. Thus, a damage index I_D is defined in order to catch the crack influence on the piezoelectric generated charge along the beam length. More particularly, the damage index for each electrodes is defined as

$$I_{\mathcal{D}} = \frac{Q_{E_i}^{cracked} - Q_{E_i}^{undamaged}}{Q_{E_1}^{undamaged}} \qquad (i = 1,...,7) \qquad (9)$$

where $Q_{E_i}^{undamaged}$ is representative of the charge generated at the electrode E_i by the undamaged structure and is computed as the integral over half the natural period of the current steady-state signal, evidenced in Fig. (7) by the hatched area. On the other hand, $Q_{E_i}^{cracked}$ is the charge generated at the same electrode when the beam is cracked and is represented in Fig. (7) by the dotted area. Four crack positions have been analyzed, namely d_{ci} . The damage indices at each electrodes have been computed for every crack positions and are reported in Fig. (8), where the maximum value of the dynamic stress intensity factors are also shown and the cracks and electrodes locations are graphically evidenced. It is worth noting that the damage index always shows a maximum at the electrode that is closest to the crack, moreover the value of the damage index appears to decrease as the maximum stress intensity factors K_{max} decrease. It follows that the through the length distribution of the proposed damage index allows to locate the crack and also gives information about stress intensification introduced by the damage.

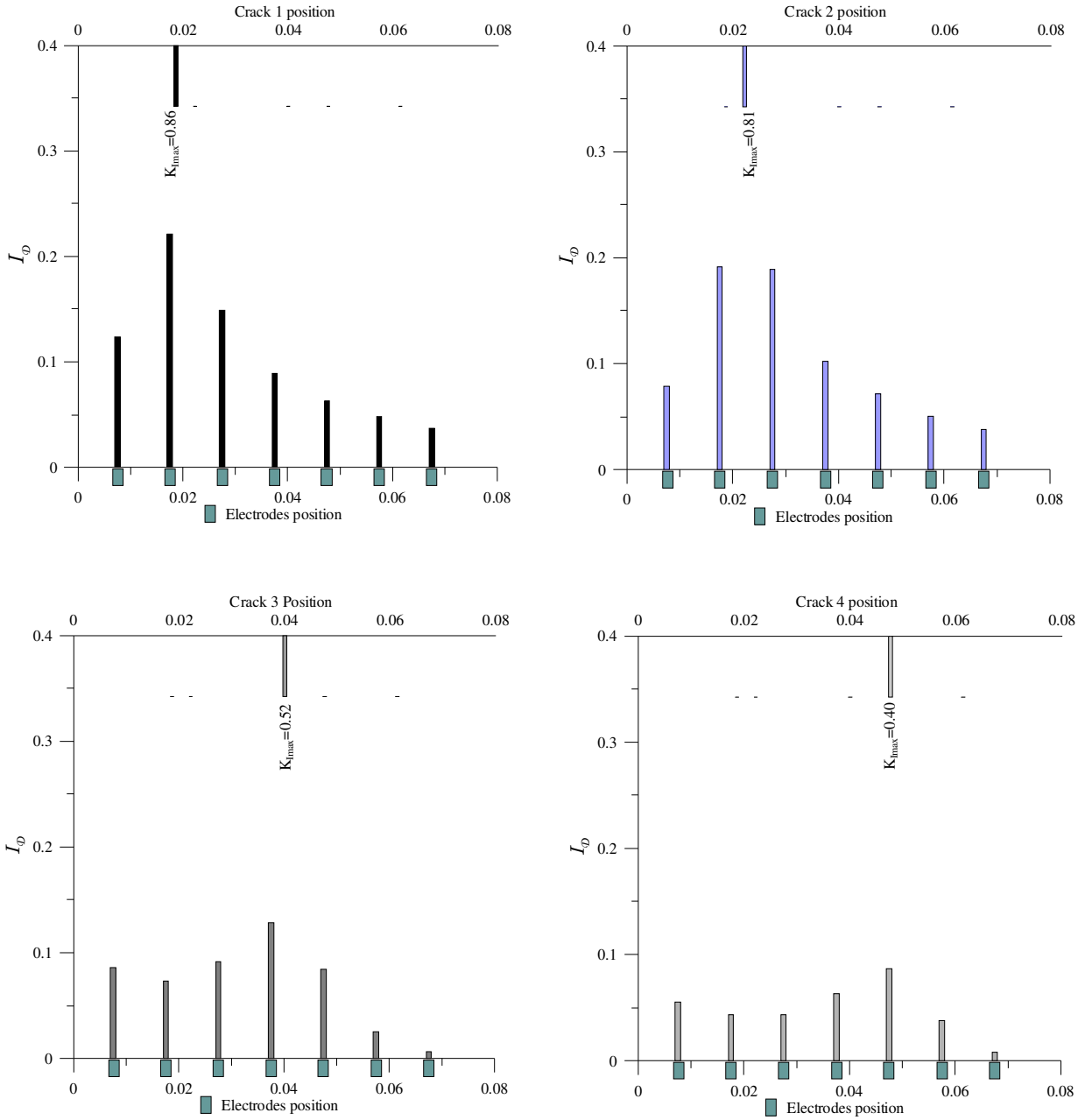


Fig. 8 Damage index for all the crack configurations.

5 Conclusions

In this paper a two-dimensional dual reciprocity boundary element model for piezoelectric problem has been presented. The proposed BE approach reveals as an effective numerical tool to model and study structural health monitoring systems based on the piezoelectric dynamic strain sensing capability. Analyses have been carried out for an isotropic beam with a piezoelectric strip attached on it and dynamical results for the undamaged structure have been compared with FE results showing good agreement. The influence of the crack location on the piezoelectric based structural health monitoring system response has been studied. It has been found that the crack presence influences the sensors time responses by changing the natural frequency of vibration of

the beam. A damage index definition has then been proposed and used to highlight that the surface mounted piezoelectric patches are capable of sensing the beam through the length strain changes between the damaged and undamaged case. In conclusion, it has been shown that results obtained in terms of the damage index can be used to locate the crack along the beam and to characterize the related amount of stress intensification introduced.

References

[1] I. Chopra, *Review of State of Art of Smart Structures and Integrated Systems*, AIAA Journal, 40 (10), 2145-87 (2002).

[2] T. Ikeda, "Fundamentals of piezoelectricity", New York: Oxford University Press, (1990).

[3] R. Smith, "Smart material systems: model development (Frontiers in applied mathematics; 32)", Philadelphia: SIAM, (2005).

[4] G. Gilardi, B.J. Buckham and E.J. Park, *Finite Element Modeling of a Slewing Non-linear Flexible Beam for Active Vibration Control with Arrays of Sensors and Actuators,* Journal of Intelligent Material Systems and Structures, 20(16): 1941-1958 (2009).

[5] Y.Y. Li, L. Cheng and P. Li, *Modeling and vibration control of a plate coupled with piezoelectric material,* Composite Structures, 62: 155–162 (2003).

[6] T. Monnier, *Lamb Waves-based Impact Damage Monitoring of a Stiffened Aircraft Panel using Piezoelectric Transducers*, Journal of Intelligent Material Systems and Structures, 17(5): 411-421 (2006).

[7] V. Giurgiutiu, *Tuned Lamb Wave Excitation and Detection with Piezoelectric Wafer Active Sensors for Structural Health Monitoring*, Journal of Intelligent Material Systems and Structures, 16(4): 291-305 (2005).

[8] D.M. Peairs, G. Park and D.J. Inman, *Improving Accessibility of the Impedance-based Structural Health Monitoring Method,* Journal of Intelligent Material Systems and Structures, 15(2): 129-139 (2004).

[9] T. Liu, V. Martin and K. Sritawat, *modelling the input-output behavior of piezoelectric structural health monitoring systems for composite plates,* Smart Materials and Structures, 12(5): 836-844 (2003).

[10] A.V. Srinivasan and D.M. McFarland, "Smart Structures (Analysis and Design)", Cambridge University Press (2001).

[11] D.M. Barnett, J. Lothe, *Dislocations and line charges in anisotropic piezoelectric insulators,* Physics State Solid (b), 67, 105-111 (1975).

[12] G. Davì, A. Milazzo, *Multidomain boundary integral formulation for piezoelectric materials fracture mechanics,* International Journal of solids and Structures, 38, 7065-7078 (2001).

[13] P.W. Partridge, C.A. Brebbia and L.C. Wrobel, The Dual Reciprocity Boundary Element Method, Computational Mechanics Publication, (1992).

[14] J.C. Houbolt, *A recurrence matrix solution for the dynamic response of elastic aircraft,* J. Aeronaut. Sci., 17, 540-550 (1950).

[15] A. Alaimo, A. Milazzo and C. Orlando, *Boundary Element Analysis of Adhesively Bonded Piezoelectric Active Repair,* Engineering Fracture. Mechanics, 76, 500-511 (2009).

[16] C. Renaud, and Z.Q. Feng, *BEM and FEM analysis of Signorini contact problems with friction*, Computational Mechanics, 31: 390-399, (2003).

[17] B. Lin, V. Giurgiutiu, *Modeling and testing of PZT and PVDF piezoelectric wafer active sensors*, Smart Materials and Structures. 15, 1085-1093 (2006).

[18] K.S.C. Kuang, S.T. Quek, W.J. Cantwell, *Use of polymer sensors for monitoring the static and dynamic response of a cantilever composite beam,* Journal of Material Science. 39, 3839-3843 (2004).

Future-oriented C4ISTAR applications. UAS versus manned aircraft. Merit analysis to consider performance/cost aspects

S. Chiesa and N. Viola
Politecnico di Torino, Italy

G.A. Di Meo and M. Fioriti
Politecnico di Torino, Italy

G. Mastrodomenico
Alenia Aeronautica S.p.A., Italy

Keywords: *UAS, Effectiveness/Cost, C4ISTAR*

Abstract

The aim of the work is to study in an economical point of view the possibility to employ UAS's (Unmanned Aircraft System) for surveillance missions typically performed by manned platforms. The reason is that UAS's concern performances (high endurance and operating altitude) suitable for surveillance missions at lower operating costs than manned aircraft. We considered the following mission profiles: AEW&C (Airborne Early Warning and Control) MPA (Maritime Patrol Aircraft) and AGS (Airborne Ground Surveillance). By considering operative requirements of each mission profile, the best unmanned platform has been preliminary designed and its DOC (Direct Operating Cost) has been evaluated by using software tools developed at Politecnico di Torino in collaboration with Alenia Aeronautica S.p.A. A comparison in an Effectiveness/Cost point of view has been conducted in order to demonstrate validity of unmanned solution versus manned one.

1 Introduction

Nowadays UAS's are successfully part of complex Systems of Systems (SoS) based on C4ISTAR (Command Control Communication Computing Intelligence Surveillance Target Acquisition and Reconnaissance) requirements. Their success in C4ISTAR applications is due to their high effectiveness in D^3 environments (Dull, Dangerous and Dirty). US Army Center of Excellence in the document *"US Army Unmanned Aicraft Systems Roadmap 2010-2035"* [9] asserts:

> Unmanned aircraft systems can provide three critical capabilities for the Army's current and future force. First, UAS reduce risks to Soldiers in the current fight (e.g. explosive hazard detection and neutralization). Second, UAS reduce the workload on Soldiers by performing routine missions and enable sustained high tempo operations (e.g., routine surveillance of forward operating bases). Third, UAS provide emerging capabilities for extended range or standoff reconnaissance operations.

These characteristics call for a further integration of UAS in C4ISTAR application where nowadays some roles are still performed by piloted platforms. In particular, this work is focused of the following C4ISTAR mission profiles:

- Airborne Early Warning and Control (AEW&C)

- Maritime Patrol Aircraft (MPA)
- Airborne Ground Surveillance (AGS)

For each mission profile we propose one or more UAS solutions in order to compare them with existing manned solutions through an Effectiveness/Cost analysis. Following sub-sections will analyze at first features of each application by presenting the state of the art.

1.1 Airborne Early Warning and Control (AEW&C)

An Airborne Early Warning and Control (AEW&C) system is an airborne radar system designed to detect hostile aircraft and missiles. An older term to designate AEW&C is AWACS (Airborne Warning and Control System).

AEW&C systems are typically jet liners converted by installing a radar system upon the fuselage. Examples are E-3 Sentry, B737 Wedgetail and B767 AWACS. There exists also regional turboprops converted for AEW&C purposes but their performances are poorer than jet platforms and they find application where there is not a wide airspace to survey. *No UAS performing AEW&C missions exists*.

We want to hypothesize and to compare two surveillance systems: the first composed of MALE (Medium Altitude Long Endurance) UAS's and the second composed of HALE (High Altitude Long Endurance) UAS's designed for AEW&C missions. A MALE typically has an operating altitude between about 8 to 13 Km, an HALE instead flies between about 15 and 19 Km. It is interesting to analyze if it is better to have a higher number of MALE UAS or a lower number of HALE UAS in an Effectiveness/Cost point of view.

1.2 Maritime Patrol Aircraft (MPA)

A Maritime Patrol Aircraft is a platform equipped to survey maritime space in order to detect hostile or illegal ships and submarines. In particular there exist two functionalities: MP (Maritime Patrol) and ASW (Anti-Submarine Warfare). The first one consists in detecting just hostile ships by using typically a SAR (Synthetic Aperture Radar) for detection and EO/IR (Electro Optic Infrared) for reconnaissance. The second one is more advanced, it comprises MP functionalities but in addition there is the possibility to detect submarines, by using MAD sensor (Magnetic Anomaly Detector), and to provide counter-attack by using missiles or torpedoes. Modern MPA platforms are typically manned regional turboprops, examples can be ATR 42 MP, ATR 72 MP and ATR 72 ASW. Even if UAS's are not typically used for MPA missions, there exist examples as MQ-4C Global Hawk which is an HALE platform equipped for broad maritime surveillance.

We want to analyze the case of a constellation of MALE platforms designed to perform MP functionality over an assigned maritime area. By analyzing its Effectiveness and its Costs, it will be interesting to make a comparison with existing manned platforms.

1.3 Airborne Ground Surveillance (AGS)

Airborne Ground Surveillance is a mission profile which consists in the surveillance of territory in order to detect possible menaces moving on the ground. Sensors used are SAR/ISAR and EO/IR. AGS platforms are typically unmanned HALE (e.g. RQ-4 Global Hawk) or MALE (e.g RQ-1 Predator). The most relevant case of AGS manned platform has been E-8 Joint Star but nowadays development of AGS aircraft is UAS-oriented.

We want to suppose a constellation of UAS MALE AGS aircraft and by analyzing Effectiveness/Cost aspects it is easy to show that even if an AGS manned aircraft can assure a good effectiveness its related costs justifies the choice to develop AGS on UAS platforms.

2 UAS preliminary design

A first step of our work is to provide preliminary design of UAS starting from surveillance requirements identification. Surveillance requirements define the technical characteristics of mission to be performed as total area to be surveyed, distance between top of climb and patrol zone (i.e mission radius). We established these requirements by hypothesizing the surveillance of Italian nation.

Table 1 Surveillance requirements

	AEW&C		*MPA*	*AGS*
Area to survey	942157 Km^2		313157 Km^2	31416 Km^2
UAS Category	MALE	HALE	MALE	MALE
Mission Radius	0 Km	0 Km	150 Km	0 Km
Operating Altitude	7,92 Km to 13,72 Km	15,24 Km to 19,81 Km	7,92 Km to 13,72 Km	7,92 Km to 13,72 Km
UAS Endurance	32 h to 40 h	24 h to 36 h	32 h to 40 h	32 h to 40 h

Furthermore, AEW&C is supposed to patrol the whole national airspace, MPA is supposed to patrol Southern Italian coasts and AGS is supposed to provide surveillance on Rome during major events. Operating Altitude and Endurance have been decided on the base of MALE or HALE typical performances.

It is possible to divide the whole design-process in two phases:

1) From surveillance requirements to aircraft requirements
2) From aircraft requirements to aircraft preliminary design

It is to notice that several aircraft requirements are able to satisfy defined surveillance requirements. As a consequence, different aircraft preliminary designs are able to satisfy imposed surveillance requirements. As an example, it is possible to perform an Airborne Early Warning surveillance mission on an assigned area by employing a high number of light and small UAS's flying at low altitude or a smaller number of heavier and bigger UAS's flying at higher altitude. Merit evaluation of every solution is the driver for the choice of the best solution.

2.1 From surveillance requirements to aircraft requirements

Starting from surveillance requirements defined in Table 1 it is possible to obtain a discrete set of aircraft requirements so that it is possible to define in a further step aircraft preliminary design of UAS platforms. A tool developed in Matlab® programming language has been used to provide calculations.

It is important to define relationships able to link surveillance requirements with aircraft requirements. Area to survey (SA), number of aircraft ($N°_{assets}$), operating altitude (h), endurance (End) and loiter speed (V_{ltr}) are linked each other by the Eq. (1) and Eq. (2).

$$SA = N°_{assets} \cdot D_{HOR} \cdot V_{ltr} \cdot End \qquad (1)$$

$$D_{HOR} = \sqrt{2 \cdot R_E \cdot h} \qquad (2)$$

D_{HOR} represents the distance of visual horizon and R_E represents the radius of the earth. Equation (1) is valid for AEW&C missions, for MPA and AGS missions D_{HOR} is to be substituted with RR_{SAR} which represents the range of SAR radar. Range of AEW&C radar is supposed to be equal to D_{HOR}, EO/IR sensors are dimensioned by considering to have NIIRS (National Image Interpretability Rating Scales) defined in Table 2 at the operating altitude and considering the target to be 10 Km distant.

Table 2 NIIRS defined for EO/IR dimensioning

EO/IR mode	NIRS
EO Spotter mode	8
EO Wide FOV mode	3
EO Narrow FOV mode	6
IR Wide FOV mode	3
IR Narrow FOV mode	7

2.2 From aircraft requirements to preliminary design

Surveillance requirements allow defining the mission profile which is at the base of preliminary design process. A statistical basis about UAS has been created by using *Jane's Unmanned aerial vehicles and targets* [7] as a data source. Methodology proposed by *Dr. Roskam J.* [3] and [4] has been used for preliminary sizing. A statistical base has been created also about mission avionics by using *Jane's Avionics* [6]. Furthermore, by defining NIIRS performance of EO/IR sensors and maximum range of AEW&C radar and SAR

radar it is possible to determine weight, electrical power absorption and dimensions of these equipments.

UAS database and Roskam methodology has been implemented in a Matlab® tool able to define a preliminary configuration of an UAS starting from aircraft requirements. It is to notice that surveillance requirements concern several sets of aircraft requirements and so several UAS preliminary designs. The best preliminary sizing has been chosen by employing a merit approach.

The aim of the approach is to evaluate merit of every surveillance solution in order to establish maximum merit solution. Parameters involved in merit evaluation have been selected and normalized. Merit of every single parameter has been evaluated by employing merit functions that are relations associating a value from 0 to 1 to a normalized parameter value. Merit functions associate 0 if the parameter assumes the minimum merit value or 1 if it assumes maximum merit value.

It is possible to calculate global merit of the whole surveillance solution by using Eq. (3) where K and a_i are defined in Eq. (4) as reference [8].

$$U(x) = \frac{1}{K}\left\{\prod_{i}^{n}\left[Ka_i U_i(x)+1\right]-1\right\} \qquad (3)$$

$$1+K = \prod_{i=1}^{n}\left(1+Ka_i\right) \qquad (4)$$

U(x) represents global merit of the surveillance solution, $U_i(x)$ represents merit related to the parameter "i", a_i is the relative importance coefficient related to parameter "i", K is a scaling constant calculated by using Eq (4). For further information on used methodology refer to the paper *System of System(SoS) project and optimization in a Network Centric Operations (NCO) environment* [1].

2.3 UAS results

Previously described methodology lead to preliminary configurations showed in Table 3.

As far as AEW&C mission is concerned it is possible to see, as already said, that the choice is between three MALE UAS and two heavier HALE UAS with a more performing radar system. So even if flight at higher altitude is convenient because of the reduced number of platforms it is to be taken into account that sensors have to be more performing, heavier and more power consuming, as a result the single platform has to be heavier and bigger and so more expensive.

As far as MPA mission is concerned it is possible to notice that Southern Italian coasts can be patrolled just by four light MALE UAS. In order to make comparisons with manned platforms it is to notice that sensors of these platforms are not so performing as typical sensors installed on manned MPA's, so it is to be taken in account that even if manned platforms are more expensive it is necessary a lower number of them to survey the same area.

As far as AGS mission is concerned it is possible to see that one light MALE UAS is sufficient for surveillance on a big town during great events. Even in this case it is to notice that for making comparisons with manned platforms same surveyed area is to be considered.

Table 3 UAS results compliant with imposed surveillance requirements.

	AEW&C		*MPA*	*AGS*
UAS cathegory	HALE	MALE	MALE	MALE
Surveillance parameters				
N° of aircraft	2	3	4	1
Operating Altitude [m]	17520	12270	12270	12200
Endurance [h]	29	41	41	38,8
Aircraft parameters				
MTOW [Kg]	59201	43671	3126	1987
OEW [Kg]	23384	16463	1595	1019
Fuel Weight [Kg]	30979	22722	1335	794
Payload [Kg]	4537	3584	187	173

Avionic sensors				
Sensor	AEW Radar	AEW Radar	SAR Radar	SAR Radar
Weight [Kg]	4237	3134	34	20
Range [Km]	472	390	158	100
Power absorption [W]	Not Available	Not Available	650 W	400 W
Antenna Dimensions [mm]	Length 15100	Length 12200	H400 D310 W150	H310 D300 W110
Sensor			EO/IR	EO/IR
Weight [Kg]			51,4	51,4
Power absorption [W]			650	650
Dimensions [mm]			Diam. 400 Height 500	Diam. 400 Height 500

In the following paragraph comparisons in terms of Effectiveness and Costs will be performed.

3 Effectiveness/Cost Analysis

Effectiveness/Cost analysis is a useful tool able to show validity of proposed solutions versus existing manned ones in an economical point of view. Effectiveness analysis assesses an overall performance of our UAS solutions compared to existing manned platforms on the market. By evaluating direct operating costs, it will be possible to assess which are advantages/disadvantages of considered aircraft-systems in a tight economical point of view. In the end, it is possible to assess "performance price" by combining these two analyses.

Effectiveness/Cost analysis has been conducted both for our UAS solution and for existing manned platform chosen as reference. E-3 Sentry has been chosen as a reference for AEW&C missions because it is the most employed platform on the market and because E-3 Sentry fleet is becoming ancient and so it is reasonable to think to replace aircraft. ATR 72 MP has been chosen as reference platform for MPA mission because it is a good example of a modern maritime patrol aircraft. E-8 Joint Stars has been chosen has reference for AGS missions because it is nearly the unique example of manned aircraft of this category.

3.1 Effectiveness Analysis

Effectiveness of an aircraft is a unique number, which expresses an overall performance assessment. The first phase of Effectiveness analysis concerns to select the most significant parameters for every surveillance mission point of view among all aircraft performance parameters. Following tables shows values of selected parameters for considered aircraft. For UAS, Crew is intended the number of operators working on the ground station.

Table 4 AEW&C performance parameters

	AEW&C		
	E-3 Sentry	UAS HALE Solution	UAS MALE Solution
Avionic mission system	AN-APY-2	AEW&C Radar	AEW&C Radar
Service Ceiling [m]	11855	17520	12270
Endurance [h]	11,4	29	41
Typical Mission Range [nm]	5000	10863	4724
TO field lenght [m]	3054	2000	700
Max Cruise Speed [m/s]	270	207	96
Crew	17	10	10

Table 5 MPA performance parameters

	MPA	
	ATR 72 MP	UAS MALE Solution
Avionic mission system	SAR: Thales Ocean Master 400 EO/IR: Galileo EOST 23	SAR and EO/IR
Service Ceiling [m]	7260	12270
Endurance [h]	9	41
Typical Mission Range [nm]	1700	3462
TO field lenght (ISA, MTOW,SL) [m]	1290	800

Max Cruise Speed [m/s]	142	68
Crew	5	3

Table 6 AGS performance parameters

	AGS	
	E-8 Joint	UAS MALE Solution
Avionic mission system	AN-APY-3	SAR and EO/IR
Service Ceiling [m]	12800	12200
Endurance [h]	11	38,8
Typical Mission Range [nm]	5000	3367
TO field lenght (ISA, MTOW,SL) [m]	3054	600
Max Cruise Speed [m/s]	270	72
Crew	20	3

By normalizing each parameter, it is possible to evaluate merit of a single parameter in a linear scale where 0 is a parameter with minimum merit and 1 is a parameter with maximum merit. It is possible to evaluate global effectiveness of a platform by using Eq. (3).

3.1.1 Effectiveness Results

Results of Effectiveness analysis are presented in the following paragraphs.

3.1.1.1 AEW&C

Figure 1 shows results of Effectiveness analysis for AEW&C platforms. It is possible to notice that both UAS platforms have a higher effectiveness than manned one due to the higher endurance and operating altitude. In particular, MALE UAS can survey an area comparable with E-3 Sentry because they have similar operating altitudes. Then, a MALE UAS has a lower maximum speed than manned AEW&C because MALE UAS's are typically equipped with turbocharged piston engine and E-3 Sentry is equipped with jet engines. MALE UAS has a higher effectiveness than E-3 Sentry due to the very higher endurance which is a relevant parameter in a surveillance mission. HALE UAS has a higher effectiveness than MALE UAS due to the higher operating altitude and the higher maximum speed that allows surveying a wider area.

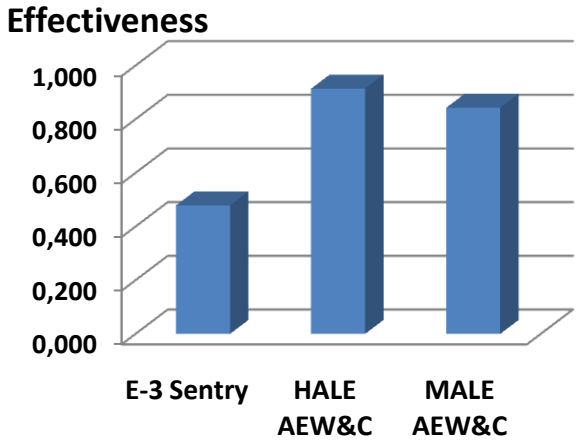


Fig. 1 Effectiveness results for AEW&C mission

3.1.1.2 MPA

Figure 2 shows Effectiveness results for MPA platforms. MALE UAS has a higher effectiveness due to its higher endurance performance and range.

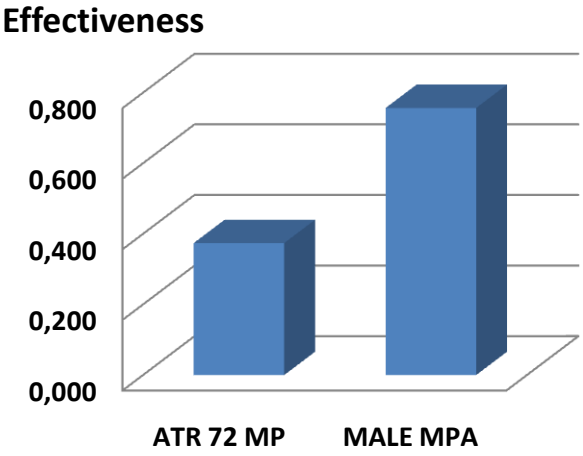


Fig. 2 Effectiveness results for MPA mission

3.1.1.3 AGS

Figure 3 shows Effectiveness results for AGS platforms. MALE UAS and E-8 Joint present similar results because they has similar operating altitude and even if MALE UAS offers an higher endurance and it requests a small number of operators (i.e. operators on the ground control station), E-8 Joint Stars is

equipped with a more performing sensor suite and it has a higher maximum speed.

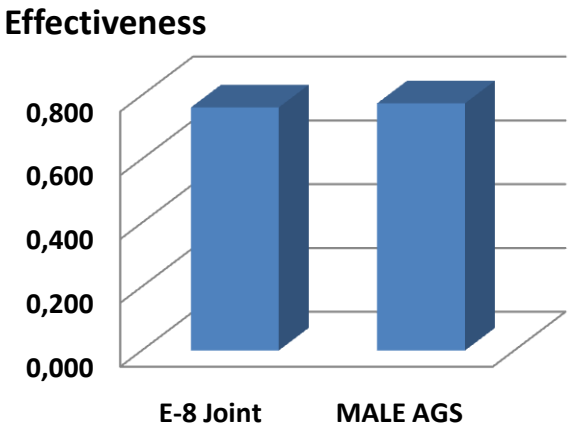


Fig. 3 Effectiveness results for AGS mission

statistical points. This technique employs a consistent database composed of all physical dimensions, characteristics and performances used as cost drivers. Defined formulas are implemented in order to evaluate cost variation with cost drivers. The maximum hydraulic flow, the electrical energy provided by generators and the fuel tanks volume are examples of cost drivers for, respectively, hydraulic, electric and fuel systems. Effects on costs of changes in aircraft design can be evaluated by using this technique, which finds application in the preliminary design phase. As a final result the technique estimates, by linear regression, the Cost-Estimating Relationships (CERs). Eq. (5) shows general form of CERs.

$$C_i = A_i \cdot Q_i^{B_i} \cdot N_p^{-K_i} \qquad (5)$$

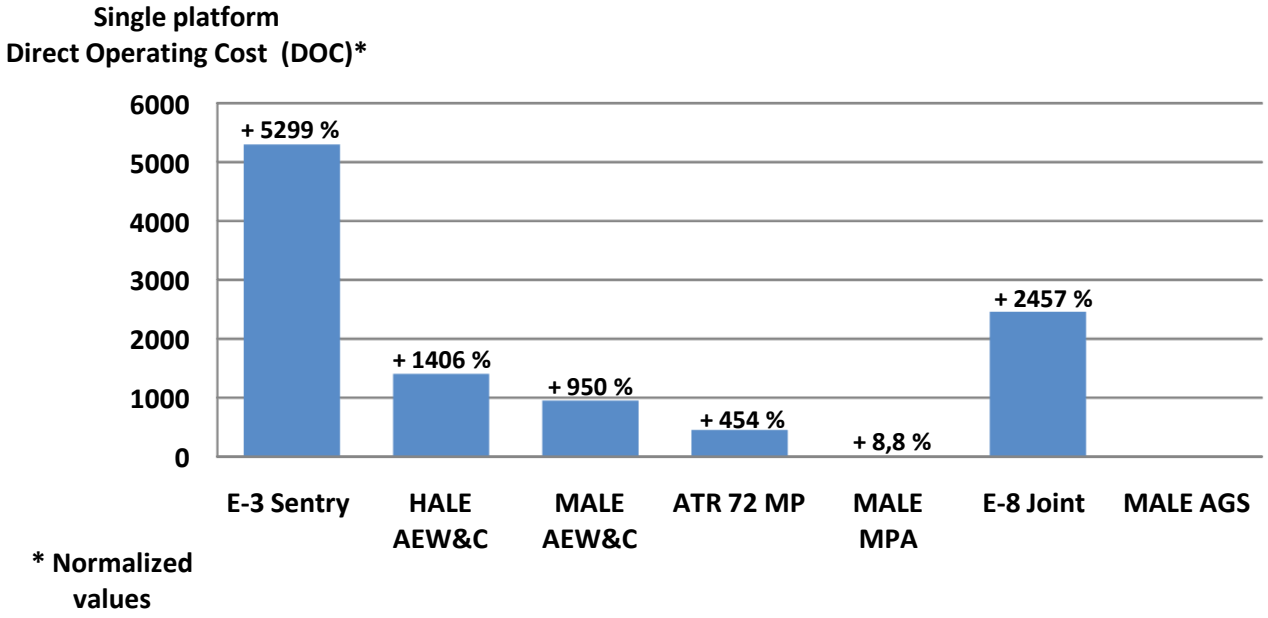


Fig. 4 Direct Operating Cost (DOC) of single platforms.

3.2 Cost Analysis

The cost estimation tool has been developed by Politecnico di Torino during a Ph.D. thesis [2] with the support of the preliminary design office of Alenia Aeronautica and starting from methodology proposed by *Dr Roskam J.* [5]. The tool is based on parametric and statistical approach and it uses a large number of

C_i represents the average unit acquisition cost or maintenance cost of the subsystem "i" (USD), Q_i is the main characteristic parameter (e.g. weight) of the subsystem "i", N_p is the production quantity, K_i is a coefficient that takes into account the learning curve effect. B_i represents a "scale factor" whose values vary from zero to one and it considers the currency saving due to the increase of dimensions,

indeed, costs increase with weight augmentation. This factor is only applicable for airframe structure acquisition CERs. A_i is the subsystem acquisition cost per unit weight (USD/Kg). By considering the maintenance cost CERs, A_i represents the specific maintenance cost related to a design parameter of the subsystem "i". In particular, A_i is relevant for the airframe acquisition CER, feeble for CERs regarding general systems and it assumes intermediate values for avionic systems. The N^{Ki} coefficient is not considered in the maintenance cost CERs.

3.2.1 Cost Results

By using described methodology it is possible to calculate DOC of all the platforms considered in the Effectiveness analysis. Figure 4 shows results of this calculation normalized by using the lowest value of DOC that is for MALE AGS asset. It is possible to notice that single UAS platforms have always a lower DOC than manned platforms. In addition the analysis confirms, as expectable, that AEW&C mission present a really high Direct Operating Cost compared to other missions. In order to make comparison among platforms, DOC of the whole surveillance system has to be considered. In particular, by considering Eq. (1) for each mission it is to calculate the number of assets necessary to perform surveillance on the area defined in Table 1 of requirements. It is to notice that AGS requirement of area to survey is too small to make a proper comparison between E-8 Joint and MALE AGS. The area that one E-8 Joint can survey has been supposed for AGS assets calculation. This supposition does not compromise validity of results because both Effectiveness and Costs are normalized. Table 8 presents results. Cost of surveillance system can be approximated by multiplying DOC of each platform and the number of assets. This approximation is valid for our purposes because we are not interested in the value of DOC itself but in the percentage difference among DOC's.

Table 8 Asset calculation for cost evaluation of surveillance systems

Asset	Number of aircraft
E-3 Sentry	3
MALE AEW&C	3
HALE AEW&C	2
ATR 72 MP	1
MALE MPA	4
E-8 Joint Stars	1
MALE AGS	6

3.2.1.1 AEW&C

Figure 5 shows results of Cost analysis for AEW&C surveillance systems. By analyzing DOC's of AEW&C surveillance systems it is possible to see that the costs of two HALE AEW&C and three MALE AEW&C are really similar. A system composed of two E-3 Sentry has a huge cost compared with UAS surveillance systems. Even if this result calls for UAS employment for AEW&C purposes, nowadays manned platforms are still used because existing UAS are too small to be converted and reduced number of AEW&C assets salable does not allow amortizing huge development costs of an unmanned platform for AEW&C purposes.

3.2.1.2 MPA

Figure 6 shows results of Cost analysis for MPA surveillance systems. By analyzing results it is possible to see that a system composed of four MALE MPA have a direct operating cost 30% lower than one ATR 72 MP. This results is justified by the fact that typically MPA's are regional turboprop platforms with not so high operating costs. This also justifies the employment both of manned platform and UAS platforms for maritime surveillance.

3.2.1.3 AGS

Figure 7 shows results of Cost analysis for AGS surveillance systems. Even if the UAS surveillance system is composed of six platforms, its cost is 326% lower than a system composed by one E-8 Joint. This result is confirmed by the modern employment of above all UAS assets for AGS purposes.

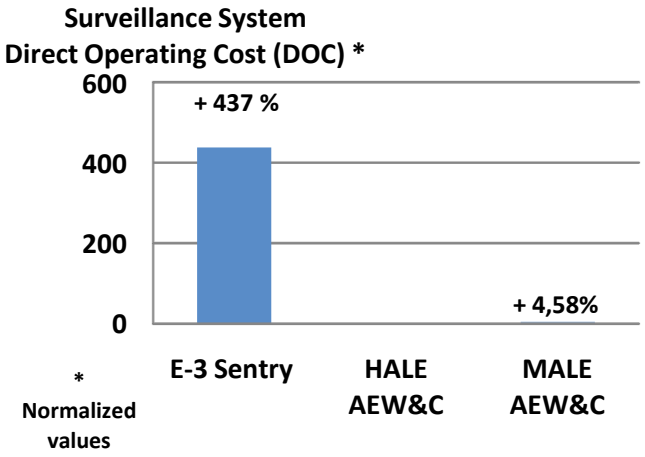


Fig. 5 Cost results for AEW&C surveillance systems

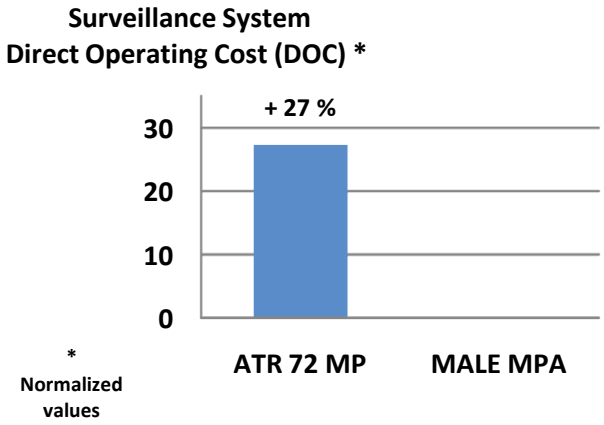


Fig. 6 Cost results for MPA surveillance systems

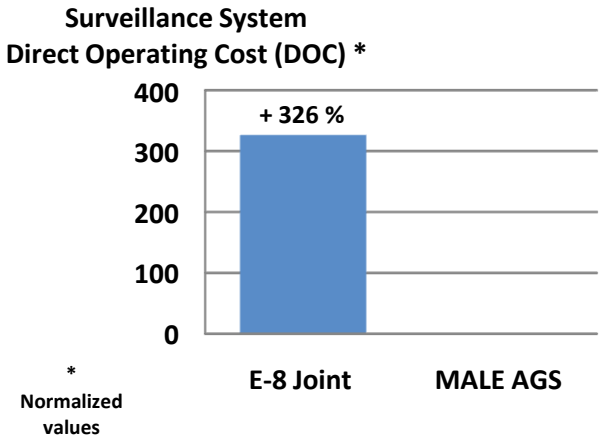


Fig. 7 Cost Results for AGS surveillance system

3.3 Effectiveness/Cost results

By dividing Effectiveness results for Cost results and by normalizing, it is possible to obtain the following Effectiveness/Cost results.

3.3.1.1 AEW&C

Figure 8 shows Effectiveness/Cost results for AEW&C surveillance systems. UAS system has the highest Effectiveness/Cost and in particular HALE AEW&C reaches the best result. E-3 Sentry has a high DOC so its result is scarce.

3.3.1.2 MPA

Figure 9 shows Effectiveness/Cost results for MPA surveillance systems. The difference between a system composed of ATR 72 MP or MALE MPA's is not so high and, as already said, it justifies the modern employment of both kind of platforms in the maritime surveillance environment.

3.3.1.3 AGS

Figure 10 shows Effectiveness/Cost results for AGS surveillance systems. A system composed of MALE AGS has a high Effectiveness/Cost result compared with a system composed by a E-8 Joint Stars. This is due to similar Effectiveness between E-8 Joint and MALE AGS and big difference in terms of DOC. This result justifies the employment of E-8 Joint Stars just when high performances of its sensors are requested.

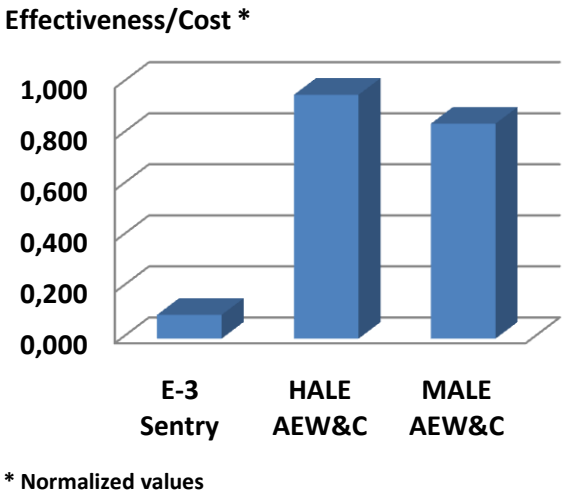


Fig. 8 Effectiveness/Cost results for AEW&C surveillance system

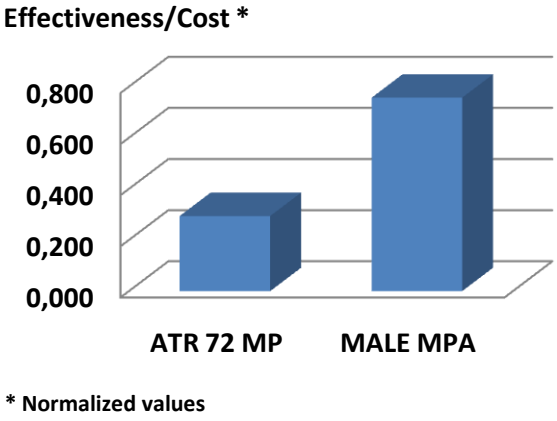


Fig. 9 Effectiveness/Cost results for MPA surveillance system

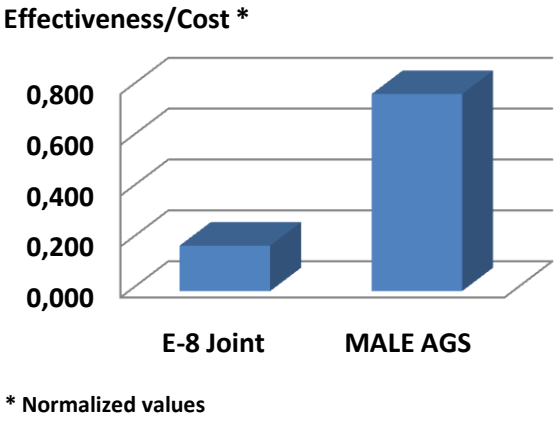


Fig. 10 Effectiveness/Cost results for AGS surveillance system

4 Conclusions

By assigning C4ISTAR surveillance requirements UAS preliminary designs have been defined for every mission profile. By comparing UAS solutions and existing manned aircraft in terms of Effectiveness/Cost it has been possible to find the following results:

- Effectiveness of UAS platforms is often higher than manned ones due to the typical high endurance and high operating altitude of patrolling missions.
- Costs of UAS platforms are smaller than manned ones even considering the whole surveillance system. In particular as far as AEW&C assets are concerned it is possible to notice that there are small differences in term of costs between a solution composed of MALE assets and one composed of HALE assets.
- Effectiveness/Cost results reveal that big difference between manned and UAS surveillance systems exists for AEW&C and AGS missions. MPA mission concern small differences that justify modern employment of both kind of platform for maritime patrol.

Acknowledgments

The authors want to thank preliminary sizing office of Alenia Aeronautica S.p.A for the contribution on Matlab® tools development and Eng. Valerio Bonifazio for its contribution in the contextualization of arguments.

References

[1] Di Meo G. A., "System of System(SoS) project and optimization in a Network Centric Operations (NCO) environment", *VII Pegasus AIAA Student Conference 2011 Torino*, Department of Aerospace Engineering, Politecnico di Torino, Italy, 2010.

[2] Fioriti M., Innovative solution for light, very light and unmanned aircraft, Ph.D. Dissertation, Department of Aerospace Engineering, Politecnico di Torino, Italy, 2010.

[3] Dr. Roskam J., *Airplane Design Part I: Preliminary sizing of airplanes*, 1st ed., Roskam Aviation Engineering Corporation, Ottawa, Kansas, USA, 1985, Chaps 1-4

[4] Dr. Roskam, J., *Airplane Design Part II: Preliminary configuration design and integration of the propulsion system*, 1st ed., Roskam Aviation Engineering Corporation, Ottawa, Kansas, USA, 1985, Chaps 6, 8, 12, 13

[5] Dr. Roskam, J., *Airplane Design Part VIII: Airplane Cost Estimation: Design Development, Manufacturing and Operating*, 1st ed., Roskam Aviation Engineering Corporation, Ottawa, Kansas, USA, 1985

[6] Wilson, M., *Jane's Avionics*, Jane's Information Group, Coulsdon, Surrey CR5 2YH, UK, 2003

[7] Munson, K., *Jane's Unmanned aerial vehicles and targets*, Jane's Information Group, Coulsdon, Surrey CR5 2YH, UK, 2002

[8] Locascio, A., Thurston, D.L., "Transforming the House of Quality to a multiobjective optimization formulation", *Structural*

optimization, Vol. 16, Springer-Verlag,1998, pp 136-146

[9] *US Army Unmanned Aicraft Systems Roadmap 2010-203*, U.S. UAS Army Center of Excellence, Bldg 5000, Luck Star Street, Fort Rucker, Alabama, USA. 2009.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Electrodynamic Tethers for Deorbiting Maneuvers

D. Zanutto[1] and G. Colombatti[2]
[1]Center of Studies and Activities for Space CISAS "G. Colombo" Research Center, Padova, Italy, denis.zanutto@studenti.unipd.it, phone: +390498276795, fax: +39 0498276788
[2]Center of Studies and Activities for Space CISAS "G. Colombo" Research Center, Padova, Italy, giacomo.colombatti@unipd.it, phone: +390498276848

E. C. Lorenzini[3]
[3]Department of Mechanical Engineering of University of Padova, Italy, enrico.lorenzini@unipd.it, phone: +390498276766

Keywords: *electrodynamic tether, deorbiting, parametric analysis, control, eigenfrequencies*

Abstract

The space debris proliferation has focused the attention of the space community on passive tethers as deorbiting system, since Earth's environment is particular favorable to fly such a system. An orbital simulator has been implemented to investigate the dynamics of the tether and its deorbiting performance. The code studies the tethered system as acted upon by gravity and forced by electrodynamic and aerodynamic forces. Moreover it includes all the upgraded environmental routines for the Earth's atmosphere, magnetic field, electron density and gravity potential, so as to simulate as best as possible the actual situation in flight. Also the thermal dynamics of the tether has been taken into account, because the electric current intensity depends on the electrical resistance that varies with temperature. Particular attention has also been paid to the bending dynamics, because the Lorentz force is a distributed load all along the tether that excites the lateral vibration modes.

1 Introduction

Since the beginning of space exploration thousands of satellites have been launched in orbit around the Earth. Nowadays, space debris (mainly consisting of non-operative dead spacecraft, spent upper stage, residues from explosions and collisions) represents a serious threat for new spacecraft because of hypervelocity impacts that can cause severe damages or even the failure of a mission. Consequently, space agencies are planning strategies to dispose or to make reenter satellites at the end of their operative life.

In the last years several authors [1-4] have studied in detail electrodynamic tethers (EDT) for deorbiting maneuvers. Such a system can carry out a propellantless and relatively quick reentry of satellites in low Earth orbits (LEO) because the Earth's environment is particularly favorable to fly an EDT. In fact the high electron density surrounding the planet at LEO orbits and the magnetic field generate a non-negligible induced potential, according to the Faraday's law, which enables the collection of ionospheric electrons at the anodic end of the tether. Consequently, when the circuit is closed an electric current [5] flows along the wire, and the interaction between the charged particles and the magnetic field generates a Lorentz force acting on the tether that progressively lowers the orbit of a satellite attached to the EDT.

A tethered system is subjected to forces and torques associated with the gravity gradient and

various perturbations. The Earth's gravitational attraction tends to stabilize the satellite keeping the wire taught and aligned with the local vertical, while perturbations (mainly due to electrodynamic forces, aerodynamic drag, and solar pressure) change its orbit. Specifically, the electrodynamic force pumps continuously energy into the system attitude motion [6], enlarges the libration dynamics and bends the tether exciting the natural modes of the lateral motion [7]. So if no-controlled an EDT system goes rapidly into instability.

In this paper we presents a parametric analysis of several parameters affecting the attitude of the tether, in order to define the main causes of instability. Then we introduce an on-off control of the electric current to provide a complete reentry of the satellite till low altitudes where aerodynamic drag is plenty to guaranty a rapid decay. At last we investigate by means lump masses approach the lateral motion of the wire highlighting the most important frequency term exciting its dynamics.

2 Tethered Satellite Dynamics

In literature we can find two main models to study the dynamics of a tethered system: the first describes the satellite like a rigid dumbbell with an inextensible wire; the second is more accurate and considers it as flexible and extensible. The dumbbell model is very useful to obtain quickly information about the system, investigate its performance, identify operative orbital scenarios and run extensive parametric analyses. On the other end, the flexible wire model gives a detailed description of attitude dynamics, studying also the lateral bending caused by the electrodynamic and aerodynamic force distributed along the tether.

2.1 Dumbbell Model

In the dumbbell model the tether is a thin rigid rod with mass m_t and length L, connecting two tips masses m_1 and m_2. The system is considered undeformable and ignores all the elastic phenomena. The simplicity of this model allows to study a few simple cases by means of analytical approach and, more generally, to derive simple equations of motion for immediate numerical implementation. This method lets to investigate the main instability associated with the libration dynamics of the system, eliminating those associated with tether flexibility.

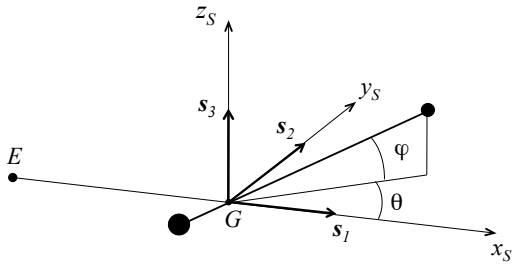


Fig. 1 Synodic reference frame

In the end, neglecting the contribution due to the torsion, the tether moves around the local vertical and this motion can be described by two libration angles: in-plane θ and out-of-plane φ (see Fig. 1).

2.2 Orbital Motion

The orbital path followed by the EDT depends on all the forces acting on the system. In a detailed analysis the contributions are several, because the dynamics is affected by various interaction with the surrounding environment. We define three main contributions to the orbital dynamics: the gravitational, electrodynamic and aerodynamic forces, because those are the most important interaction affecting the satellite in our orbital scenario.

Gravitational Force

A tethered system can be several kilometers long, so the attraction of Earth's gravitational field must be studied as that of an extended body. To investigate the gravitational effects some assumptions must be made: in first approximation, we consider the potential model of the gravity as spherical, and the higher harmonics of the gravitational potential neglegible. The gravitational force will be provided by two main terms: the former explains the homogenous gravity attraction, the latter the mass distribution of the body.

$$\vec{\mathbf{F}}_{gr} = -\frac{\mu}{r^2} m[\hat{\mathbf{r}} + \sum_{n=2}^{\infty} (-1)^n \left(\frac{L}{r}\right)^n a_n (S_n[\cos\alpha]\hat{\mathbf{r}} + -S_{n-1}[\cos\alpha]\hat{\mathbf{u}})] \qquad (1)$$

Long tether affects the orbital motion of the whole system with second order terms causing small perturbations, whose magnitude is proportional to the $(L/R)^n$ ratio.

Electrodynamic Force

Along the orbit the tether crosses continuously the magnetic field lines generating according to the Faraday's law a differential of potential *EMF* between the two ends:

$$EMF = \left(\vec{\mathbf{v}}_{rel} \times \vec{\mathbf{B}}\right) \cdot L\hat{\mathbf{u}} \qquad (2)$$

where $\vec{\mathbf{v}}_{rel}$ is the satellite relative velocity with respect to the magnetic field, and $\vec{\mathbf{B}}$ the magnetic field of the Earth evaluated at the center of mass. If the electric circuit is closed through the ionosphere by ejecting the electrons at the cathode, a conventional current $I(s)$ of positive charges can flow along the wire in the direction from the Earth outwards, supposing the tether moves eastward on a prograde orbit.

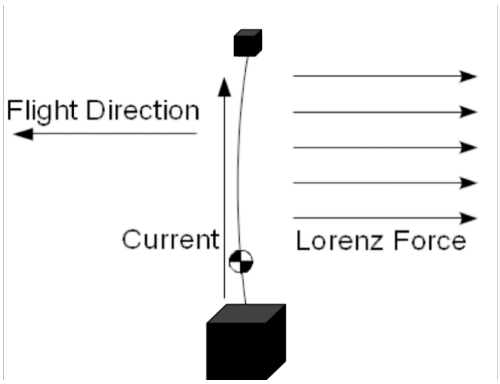


Fig. 2 Electrodynamic force

This current interacts with the geomagnetical field and a Lorentz drag force is induced in the opposite direction of the orbital motion (see Fig. 2):

$$\vec{\mathbf{F}}_{el} = \int_L I(s)\hat{\mathbf{u}} \times \vec{\mathbf{B}} ds \qquad (3)$$

Aerodynamic Force

The atmospheric density decreases exponentially versus the height above the surface, therefore its effects become predominant at low altitudes, causing a rapid decay of the satellite. The magnitude of aerodynamic forces depends strongly on the geometry of the vehicle and in particular on the frontal surface:

$$\vec{\mathbf{F}}_a = -\frac{1}{2}\int_A \rho c_D \left(\vec{\mathbf{v}}_{rel} \cdot \hat{\mathbf{n}}\right)\vec{\mathbf{v}}_{rel} dA \qquad (4)$$

2.3 Attitude Motion

The attitude dynamics of the tether is governed by the angular momentum $\vec{\mathbf{H}} = \mathbf{I}\vec{\Omega}$ with $\mathbf{I}$ the inertial tensor and $\vec{\Omega}$ the angular velocity of the satellite. Its variation is due to the gravitational, electrodynamic and aerodynamic torques acting on the satellite:

$$\frac{d}{dt}\vec{\mathbf{H}} = \vec{\mathbf{M}}_{gr} + \vec{\mathbf{M}}_{el} + \vec{\mathbf{M}}_a \qquad (5)$$

Assuming that the tether cannot lengthen or shorten, the attitude dynamics can be expressed in the body reference frame as a function of libration angles in two non-linear and coupled differential equations

$$\ddot{\theta} = -\dot{\omega}_{orb} + 2(\dot{\theta} + \omega_{orb})\dot{\phi}\tan\phi + \frac{M_\theta}{I_s \cos\phi} \qquad (6)$$

$$\ddot{\phi} = -\frac{1}{2}\sin 2\phi\left(\dot{\theta} + \omega_{orb}\right)^2 + \frac{M_\phi}{I_s} \qquad (7)$$

Where M_θ and M_ϕ are the sum of external torques affecting θ and ϕ, respectively, and ω_{orb} the orbital angular velocity.

Gravitational Torque

The gravity gradient torque depends on how the mass is distributed with respect to the CM of the satellite. The gravitational forces generate a

momentum that tends to stabilize the satellite aligning it with its minimum inertia axes.

$$\vec{\mathbf{M}}_{gr} \approx \frac{3\mu}{r^3}\hat{\mathbf{s}}_1 \times (\mathbf{I} \circ \hat{\mathbf{s}}_1) = \frac{3\mu}{r^3} I_s (\hat{\mathbf{u}} \times \hat{\mathbf{s}}_1)(\hat{\mathbf{u}} \cdot \hat{\mathbf{s}}_1) \quad (8)$$

Where the unit vector $\hat{\mathbf{u}}$ indicates the orientation of the wire.

Electrodynamic Torque

The Lorentz torque $\vec{\mathbf{M}}_{el}$ caused by the current $I(s)$ flowing along the wire is:

$$\vec{\mathbf{M}}_{el} = J_1 \hat{\mathbf{u}} \times \left(\hat{\mathbf{u}} \times \vec{\mathbf{B}}\right) \quad (9)$$

Where J_1 a coefficients explaining the electric current distribution along the wire:

$$J_1 = \int_{-\sin^2\phi}^{\cos^2\phi} I(s)(h_G - Ls)ds \quad (10)$$

Aerodynamic Torque

The aerodynamic torque is predominant at lower altitude where the electron population is less dense, and can be evaluated as:

$$\vec{\mathbf{M}}_a = -\frac{1}{2}\int_A (\vec{\mathbf{r}} - \vec{\mathbf{r}}_G) \times \rho c_D (\vec{\mathbf{v}}_{rel} \cdot \hat{\mathbf{n}}) \vec{\mathbf{v}}_{rel} dA \quad (11)$$

It brings quickly the system to instability, because it can't be controlled in the same way as electrodynamic one.

2.4 Thermal Model

The temperature is important for the dynamical motion because it affects the electrical and mechanical properties of the tether and, consequently, the electrodynamic interaction. Its variation with time depends on several terms [8-9], mainly the solar radiation and radiative cooling:

$$\frac{dT}{dt} = \frac{1}{m_k c_k}(Q_{sol} + Q_{al} + Q_e + Q_a + Q_{ohm} - Q_{rad}) \quad (12)$$

Where Q_{sol}, Q_{al}, Q_e, Q_a, Q_{ohm} and Q_e are the thermal fluxes due the Sun, Earth's albedo, the Earth's IR radiation, the heating related to atmospheric drag, the ohmic losses the radiative cooling, and c_k the thermal heat capacity

2.5 Lump Masses Model

To solve the motion of the whole system, avoiding mathematical complications, an effective strategy is to discretize the wire as a series of lumped masses connected by massless elastic springs and dampers [9-10]. The model spatial resolution becomes higher with more nodes/lumps discretizing the wire, but this increases rapidly the computational time required for a solution. So the number of lumped masses will be a tradeoff between a good prediction of the real motion and the minimization of the integration time without loss of fundamental information about the critical aspects of the system.

In a tethered system, two masses orbiting at different heights with a common orbital rate ω_{orb} must be subjected to a tether tension to compensate for the excess or reduction in their velocity. In fact the upper mass runs at higher velocity in order to follow the same orbit, while the lower mass must be orbiting at a slower velocity than the orbital velocity at that altitude.

The motion of the tether relative to the inertial reference system can be described by a set of partial differential equations:

$$\frac{\partial}{\partial s}\vec{\mathbf{T}} + \vec{\mathbf{f}}_{el} + \vec{\mathbf{f}}_a = \rho_T \frac{\partial^2}{\partial t^2}\vec{\mathbf{r}} \quad (13)$$

where $\vec{\mathbf{T}}$ is the tether tension, $\vec{\mathbf{f}}_{el}$ and $\vec{\mathbf{f}}_a$ the Lorentz and aerodynamic forces per unit of length at the element dm, and s the position along the wire.

The tension is a fundamental aspect to take into account when we work on flexible tether because it drives the lateral dynamics response

of the wire. In fact the eigenfrequencies of lateral oscillations are function of the same tension along the rope, which is function of the local deformation of the tether. In a visco-elastic model the tether internal force can be separate in two main contributions: the elastic and damping term:

$$\vec{\mathbf{F}}_t = \vec{\mathbf{F}}_{elas} + \vec{\mathbf{F}}_{damp} \qquad (14)$$

Where

$$\vec{\mathbf{F}}_{elas}(\vec{\mathbf{r}}_k, \vec{\mathbf{r}}_{k+1}) = \frac{E_k A_k}{l_k^0}\left\{|\vec{\mathbf{r}}_{k+1} - \vec{\mathbf{r}}_k| - l_k^0\left[1 + \alpha_k\left(T_k - T_k^0\right)\right]\right\}\hat{\mathbf{u}}_k \qquad (15)$$

$$\vec{\mathbf{F}}_{damp}\left(\dot{\vec{\mathbf{r}}}_k, \dot{\vec{\mathbf{r}}}_{k+1}\right) = b_k\left\{|\dot{\vec{\mathbf{r}}}_{k+1} - \dot{\vec{\mathbf{r}}}_k|\right\}\hat{\mathbf{u}}_k \qquad (16)$$

With α_k, l_k^0, b_k and $\hat{\mathbf{u}}_k$ the thermal expansion coefficient, the unstretched length of tether element k, the material damping constant and the unit vector between the masses m_k and m_{k+1}:

$$\hat{\mathbf{u}}_k = \frac{\vec{\mathbf{r}}_{k+1} - \vec{\mathbf{r}}_k}{|\vec{\mathbf{r}}_{k+1} - \vec{\mathbf{r}}_k|} \qquad (17)$$

2.6 Electric Circuit Model

The tether model that we consider is a bare EDT of length L, width w and thickness h, made of conductive material of density ρ and electrical conductivity σ, which are function of temperature. In this study we will be working with relatively short wires, hence it is reasonable to assume that the quantities N_e and B are constant along the tether, because their scale heights are too long to force meaningful variations along the tether length.

A totally or partially bare tether is an efficient anode. Moreover, the electron collecting area can be increased by using longer tether and satisfy at the same time the condition that the Debye length be larger than the tether radius. The motion of the wire with respect to the magnetic field generates the potential difference with respect to the plasma. As explained by the Faraday's law, the relative motion causes an electric field (called motional electric field), along the wire, provided by the induced potential:

$$EMF = \int_L \vec{v}_{rel} \times \vec{B} \cdot d\vec{l} \qquad (18)$$

where $\vec{v}_{rel}$ is the relative velocity of the tether respect the magnetic field lines. Plasma has been assumed fixed to the magnetic field and corotating with the planet.

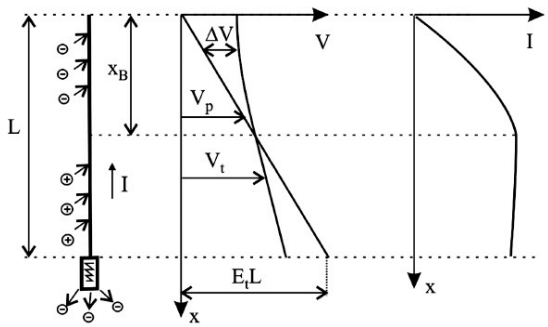


Fig. 3 Electric Current Profile.

When the electrical circuit is closed electrons are continually collected in the upper part of the wire along the positively biased segment. The electric current profile is non-linear and reaches the maximum at x_B where the potential bias ΔV goes to zero. Then in the general case in which the potential drop of the plasma contactor V_C and/or the impedance Z_{load} of the applied load is not zero the lower portion of the tether will be negatively biased with respect to the plasma and will collect positive ions, producing a very small decrease of the electric current, as illustrated in Fig. 3.

The current collection is here assumed to take place in the orbital motion limited (OML) regime, and the equations governing the process along the wire are [5][11]:

$$\frac{d\Delta V}{dx} = \frac{I}{\sigma w h} - E_t \qquad (19)$$

$$\frac{dI_a}{dx} = \frac{p}{\pi} q_e N_e \sqrt{\frac{2q_e}{m_e}\Delta V} \qquad (20)$$

$$\frac{dI_c}{dx} = -\frac{p}{\pi} q_e N_e \sqrt{-\frac{2q_e}{m_i}\Delta V} \qquad (21)$$

Where I is the electric current (I_a and I_c respectively the current flowing in the anodic and cathodic segments), p the perimeter of the wire [$p = 2\ (w + h)$], q_e the electron charge, m_e

the electron mass, m_i the mass of the most abundant ion species.

This continuous replenishment of electrons is driven by the motional electric field projected along the wire and the local electron density of the Earth's ionosphere. In order to obtain the electric profile along the tether we must solve a set of two differential equations with boundary conditions that have also to satisfy the circuit equation:

$$E_t L = \Delta V_A + \Delta V_{TETHER} + \Delta V_{LOAD} + \Delta V_C \qquad (22)$$

with the following boundary conditions:

$$\begin{cases} I = 0, \Delta V = \Delta V_A & \text{at} \quad x = 0 \\ \Delta V = \Delta V_C + I_{(x=L)} Z_{load} & \text{at} \quad x = L \end{cases} \qquad (23)$$

where $E_t L$ is the difference of potential induced by Faraday's law, while ΔV_A, ΔV_{TETHER}, ΔV_{LOAD} and ΔV_C are the potential drops respectively at anode, along the tether, the load and cathode. The drop along the tether depends on its electric resistance, and so on the tether cross-section, length, material electrical properties and temperature.

3 Parametric Analysis

An EDT is subjected to a continuously changing electrodynamic torque, whose equilibrium position cannot be described by a static solution. This torque pumps energy into the system enlarging the tether oscillation till transform it into a rotation. The aim of this section is to identify the critical parameters that govern the EDT dynamics along the orbit studying the decay performance for each configuration, and how long the system can work before going to instability.

The key parameters considered in this work for a parametric analysis of deorbiting maneuvers can be grouped in three main categories: tether, satellite and environment parameters. In the first group we consider the length, the thickness and the width of the tether; the second group comprises the satellite mass distribution and the orbital inclination and altitude; the third set consists of the Solar activity index. The tether has been tested among all of the above.

It is possible to observe, in Fig. 4, how important the tether length is for the evolution of the orbit: the longer the tether, the faster the orbital decay; in fact, the longer the tether the higher the current flowing in the tether and the higher is the destabilising electrodynamic force. In fact the current results in being proportional to the tether length $I \propto p \cdot L^{3/2}$ [5][11]. Tether width has a very similar effects, because it augments the collection area and so the electric current. As well thickness increases the current profile because enlarge the sectional area enhanced tether conductivity.

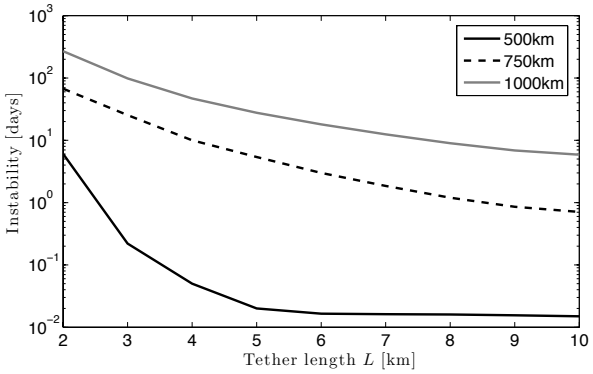


Fig. 4 Instability vs tether length

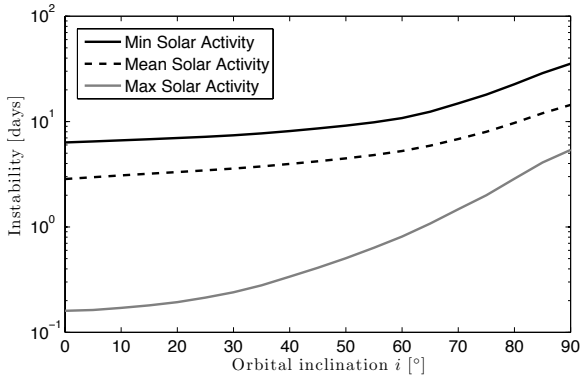


Fig. 5 Instability vs orbital inclination

The mass ratio m_{sat}/m_b, between the satellite and ballast mass, is very significant because it changes the position of barycenter of the whole system and so the value of term J_1 (Eq. 10). High mass ratios mean extremely strong electrodynamic interaction and so little stability.

The altitude evolution depends strictly on the starting altitude: if the altitude is too high (≥ 1400 km) the system lifetime is very long (tens of years). This is, of course, due to the fact that

the produced current is greater at low altitudes, as shown in Fig. 4, where the electromagnetic field is higher.

On the other hand, one of the most significant parameters is the initial orbital inclination because it is strongly related to the Earth electromagnetic field and consequently on tether current and direction of Lorentz force (see Fig. 5). The inclination values selected for the simulations are coherent with the orbital debris distribution as measured in the past years [12], [13]. Starting orbit inclinations are: 0, 30, 60 and 85 deg. As expected the higher the inclination the lower is the current because of the geometry of the Earth magnetic field.

The solar activity is the most significant environmental variables, since the 11years solar cycle changes the electron density in the ionosphere. Figures 5 shows the effect of solar activity at three different data:

- 01 January 1996: Min Solar activity
- 01 January 1998: Mean Solar activity
- 01 January 2000: Max Solar activity

When the solar activity is maximum the tether can collect a major number of electrons [14], therefore higher the solar flux, higher the current is and consequently the destabilising electrodynamic force.

4 Control

The main objective of an EDT system is to fulfill the decay maneuvers till low altitudes where the aerodynamic drag is plenty to ensure a rapid reentry. As just shown, without an appropriate control strategy the libration dynamics goes rapidly in instabilities causing a drastic decrease of the deorbiting performance. To avoid such a critic situation in this work we have considered a control strategy based on energetic approach [12-13]. For monitoring the stability of the satellite its Halmitonian function is evaluated by using information about the libration energy of the tether.

$$V_{stab} = 4 + \frac{\dot{\phi}^2 + \cos^2\phi\left(\dot{\vartheta}^2 - \omega_{orb}^2 - 3\omega_{orb}^2\cos^2\vartheta\right)}{\omega_{orb}^2} \qquad (24)$$

When V_{stab} overcomes a threshold value V_{th} the code must look at the direction of the oscillation respect to Lorentz force. In fact when both are in the same direction the electrodynamic torque causes a positive work that amplifies the oscillation, while in the other case the work becomes negative and decreases the rotational kinetic energy. Hence the control routine has to observe the attitude dynamics switching on or off the electric current, by closing and opening the circuit, as a function of the threshold value and oscillation direction.

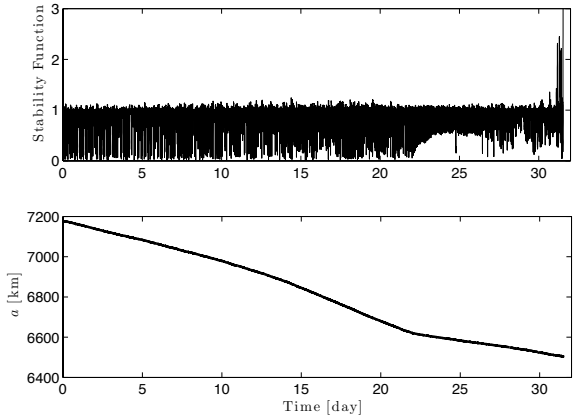


Fig. 6 Stability function V_{stab} and orbital major semiaxes

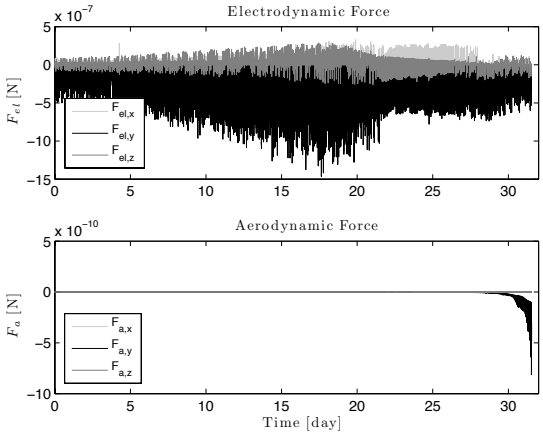


Fig. 7 Electrodynamic and aerodynamic forces

In the next figures an example of controlled reentry is shown. The tethered system has been chosen 5km long, 1cm width and 0.1mm thick, the satellite and ballast masses are 1000 and 20kg, respectively. The satellite starts from a circular equatorial orbit at an altitude about 800km and the threshold value of control

function is set at 1 (see Fig. 6). The deorbiting is simulated during a maximum of solar activity.

The electrodynamic force, whose average value is shown in Fig. 7, carries out the deorbiting till low altitudes where the atmosphere becomes quite dense to complete the reentry. In fact at this height aerodynamic load becomes very strong and the torque due to it brings the satellite into instability, since it cannot be controlled as well as electrodynamic one. But this has little importance, because our aim has been reached and the satellite from its initial operative orbit has been brought down at heights where aerodynamic drag can quickly complete the maneuver. Figure 8 describes the libration dynamics of the wire. The oscillation is mainly around the local vertical because the control forces the electric circuit to stay open for long time.

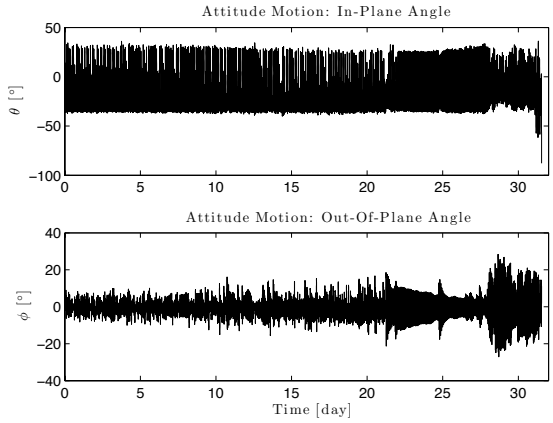


Fig. 8 In-plane and out-of-plane libration angles

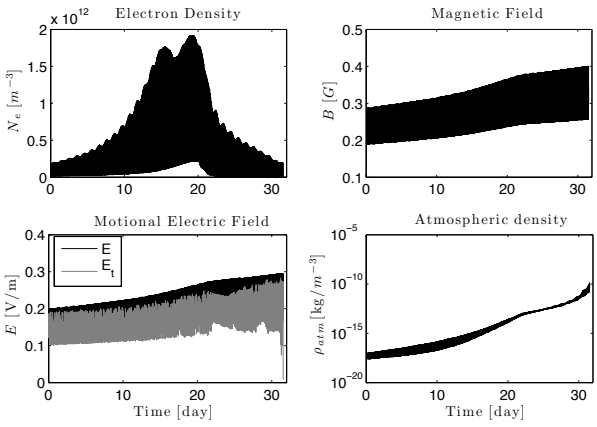


Fig. 9 Environment Variables

The environment variables have been reported in Fig. 9. The electron density has a maximum when the satellite is at roughly 300km, while the magnetic field, the motional electric field and aerodynamic density increase as the major semiaxes diminishes.

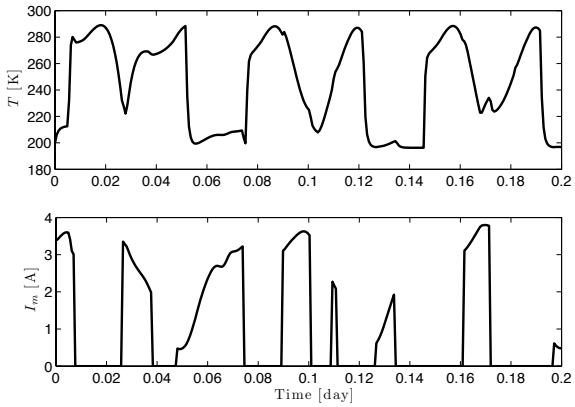


Fig. 10 Zoom of tether temperature and current profile

At least in Fig. 10 there is a zoom of tether temperature and average electric current. The temperature changes a lot because of solar radiation. When illuminated by sun the tether reaches the maximum temperature, otherwise it goes to equilibrium with Earth's radiation, that is the second most important thermal flux. The electric current is affected by thermal behavior, because tether electric resistance is function of temperature. As it can be seen in Fig. 10 the control algorithm forced the circuit to be open for a long time.

5 Lateral Dynamics

Lateral dynamics is an important issue for tethered satellite, because the transversal oscillation can introduce new instability that a rigid dumbbell model cannot see [14-16]. In order to investigate the main features of in-plane and out-of-plane motion when perturbed by Lorentz force a lump mass approach has been used, discretizing the whole tether in several nodes, that substitute the inertial, electric and thermal properties of the same wire. By adding nodes into the model we can improve the description of the whole dynamics till to higher frequencies, which however are much less excited than first three modes and librational one. For this reason a model with at least 5 nodes is plenty for an EDT satellite, and a smart

tradeoff between accuracy and computational time. In fact the cut-off frequency is much higher than first eigen-modes, so their value is rather representative of those of a real system.

In Figs. 11 – 12 the trend of in-plane and out-of-plane is shown for the following initial conditions: circular equatorial orbit at 1000km of altitude, masses m_1 and m_2 about 1000 kg and 25kg, respectively, and tether 5km long.

In the spectral response of the in-plane motion, reported in Fig. 13, we can notice several terms among which the in-plane and out-of-plane librations at frequency $\sqrt{3}f_{orb}$ and $2f_{orb}$, respectively (coupling between the motion is just provided by dumbbell model equations). The most important bending mode of the tether is the first (f_1≈1.16Hz for a mean tension about 0.44N), while the others at higher frequency are much less excited by electrodynamic force. The Earth's environment introduces two important components due the rotation of the planet around its spin axis (f_{Earth}=1.2x10^{-5}Hz) and the relative motion f_{rel} between the spacecraft and surrounding plasma, which affect the motional electric field and consequently the electric current:

$$f_{rel} = f_{orb}\left(1 + \frac{|\vec{v}_{rel}|}{|\vec{v}_{S/C}|}\right) \qquad (25)$$

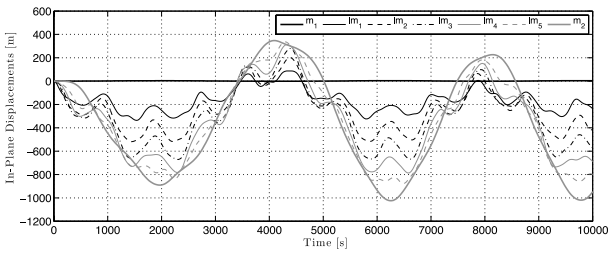


Fig. 11 In-Plane motion - 5 lump masses

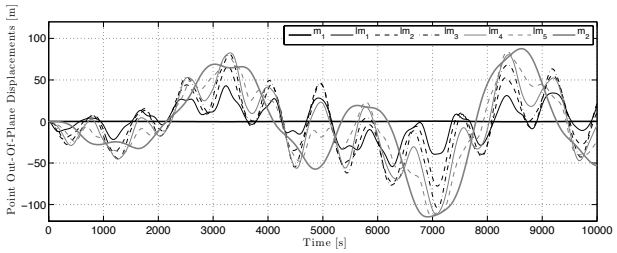


Fig. 12 Out-Of-Plane motion - 5 lump masses

The Earth's rotation frequency is very important, because the electron density, magnetic field, motional electric field and atmosphere density have a diurnal variation between light and dark side of the Earth, and this affects the collection of electron toward the anodic end.

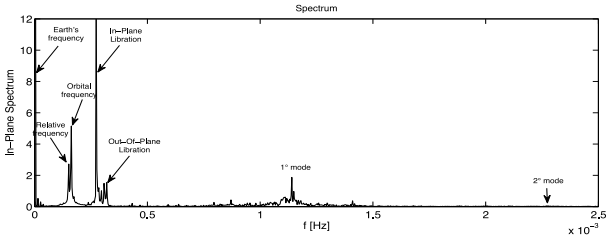


Fig. 13 Spectrum of in-plane lateral dynamics

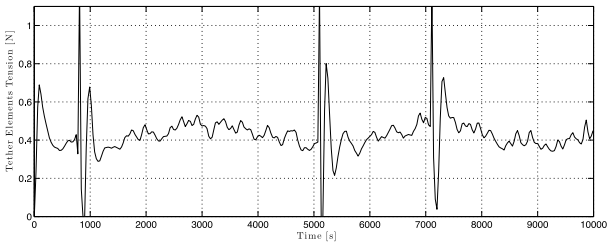


Fig. 14 Tension at the middle of the wire

As shown in **Fig. 14**, the tension along the wire is variable in the time because it derives from the elastic longitudinal deformation of the tether. It has a transitory, where the free component is extinguished by the viscous term, then it goes to a stationary condition, forced by the gravity gradient term, and oscillates around a mean value about 0.44N. The peaks are generated by rapid temperature variation when satellite passes from sunlight to dark side, and vice versa.

6 Conclusion

An orbital simulator has been developed in order to simulate and investigate the dynamics, thermodynamics and electrodynamics of a bare tethered satellite operating in Earth environment. The simulator can model the tether either like rigid dumbbell system or by means of lump mass approach. The former lets obtain quickly results about the deorbiting performance, while the latter can better approximate the lateral oscillation of the wire. Both the models utilize up-to-date environmental routines for magnetic field,

ionospheric model, atmospheric model and gravity field.

A parametric simulation campaign has been run in order to find out the main destabilizing factors. Some key simulations have been run for an uncontrolled system to highlight the issue of attitude instabilities over time and the need for control.

An energetic control algorithm has been implemented to control the electric current profile along the wire. In this way the EDT system can decrease its orbit till low heights where the atmospheric drag is plenty to complete the reentry.

The lump masses simulator has been verified for its numerical integration accuracy, spectral content and resolution of its dynamic response, and realism of its numerical output. The in-plane, out-of-plane and radial motion of each tether element has been studied in detail revealing that electrodynamic interaction tends to excite only the first vibration modes.

References

[1] Iess, L., Bruno, C., Ulivieri, C., et al., "Satellite De-Orbiting by Means of Electrodynamic Tethers Part I: General Concepts and Requirements", Acta Astronautica Vol. 50, No. 7, April 2002, pp. 399–406.

[2] Iess, L., Bruno, C., Ulivieri, C., and Vannaroni, G., "Satellite De-Orbiting by Means of Electrodynamic Tethers Part II: System Configuration and Performance", Acta Astronautica, Vol. 50, No. 7, pp. 407-416, 2002.

[3] Forward, R. L., and Hoyt, R. P., "The Terminator Tether™: A Spacecraft Deorbit Device", Journal of Spacecraft and Rockets, Vol. 38, No. 2, pp. 187–196, 2000,.

[4] Yamaigiwa, Y., Hiragi, E., and Kishimoto, T., "Dynamic Behavior of Electrodynamic Tether Deorbit System on Elliptical Orbit and its Control by Lorentz Force", Aerospace Science and Technology, Vol. 9, No. 4, pp 366-373, 2005.

[5] Sanmartín, J. R., Martínez-Sanchez, M., and Ahedo, E., "Bare Wire Anodes for Electrodynamic Tether", *Journal of Propulsion and Power*, Vol. 9, pp. 352–320, 1993,.

[6] Peláez, J., Lorenzini, E.C., López-Rebollal, O., and Ruiz, M., "A new Kind of Dynamic Instability in Electrodynamic Tethers", *The Juornal of the Astronautical Sciences*, Vol. 48, No. 4, pp. 449-476, 2000.

[7] Peláez, J., Ruiz, M., López-Rebollal, O., Lorenzini, E. C., and Cosmo, M. L., "A Two-Bar Model for the Dynamics and Stability of Electrodynamic Tethers", *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 6, pp. 1125–1135, 2002.

[8] Cosmo, M., and Lorenzini, E. C., "Electrodynamc Tethers in space handbook", Special Report under Grant NAG8-1160 Third Edition, NASA Marshall Space Flight Center, Dec. 1997.

[9] Colombo, G., et al., "Study of the Dynamics of a Tethered Satellite System (Skyhook)", Final Report, Contract NAS8-32199, NASA Marshall Space Flight Center, Prepared by Smithsonian Astrophysical Observatory, Cambridge, MA, 1978.

[10] French, A. P., "Vibrations and Waves", MIT Introductory Physics Series, W. W. Norton & Company, 1971.

[11] Bombardelli, C., Peláez, J., Sanjurjo, M., "Asymptotic Solution for the Current Profile of Passive Bare Electrodynamic Tethers", Journal of Propulsion and Power, Vol. 26, No. 6, pp. 1291-1304, 2010.

[12] Cleghorn, G. et al., "Orbital Debris: a Technical Assessment", National Academy Press, Washington, 1995.

[13] NASA, "Orbital Debris Program Office", United Nations, New York, 1999, Sales No. E.99.I.17, ISBN 92-1-100813-1.

[14] Bilitza, D., "International Reference Ionosphere 2000", *Radio Science*, Vol. 36, No. 2, pp. 261-275, 2001.

[15] Corsi, J., and Iess, L., "Stability and Control of Electrodynamic Tethers for De-orbiting Applications", *Acta Astronautica*, Vol. 48, No. 5-12, pp. 491–501, 2001.

[16] Takeichi, N., "Practical Operation Strategy for Deorbit of an Electrodynamic Tethered System", Journal of Spacecraft and Rockets, Vol. 43, No. 6, pp. 1283-1288, 2006.

[17] Dobrowolny, M., "Lateral Oscillations of an Electrodynamic Tether", Journal of the Astronautical Sciences, Vol. 50, No. 2, 2002, pp. 125-147.

[18] Somenzi, L., Iess, L., and Peláez, J., "Linear Stability Analysis of Flexible Electrodynamic Tethers", American Astronautical Society Paper AAS03-539, 2003.

[19] Vestroni, F., Luongo, A., and Pasca, M., "Stability and Control of Transversal Oscillations of a Tethered Satellite System", Applied Mathematics and Computation, Vol. 70, 1995, pp. 343-360.

Linearized theory of supersonic potential flow: recent results of studies on Volterra's equation and its applications

Luigi Polito
Department of Aerospace Engineering , University of Pisa , Italy

Keywords: *Potential flow ,Wave drag ,Panel method*

Abstract

Main purposes of the author's recent studies are both to improve the preliminary evaluation of the aerodynamic forces acting on aircrafts at supersonic speeds, particularly the wave drag, and to give some contribution to the potential flow theory. With regard to the induced kinetic potential, both the formula of equivalence between double-layer and boundary / surface vortices for any layer shape and, for its gradient, the formula of equivalence, through exchanges among components, with the induction of sources are given. As, by the impermeability boundary condition, double-layers are sufficient to simulate the flow perturbation, these results allow a panel method formulation to be developed without layer strength discontinuities on the panel borders; this is important for supersonic flows. With regard to the wave drag analysis, evaluation is given of the rate at which the induced velocity comes down when the distance from the aircraft increases.

1 Introduction

The linearized kinetic potential equation

$$(1 - M^2)\varphi_{xx} + \varphi_{yy} + \varphi_{zz} = 0 \qquad (1)$$

plays an important role in the aircraft preliminary design stage, above all because gives plentiful and fundamental analytical results which guide crucial choices for the aircraft configuration, concerning the aerodynamic efficiency in cruise conditions. The fluid velocity can be expressed in the form

$$\boldsymbol{v} = (V + \varphi_x)\boldsymbol{i} + \varphi_y\boldsymbol{j} + \varphi_z\boldsymbol{k} = 0, \qquad (2)$$

where $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ are the unit vectors of the coordinate axes x, y, z , respectively, and V is the flight speed corresponding to the Mach number M.

For the flight at supersonic velocities particularly important are the analytical results for the evaluation of the wave drag [1;2]. When the flight conditions are such that the boundary layer does not anticipate the separation and the trailing edges can be suitably predicted, then

also the calculation by the panel method [3;4;5] for Eq. (1) offers many opportunities at preliminary design stage, particularly for the configuration optimization. In the following the author presents some analytical results of his more recent studies based upon Eq.(1), intended both to improve the evaluation of the aerodynamic forces acting on aircrafts at supersonic speeds, particularly the wave drag evaluation, and also to give some contribution to the general theory. These results have been obtained starting from the expressions which give by Volterra's method [6;7] the perturbation kinetic potential either through the boundary values or through the strength of doublets (of sources). This method does not require the use of the notion of Hadamard's finite part of diverging integrals [8].

In the following the expression "double-layer" indicates a doublets distribution on a regular surface (a "panel" with arbitrary shape) and the expression "singularity" indicates a source, a doublet, etc.

The first part of the presented results concerns the induction of a double-layer. The induction formulas, originally expressed directly through the double-layer local strength, are given expressed through this strength along the border and its gradient on the surface, so demonstrating for the induced potential the equivalence with the induction of vortices [9]. As for the induced velocity components, the reported expression in vector form [10,11] shows the equivalence with the induction of sources, through exchanges among the components (likewise it happens between source and vortex in planar incompressible flow). Analogous results hold for subsonic flow [9;10;11]; for incompressible flow they correspond to results of the Electromagnetism (such as the Ampère's "equivalence principle"). As the perturbation introduced by an impermeable surface in the flow can be represented through the double-layer induction only [11;12], these results allow the development of a simplified formulation of the panel method, with panels covered by doublets, without any sources and singularity strength discontinuities along the panel borders.

The second part of the presented results concerns the analysis of the wave drag. For the velocity induced by an arbitrary singularities distribution inside a bounded region of the space the evaluation of the attenuation due to increasing distance from the singularities is given. The total change of the fluid velocity due to the aircraft is obtained by summing to the aforesaid velocity, computed by the appropriate singularities distribution, the velocity induced by the part of the wake double-layer which is far from the aircraft. As known, according to the small perturbations theory, the part of the wake whose distance from the aircraft is great enough can be represented by half straight lines, whose origins make up a line and which are mutually parallel; the double-layer strength is constant along each of these half straight lines (corresponding to boundary vortices). The induction of this part of the wake is a classical item not considered here.

2 Double layer induction

In supersonic flow the potential induced at the point $P(x, y, z)$ by a double-layer on any regular surface, S , if P is not on S is given by

$$\varphi(P) = \frac{U_x(P)}{2\pi} \qquad (3)$$

$$U(P) \equiv \int_{SP} m(Q)\boldsymbol{n}(Q) \cdot \boldsymbol{G}(P) h dS, \qquad (4)$$

where $Q(u,v,w)$ is the centre of the integration element, that is dS on S and dl on the boundary of S, $m(Q)$ is the double-layer strength, $\boldsymbol{n}(Q)$ is the unit vector orthogonal to S at Q, SP is the part of S such that $x > u + B|\boldsymbol{D}|$, and, moreover,

$$\boldsymbol{G}(P) \equiv -B^2\boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z} \qquad (5)$$

$$B \equiv \sqrt{M^2 - 1} \qquad (6)$$

$$h \equiv ln\frac{x - u + \sqrt{\Omega}}{B|\boldsymbol{D}|} \qquad (7)$$

$$\boldsymbol{D} \equiv (y - v)\boldsymbol{j} + (z - w)\boldsymbol{k} \qquad (8)$$

$$\Omega \equiv (x - u)^2 - B^2|\boldsymbol{D}|^2 \qquad (9)$$

If *CP* indicates the part of the border of *S* where $x > u + B|\boldsymbol{D}|$ and *K* is a suitable constant, when on *CP* is $|\boldsymbol{D}| > 0$ it results, by known identities between integrals and by suitably handling the irregularities,

$$\varphi(P) = K + \frac{I_C(P) - I_S(P)}{2\pi} \qquad (10)$$

$$I_C(P) \equiv \int_{CP} m(Q)\boldsymbol{z}\cdot\boldsymbol{n}(Q)\wedge\boldsymbol{e}(Q)dl \qquad (11)$$

$$I_S(P) \equiv \int_{SP} \boldsymbol{z}\cdot\boldsymbol{n}(Q)\wedge\boldsymbol{g}m(Q)dS \qquad (12)$$

$$\boldsymbol{z} \equiv \frac{x - u}{|\boldsymbol{D}|^2}\frac{\boldsymbol{D}\wedge\boldsymbol{i}}{\sqrt{\Omega}}, \qquad (13)$$

where $\boldsymbol{e}(Q)$ indicates the local unit vector orthogonal both to the border and to $\boldsymbol{n}(Q)$, oriented outside *S*, and $\boldsymbol{g}m(Q)$ indicates the tangential gradient of $m(Q)$. The constant *K* is zero if on *SP* is $|\boldsymbol{D}| > 0$ everywhere, otherwise results by a linear combination of the values of $m(Q)$ at the points of *SP* where $\boldsymbol{D} = 0$; if all these points are intersection points the coefficient of the linear combination is either -1, where $\boldsymbol{n}\,.\,\boldsymbol{i} > 0$, or +1 where $\boldsymbol{n}\,.\,\boldsymbol{i} < 0$; see [9] for more details. This result demonstrates the equivalence with the induction of vortices.

As for the double-layer induced velocity it results [10;11]

$$\nabla\varphi(P) \equiv \boldsymbol{G}(P)\wedge\boldsymbol{Z}(P) \qquad (14)$$

$$\boldsymbol{Z}(P) \equiv \frac{J_C(P) - J_S(P)}{2\pi} \qquad (15)$$

$$J_C(P) \equiv \int_{CP} \frac{m(Q)\boldsymbol{n}(Q)\wedge\boldsymbol{e}(Q)}{\sqrt{\Omega}}dl \qquad (16)$$

$$J_S(P) \equiv \int_{SP} \frac{\boldsymbol{n}(Q)\wedge\boldsymbol{g}m(Q)}{\sqrt{\Omega}}dS \qquad (17)$$

3 Induced velocity far field

Let be

$$X \equiv x/B \qquad (18)$$

$$q \equiv X - A \qquad (19)$$

$$y\boldsymbol{j} + z\boldsymbol{k} \equiv r(\boldsymbol{j}\cos t + \boldsymbol{k}\sin t), \qquad (20)$$

with *A* any positive constant and *r* positive, if not zero. A solution of Eq. (1) such that

$$\varphi(x \le Br, y, z) = 0 \qquad (21)$$

$$\varphi(x \ge BA, y, z) = c.f. \qquad (22)$$

$$\varphi[x \ge max\,(BA, Br), y, z] = r.f.\,, \qquad (23)$$

where *c. f.* indicates any continuous function and *r. f.* indicates any regular function, may represent the kinetic potential induced by a singularities distribution inside the cone ($X > r$, $X < A$). Let be

$$v\boldsymbol{j} + w\boldsymbol{k} \equiv c(\boldsymbol{j}\cos\psi + \boldsymbol{k}\sin\psi) \qquad (24)$$

$$\omega \equiv \psi - t \quad , \qquad (25)$$

with *c* positive, if not zero. Moreover, let be

$$2\pi L(P; f) \equiv \int_{SP} f(c, \psi)H\,dS \qquad (26)$$

$$H \equiv ln\frac{q + \sqrt{\Omega}/B}{|\boldsymbol{D}|}, \qquad (27)$$

with *SP* indicating the part of the circle *S*: ($q = 0$, $c < A$) where $q > |\boldsymbol{D}|$. Following the Volterra's method [6;7] and taking into account that the co-normal of the conical surface $X = r$

lies on the generating half-straight-line, it results [13]

$$\varphi(P) = L_{XX}(P;\varphi_0) + L_X(P;\varphi_{q0}) \quad (28)$$

$$\varphi_0(c,\psi) \equiv \varphi(BA, c\cos\psi, c\sin\psi) \quad (29)$$

$$\varphi_{q0}(c,\psi) \equiv \varphi_X(BA, c\cos\psi, c\sin\psi) \quad (30)$$

Let be

$$F(x,r,t) \equiv \varphi(x, r\cos t, r\sin t) \quad (31)$$

By suitable manipulations, from Eq. (28) the first partial derivatives of F and, consequently , the function

$$Z(x,r,t) \equiv BF_x(x,r,t) + F_r(x,r,t) \quad (32)$$

result given by final expressions containing only integrals which are either proper or improper-convergent and, moreover, devoid of derivatives of integrals [13].

For each of these four intervals of values of r

$$J1: [0 \le r \le R_0 > A] \quad (33)$$

$$J2: [R_0 \le r \le R_1(q)] \quad (34)$$

$$J3: [R_1(q) \le r \le q - A] \quad (35)$$

$$J4: [q - A \le r \le X], \quad (36)$$

where

$$R_1(q) \equiv q - A - e(q) \quad (37)$$

and *e(q)* is any function with positive lower bound and such that

$$\lim_{q\to+\infty} e(q)/q = 0 \ , \quad (38)$$

the aforesaid final expressions give what follows.

With the abbreviation "*b.f.* " indicating any bounded function and the abbreviation

$$k \equiv \sqrt{[q^2 - (r-A)^2]/(4Ar)} \ , \quad (39)$$

for great values of $|q|+r$ the first partial derivatives of $F(x, r, t)$ are given by [13] :

A) inside *J1*

$$F_x = \frac{C^{(x)} + b.f./q}{2\pi B q^2} \quad (40)$$

$$F_r = \frac{C^{(r)} + b.f./q}{2\pi q^3} \quad (41)$$

$$\frac{F_t}{r} = \frac{C^{(t)} + b.f./q}{2\pi q^3} \ , \quad (42)$$

with

$$C^{(x)} \equiv -\int_S \varphi_{q0}(c,\psi)\, dS \quad (43)$$

$$C^{(r)} \equiv \int_S \varphi_{q0}(c,\psi)(r - c\cos\omega)\, dS \quad (44)$$

$$C^{(t)} \equiv -\int_S \varphi_{q0}(c,\psi) c\sin\omega\, dS \quad (45)$$

B) inside *J2*

$$F_x = \frac{qE^{(x)}}{R_{h0}^3} \quad (46)$$

$$F_r = \frac{rE^{(r)}}{R_{h0}^3} \quad (47)$$

$$\frac{F_t}{r} = \frac{E^{(t)}}{R_{h0}^3} \ , \quad (48)$$

with

$$E^{(..)} = b.f. + \frac{qb.f.}{R_{h0}^2} + \frac{rb.f.}{R_{h0}^2} + \frac{qrb.f.}{R_{h0}^4} \quad (49)$$

$$R_{h0} \equiv \sqrt{q^2 - r^2} \quad (50)$$

C) inside *J3* $(k \ge 1)$

$$Z = \frac{b.f.}{r\sqrt{r}}\boldsymbol{P}(k) + \frac{b.f.}{r\sqrt{r}} \qquad (51)$$

$$F_r = \frac{b.f.}{\sqrt{r}}\boldsymbol{P}(k) + \frac{b.f.}{\sqrt{r}} \qquad (52)$$

$$\frac{F_t}{r} = \frac{b.f.}{r\sqrt{r}} \quad , \qquad (53)$$

with

$$\boldsymbol{P}(k) \equiv \int_0^{1/k} \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}} \qquad (54)$$

D) inside *J4* $(k \leq 1)$

$$Z = \frac{b.f.}{r\sqrt{r}}\boldsymbol{K}(k) + \frac{b.f.}{r\sqrt{r}} \qquad (55)$$

$$F_r = \frac{b.f.}{\sqrt{r}}\boldsymbol{K}(k) + \frac{b.f.}{\sqrt{r}} \qquad (56)$$

$$\frac{F_t}{r} = \frac{b.f.}{r\sqrt{r}} \quad , \qquad (57)$$

with

$$\boldsymbol{K}(k) \equiv \int_0^{1} \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}} \qquad (58)$$

As for the evaluations relevant to the intervals *J3* and *J4* is to be noted that for k approaching 1, i.e. for r approaching q - A , both the functions $\boldsymbol{K}(k)$ and $\boldsymbol{P}(k)$ have the singularity ln | k-1 | , i.e. the singularity ln | r-(q-A) |. In equations (51), (52), (55), (56) the terms containing either the function $\boldsymbol{P}(k)$ or the function $\boldsymbol{K}(k)$ are given by line integrals on the boundary of the circle *S*. For the moment their evaluation is based only on the continuity of the potential on this boundary; the analysis of the behaviour for k near to 1 of these terms has not been performed yet. Moreover is to be noted that the value of the function $Z(x, r, t)$ comes down more quickly of the values of both the axial and radial velocity components when r increase. This can be inferred from the examination of Eq. (1) expressed by cylindrical coordinates

$$(1-M^2)F_{xx} + F_{rr} + \frac{F_r}{r} + \frac{F_{tt}}{r^2} = 0 \qquad (59)$$

by assuming that this equation can be replaced for great values of r by the equation

$$(1-M^2)F_{xx} + F_{rr} = 0 \qquad (60)$$

whose continuous solutions such that $F(x < Br, r, t) = 0$ are given by $F = T(x - Br , t)$ where T is an arbitrary function, so that

$$BF_x + F_r = 0 \qquad (61)$$

4 Conclusions

The results presented contribute both to improve the valuation of the aerodynamic forces acting on aircrafts at supersonic speeds when the linearized potential flow equation is used, particularly the evaluation of the wave drag, and to the general potential flow theory.

Acknowledgements

This report describes developments of researches carried out thanks to the financial support of the fund of the University of Pisa (Italy) (*fondi di Ateneo dell'Università di Pisa*) . The author will be grateful to the readers for any critical remark they can point out directly to him.

References

[1] Hayes W. D., "Linearized supersonic flow", Thesis, California Institute of Technology, Pasadena California, 1947. (Thesis, reprinted as North American Aviation Report AL-222, 1947).

[2] Kulfan B., "New Supersonic Wing Far-Field Composite-Element Wave-Drag Optimization Method", J. of Aircraft, Vol. 46, No. 5, 2009, pp.1740-1758.

[3] Morino L.,“A general theory of unsteady compressible potential Aerodynamics”, NASA CR-2464, 1974.

[4] Polito L., Lombardi G., Bulletti G., “Metodo dei pannelli: attualità ed applicazioni per flusso subsonico tridimensionale con risultati di una nuova formulazione”, Atti del XIX Congresso Nazionale AIDAA, Forlì (FC),F-075 pp. 1-12.

[5] Epton M.A., Magnus A.E., “PAN AIR- A Computer Program for Predicting Subsonic or Supersonic Linear Potential Flows about Arbitrary Configurations Using a Higher Order Panel Method”, NASA CR-3251, 1990.

[6] Volterra V., “Leçons sur l’intégration des équations différentielles aux derive partielles-professées a Stockholm”, Hermann et fils, Paris, 1912, Leçons 8,9.

[7] Tricomi F.G., “Equazioni a derivate parziali”, Cremonese, Roma, 1957, Chap. 2.

[8] Ward G.N., “Linearized theory of steady high-speed flow”, Cambridge Univ. Press, 1955, Chap. 3.

[9] Polito L., “On the subsonic and supersonic double-layer potential for the linearized calculation of steady high-speed flow”, Atti del Dipartimento di Ingegneria Aerospaziale, Università di Pisa, ADIA 2010-5, 2010.

[10] Polito L., “Some equivalence theorems in the linear theory for subsonic and supersonic flow”, Atti del Dipartimento di Ingegneria Aerospaziale, Università di Pisa, ADIA 2003-1, 2003.

[11] Polito L. ,“Analytical contributions to the calculation of linearized subsonic and supersonic flows and to the panel method”, Atti del XX Congresso Nazionale AIDAA, Milano (MI), AE 19 pp.1-20.

[12] Polito L., “Contribution to the panel method for sbsonic and supersonic flows” , Atti del Dipartimento di Ingegneria Aerospaziale, Università di Pisa, ADIA 2000-4, 2000.

[13] Polito L. , Report in preparation.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

The Influence of Load Factors on the Turn Manoeuvres of Agricultural Aircraft: An Overview

Bosko Rasuo
University of Belgrade, Faculty of Mechanical Engineering, Aeronautical Department, Belgrade, Serbia

Keywords: *Load Factors, Safe Turn Manoeuvres, Agricultural Aircraft.*

Abstract

In this paper, a theoretical study of the turn manoeuvre of an agricultural aircraft is presented. The manoeuvre with changeable altitude is analyzed, together with the, effect of the load factors on the turn manoeuvre characteristics during the field-treating flights. The mathematical model used describes the procedure for the correct climb and descent turn manoeuvre. For a typical agricultural aircraft, the numerical results and limitations of the climb, horizontal and descending turn manoeuvre are given.

1 Introduction

The problem of turning flight with changeable altitude is described by the system of differential equations which describe the influence of the normal and tangential load factors on velocity, the path angle in the vertical plane and the rate of turn, as a function of the bank angle during turning flight [1-6].

The system of differential equations of motion was solved on a personal computer with the Runge-Kutta-Merson numerical method. Some analytical and numerical results of this calculation are presented in this paper [6].

Ariel treating has become an integral part of modern agriculture. However, agricultural flying has been plagued by a great number of accidents. The manoeuvres of an agricultural aircraft are divided into those carried out while entering or leaving the spraying line. Despite low height of the aeroplane while spraying, this part of the flight is considered to be the safest because of an appreciable flight velocity. Upon completing a spraying run, the aircraft enters the turn manoeuvre procedure.

The low altitude of the turn manoeuvre procedure and the vicinity of terrain are main causes of numerous accidents. The pilots assert that the procedure is the most dangerous manoeuvre and most of the accidents occur when these manoeuvres are performed, due to low altitude and the possibility of stall during the manoeuvres. There are many secondary factors that influence the safety of agricultural flying and some of them will be treated in this paper (see Figure 1.).

2 Analytical and Numerical Calculation

In accordance with the above mentioned considerations, two segments of the turn procedure are considered to be critical: (i) climb while turning, and (ii) rolling. In these phases of the manoeuvre, the aircraft may encounter stalling. The aim of this paper is to determine those factors that provoke the appearance of stall and to estimate their magnitude. A further aim is to investigate the most convenient turn manoeuvre procedure regarding the safety of the manoeuvre and its economical realisation. Furthermore, the analytical and numerical calculation methods are analyzed [6].



Fig. 1 An agricultural aircraft in flight.

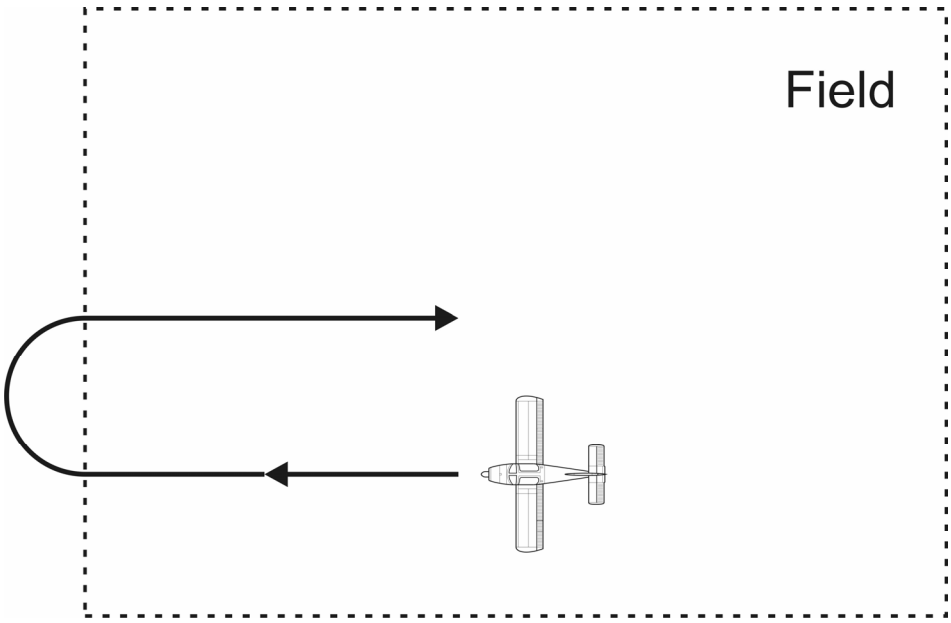


Fig. 2 Turn procedure.

The analysis has been restricted to the climb in the first phase, and the descent at the end of the turn manoeuvre. A special type of the turn manoeuvre has also been analyzed, i.e. the case in which the x-direction forces are balanced. This case involves a close balance between

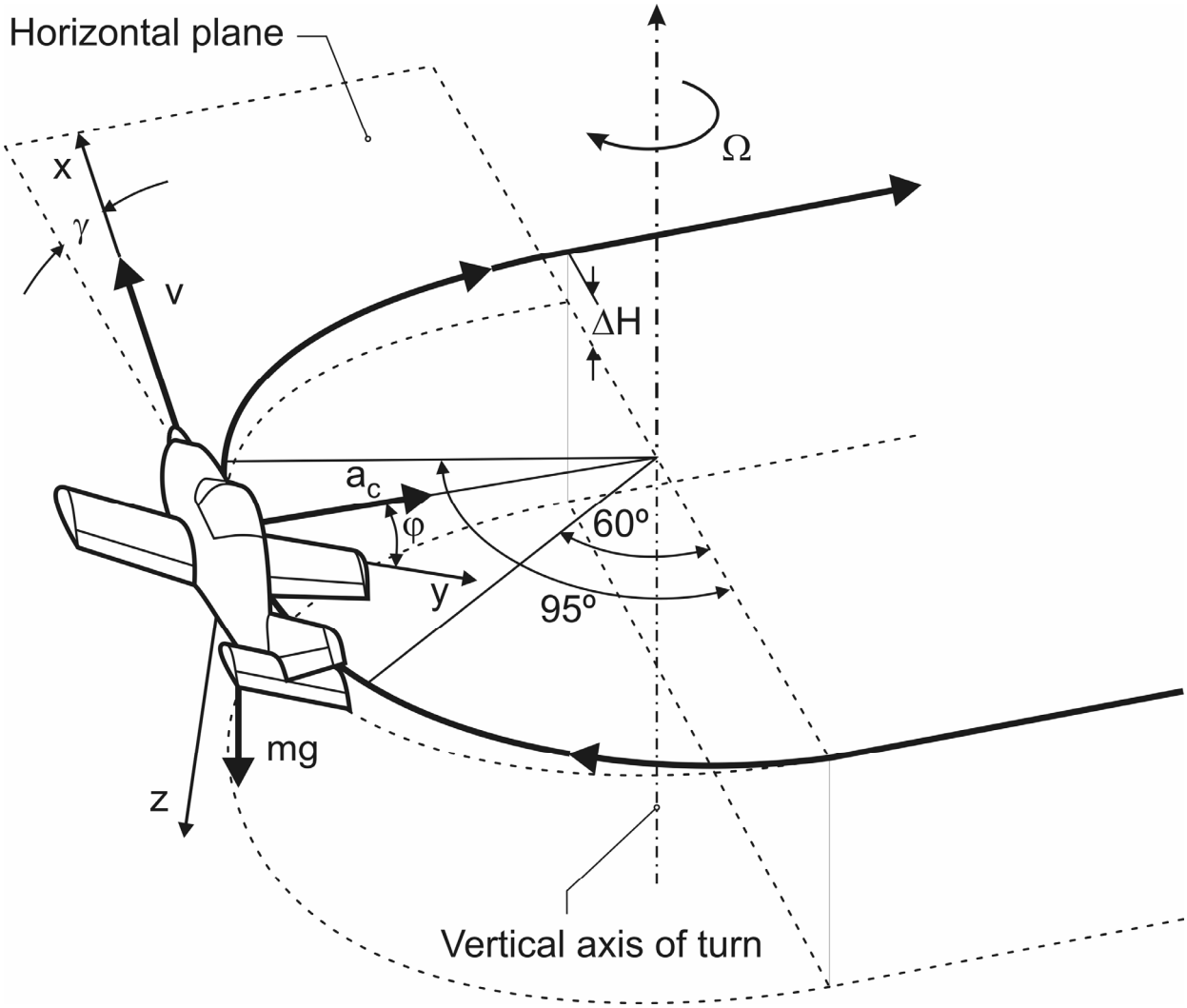


Fig. 3 The climb turn of agricultural aircraft.

thrust and drag (T ≈ D). The analytical solutions for the climb turn manoeuvre at a constant normal load coefficient have been obtained and analyzed [6].

For the analysis and calculation of the turn manoeuvre, the normal and tangential load factors have been chosen, because these combine all essential parameters for the turn manoeuvre procedure (energy potential – powerplant, aerodynamic and structural characteristics of aircraft, as well as the manner of piloting an aeroplane). That leads to universal solution of the turn manoeuvre parameters for all categories of the agricultural aircraft.

The usual treatment (spray) run is carried out along the larger dimension of the field. At the end of a treatment run, a turn procedure is made, and the pilot enters the treating line shown in Figures. 2 and 3.

The analysis and calculation of the climb (and descent) turn manoeuvre, by involving the normal and tangential load factors, provides general results. In this way, observation and analysis of behavior of all categories of agricultural aircraft are possible.

The analysis shows that the stalling speed and the start altitude at the beginning of the turn manoeuvre are limiting factors. For this kind of aircraft, the stalling speed of approximately 30 ms^{-1} is a typical value. Accordingly, at the end of the treating (spraying) run, and beginning of

the turn manoeuvre, the velocity should exceed 40 ms^{-1}(see for more details in reference [6]).

3 Conclusion

In this paper a criterion has been established for determining the margin of the speed over the stalling speed (safety speed) after the agricultural aeroplane increases its height above the ground by a given amount. The required altitude may be different for different types of aeroplanes and will depend on their capability to perform rolling and turning manoeuvres. The minimum margin of velocity allowed is determined from the requirements of the rolling manoeuvre [6].

References

[1] Etkin, B. and Reid, L.D., *Dynamics of Flight*, 3rd ed., John Wiley & Sons, New York, 1996, pp. 238-242.

[2] Mair, W.A. and Birdsall, D.L., *Aircraft Performance*, (Reprint), Cambridge University Press, 1996, pp. 214-236.

[3] Vinh, N.X., *Flight Mechanics of High Performance Aircraft*, (Reprint), Cambridge University Press, 1999, pp. 221-252.

[4] Phillips, W.F., *Mechanics of Flight*, John Wiley & Sons Inc, Hoboken, New Jersey, 2004, pp. 281-299, 804-819.

[5] Filippone, A., *Flight Performance of Fixed and Rotary Wing Aircraft*, 1st ed., Elsevier, Oxford, 2006, pp. 251-265.

[6] Rasuo, B., "Some analytical and numerical solutions for the safe turn manoeuvres of agricultural aircraft – an overview", Aeronaut J, September 2007, Vol. 111, (1123), pp. 593-599.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Benchmark Study of High Speed External Flow over Rotating Missile-Like Body using First- and Second-Order Turbulence Models

A. Ridluan, W. Charubhun, and P. Chusilp
Aerodynamics Group, Design and Simulation Laboratory, Defense Technology Institute, Nonthaburi, Thailand 11120

Keywords: *Compressible flow, Missile design, and Missile aerodynamics*

Abstract

This paper presented benchmark study of missile-like body using the Steady Reynolds-Averaged Navier-Stokes equations (SRANs) with first-order turbulence models (standard k-ε *and Renormalized Group (RNG)* k-ε, SST k- ω, *and transitional* SST k-ω *models) and a second-order turbulence model, (Reynolds Stress Model, RSM). The Mach number investigated ranges from 0.6 to 4.5.*

The numerical showed that all turbulence models were able to predict the trend of axial coefficients. However, based on the percent disagreements computed, the standard k-ε *and Reynolds Stress Model (RSM) agreed very well with the experiment database. The deviation of CFD from experimental data was observed at low speed Mach flow (Mach number of 0.6). The percent disagreements between experimental and computational data are approximately 10.*

Based on visualization, all turbulence can capture the development of shock and downstream on wake flows. For low speed incoming Mach number, the local shock pocket was observed. The local shock on the missile-like body was stronger and stronger as the incoming shock increased, resulting in the increase of the aerodynamic drag coefficient. At Mach number 1.02, the bow shock was observed and this shock moved close to the body at Mach number greater than 1.02, attributed to the reduction of the drag coefficient.

1 Introduction

Defense Technology Institutes (DTI) is initiating a number of activities in missile design and development. One of the key initiatives is design and optimization of missile aerodynamics, which is responsible by Aerodynamics group. To meet the institute goals, the Aerodynamics group has developed the roadmap including conceptual, preliminary, and detail designs as shown in Fig. 1. In conceptual design phase, the analysis involves a trial and error process applying simple analytical and empirical relations to meet constraints in order to achieve the design objective.

To improve traditional design optimization process in conceptual phase, we thus proposed developing CFD-integrated optimal design in the preliminary design. In this phase, CFD plays a major role to generate dataset of the objective design parameters. A Response Surface (RS) based polynomial and Artificial Neural Network

(ANN) are purposely applied to construct the so-called Surrogate model.

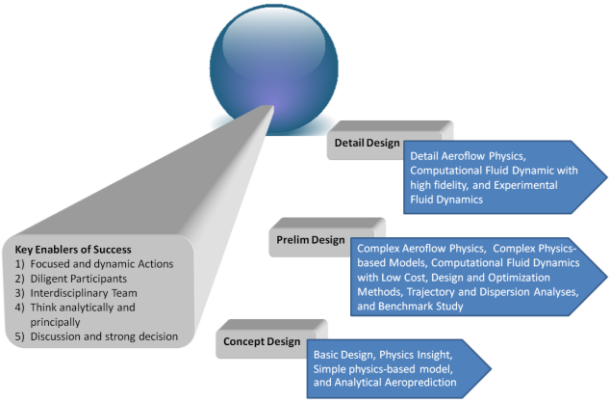


Fig. 1 Roadmap for Aerodynamic design and optimization (Ridluan)

Briefly, the application of ANN approach is to reduce a number of CFD simulations and to enhance the construction of a RS equation. Finally, the anticipated aerodynamically-designed missile configuration is investigated using very high reliable means such as CFD simulation with high fidelity and Experimental Fluid Dynamics (EFD).

It is clearly demonstrated in the road map itself that to trustfully integrate CFD methodology as the design tool in both preliminary and detail design phases, benchmarking CFD simulation against generally accepted experiments is the significant first step. A particular challenge in turbulent flow around a missile is the accurate aerodynamic prediction of turbulent flow. As the accurate simulation is depended on the turbulence model implemented, the exercise in using CFD as the design and optimization tool is one of evaluating the application of the appropriate turbulence model in terms of its accuracy and time consuming.

The present paper thus objectively investigates accuracy of CFD prediction for the flow field of rotating projectile using Steady Reynolds-Averaged Navier-Stokes (SRANS) equations with first-order turbulence models (standard k-ε and RNG k-ε, *SST* k- ω, and transitional *SST* k- ω models) and a second-order turbulence model (Reynolds Stress model, RSM) at Mach numbers ranging from 0.6 to 4.5. In this simulation, we applied the multi-reference frame approach to model the spinning motion of the projectile.

2 Mathematic model

In order to predict turbulent flow, the modification of Navier-Stokes equations (NSEs) is needed. The modified equations are often called the Reynolds-Averaged Navier-Stokes (RANS) equations.

The Reynolds averaging process introduce the unknown term. Here, the unknown terms are called the Reynolds Stress terms and defined as,

$$\tau_{ij} = -\rho \cdot \overline{u_i' \cdot u_j'} \qquad (1)$$

The traditional approach in modeling Reynolds Stress is to employ the Boussinesq hypothesis in order to relate the Reynolds Stress with the velocity gradients. However, Boussinesq hypothesis leads to a new unknown term, turbulent viscosity term.

In order to simulate turbulent flow, the turbulent eddy motion is here described the turbulent viscosity term. We note that in contrast to the molecular viscosity, the eddy viscosity depends on the flow dynamics. Therefore, selection of the appropriate turbulence model characterizing the eddy viscosity is directly linked to the accuracy of the simulation relative to a benchmark experiment or simulation.

To objectively assess the accuracy of the simulations of turbulent flow over a missile-like body, here three popular first-order turbulence models are initially chosen. The models in this study are as follows: standard k-ε, RNG k-ε, SST k-ω, and Reynolds Stress Model (RSM). The turbulent model for each model is briefly described below.

For brevity, we refer the reader to classic texts on turbulent theory and models such as Tennekes and Lumley (1976) and Rodi (1993). We consulted additional brief notes on turbulent flow models in manuals for FLUENT.

3 Detailed Formatting Instructions

In order to be able to apply commercial CFD codes in the missile design and optimization, all codes need to be validated and verified against experimental data from a flow configuration similar to the anticipated design flow. In the present case, we report the results to date on flow through the missile-like body, defined by conical and cylindrical elements (see Fig. 2).

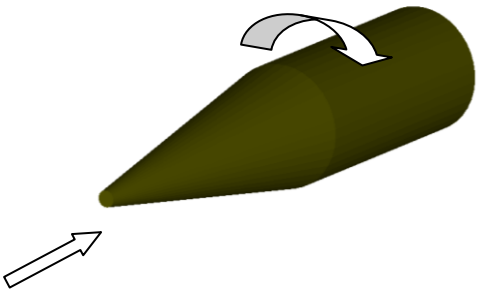
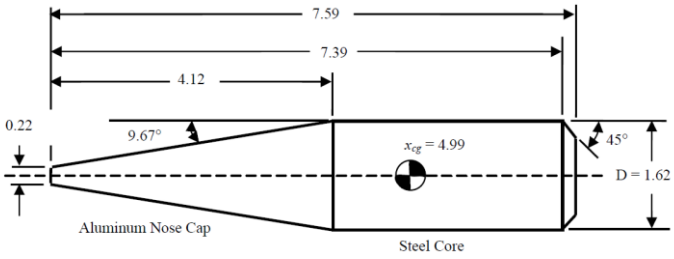


Fig. 2 The schematic of the missile-like body flow

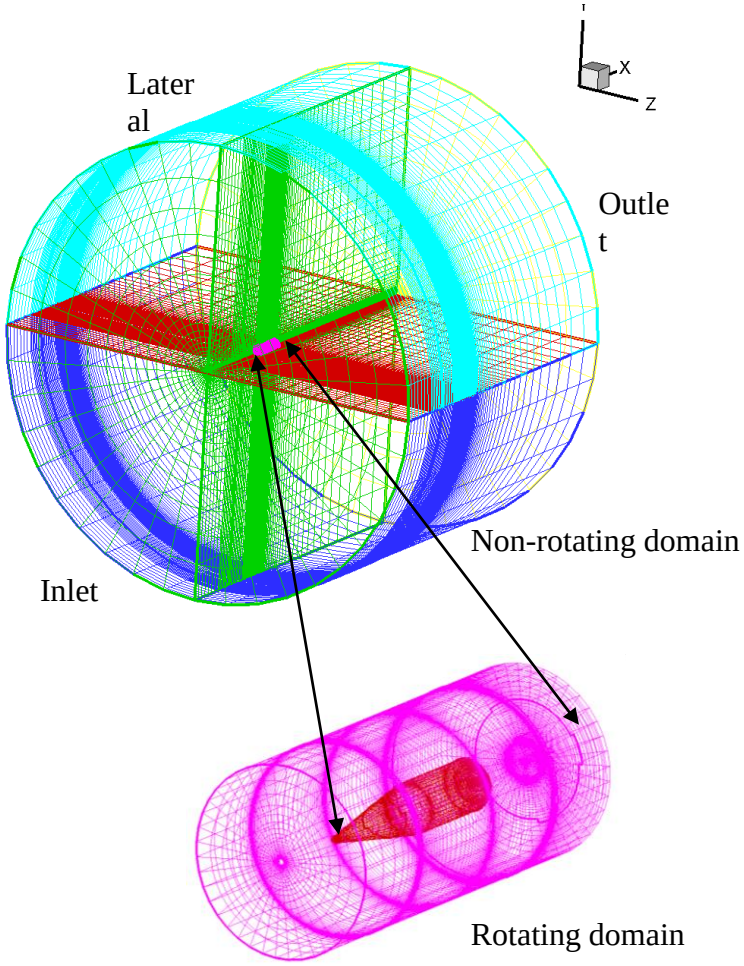
The flow over the missile-like body shown schematically in Fig. 2, reproduces the configuration of flow through a M910 subprojectile by Plostins et al. [3]. This investigation and the dataset which is accessible is widely accepted as a benchmark for CFD codes. DeSpirito and Heavey [4] provided additional information on the flow modeling. Fig. 3 (a) shows the dimensions and unit cell of the experimental configuration. The detail dimensions were schematically given by DeSpirito and Heavey [4]. The model consists mainly of conical and cylindrical sections with 4.12 and 3.27 centimeters long. The nose diameter and cylindrical diameters are 0.22 and 1.62 centimeters, respectively.

In the present study, we mainly used the commercial CFD code, the codes use a time-independent compressible Navier-Stokes formulation and the turbulent models that are discretized using the finite volume method. Further, a second-order upwind and central differencing flow numerical schemes were applied for convective and diffusive terms, respectively. The discrete nonlinear equations were implemented implicitly. To evaluate the pressure field, the ideal gas law was used. The linearized equations were solved by multi-grid method. The numerically approximated equations were performed on the collocated viscous mesh, approximately totaling 500,000 cells for the cell shown in Fig. 3 (b). For viscous mesh topology, the structured hexahedral cells are generated in the boundary layer region, of which y^+ of the missile surface was less than 5 to ensure the resolution of the viscous sublayer.

To model the rotating effect, the external domain is defined as non-rotating frame, while the inner domain is rotated at constant rotational speed. The computational grids of external and rotating regimes were constructed separately as shown in Fig. 3 (b). The interface mesh generated between two dolmans is conformal.



(a)



(b)

Fig. 3 Schematic configuration of (a) the M910 subprojectile and (b) the computational domains and unit cell

The boundary conditions used in the research were as follows. A streamwise far field condition was applied at the inlet and lateral boundaries. This specified constant Mach

number. The outlet boundary condition is taken as nonreflecting boundary condition. At the interface between rotating and non-rotating domains, the flow properties and turbulent quantities are circumferentially averaged.

4 Results and Discussion

Since aerodynamic drag is important to determine missile flight characteristics such as range, speed, maneuverability, propulsion size, and missile carriage [5]. The axial force coefficient was then selected to evaluate the predictive ability of the fist- and second-order turbulence models. Here, the first-order turbulence models selected are standard *k-ε* and RNG *k-ε*, *SST k- ω*, and transitional *SST k- ω* models. For second-order turbulence model, Reynolds Stress Model (RSM) was chosen.

4.1 Quantitative analysis

Fig. 4 exhibits the plots of axial force coefficients against Mach numbers. It is evidential that the coefficient abruptly increases from subsonic to transonic regimes and gradually decreases from transonic to supersonic flows.

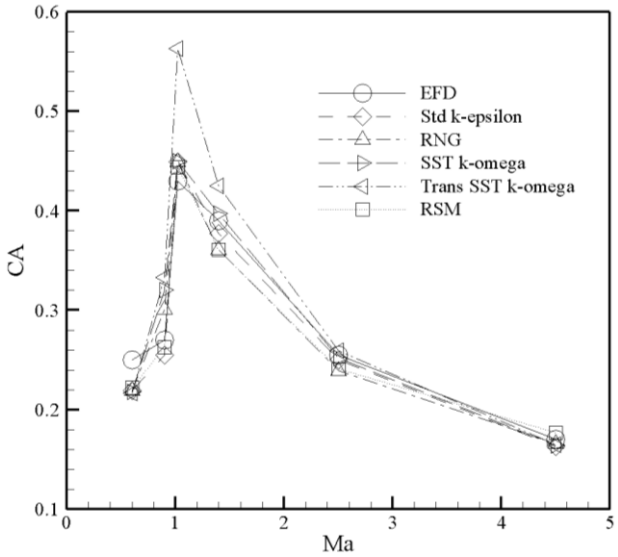


Fig. 4 The plots of experimental and computational Axial force coefficient (C_A) versus Mach number (Ma)

Based on the Figure above, all turbulence models well predict the trend of the axial coefficients. However, the transitional SST k-*ω* poorly predicts the maximum axial coefficient at Mach number 1.02.

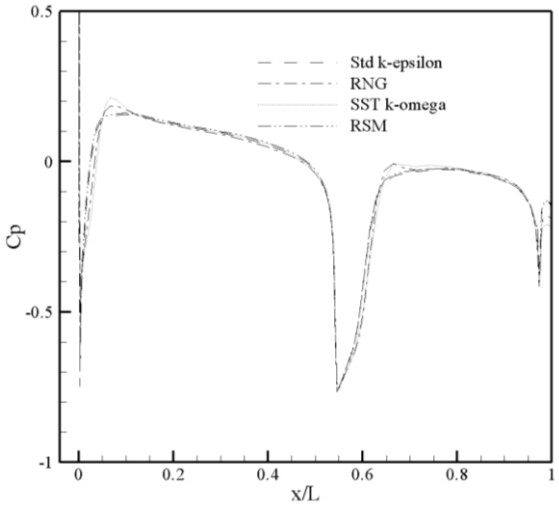
To quantitatively assess the suitability of turbulence model, the percent disagreement is computed and given in Table 1. Surprisingly, the standard *k-ε* turbulence model shows the best capability in predicting the axial coefficients. Generally, the standard *k-ε* model disagrees with the measured data by 5% whereas. We also observe that at low speed flow, the disagreement between predicted and measured data is by 13%.

Table 1 The difference of axial force coefficients between EFD and CFD

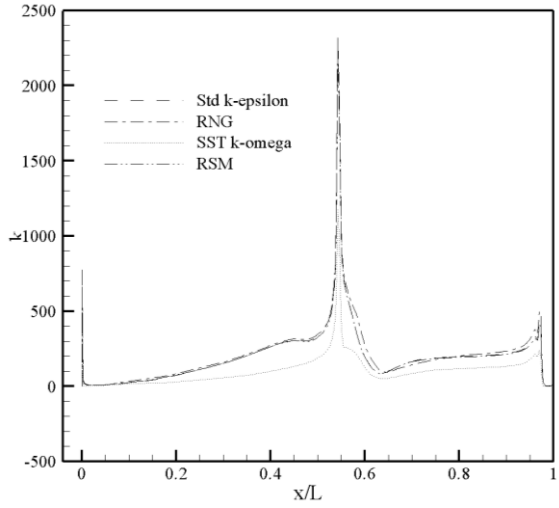
Ma	EFD&CFD Differences [(EFD-CFD)/EFD]%				
	Std k-epsilon	RNG	SST k-omega	Trans SST k-omega	RSM
0.6	12.400	12.400	12.800	13.200	11.200
0.9	5.556	-11.481	-18.519	-23.333	2.593
1.02	-4.419	-4.419	-4.651	-30.930	-3.256
1.4	3.333	7.436	-1.795	-8.974	7.436
2.5	1.961	6.275	1.176	-1.569	5.490
4.5	4.118	2.353	3.529	2.941	-4.118

4.2 Flow physics analysis

The flow around the projectile at Ma of 0.9 is first analyzed. Fig. 5 demonstrates the distributions of the pressure coefficient (C_p) and turbulence kinetic energy (k) along the projectile centerline. Based on Figs. 5 (a) and (b), there are the sudden decrease and increase of the pressure coefficient and turbulence kinetic energy at x/L = 0.54 and 0.97. We note that the positions at x/L=0.54 and 0.97 are where the geometrical transition exists; that is, from conical ogive to cylindrical body and from cylindrical body to ogive boat tail, respectively. It can observe that all turbulence models predict similar trends. However, the lowest pressure coefficients are predicted by SST *k- ω*.



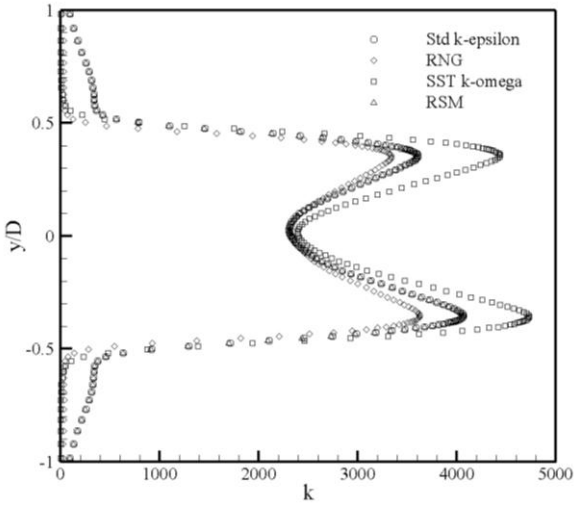
(a)



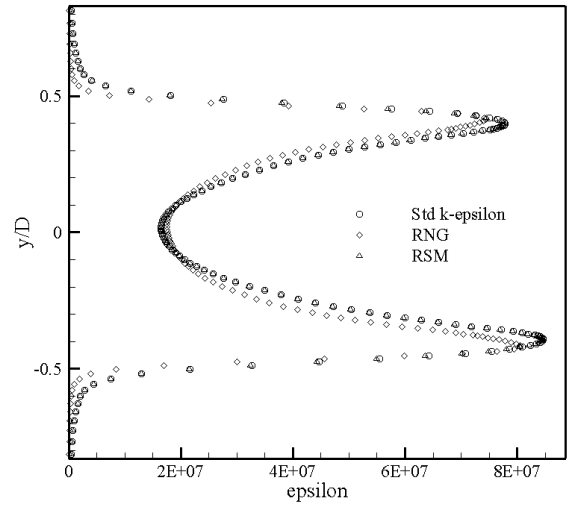
(b)

Fig. 5 The centerline plots of (a) pressure coefficient and (b) turbulence kinetic energy along the projectile.

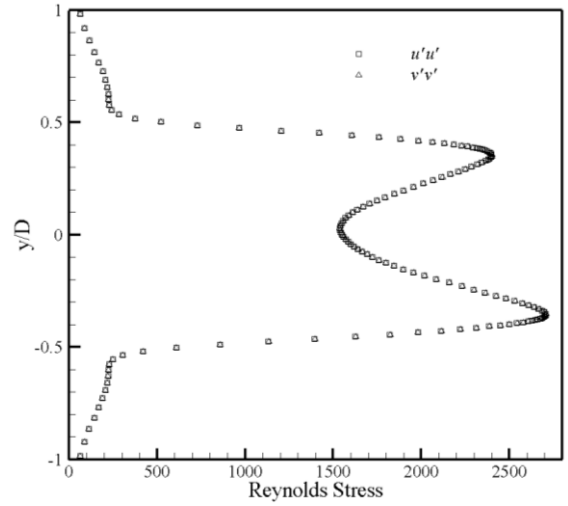
The turbulence flow in wake region is also investigated. As shown in Figs. 6 (a) and (b), all turbulence models provide the similar trends of kinetic energy and dissipation rate. The maximum kinetic energy and dissipation rate of turbulence locate inside recirculation flow (we define this flow pattern as the very large eddies) in the wake region. The predicted kinetic energy of turbulence flow using Standard k-ε and RSM is identical. SST k- ω predicts the higher turbulence kinetic energy, while RNG yields the slightly-smaller kinetic energy value, comparative to the calculation based on standard k-ε and RSM.



(a)



(b)

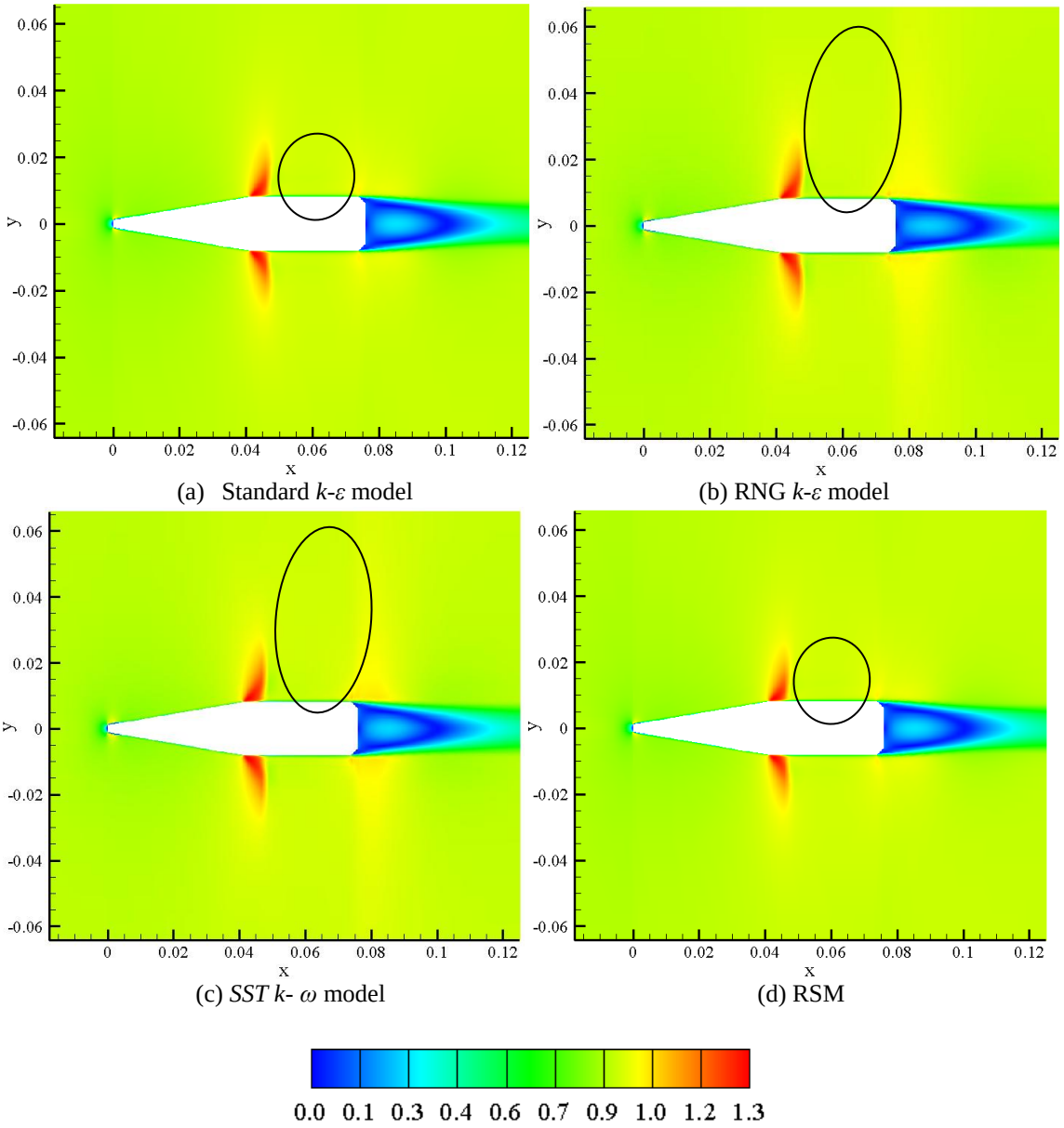


(c)

Fig. 6 The plots of (a) turbulence kinetic energy, (b) epsilon, and (c) Reynolds stress in wake region.

We further investigate the Reynolds Reynolds stress behavior in the wake region as shown in Fig. 6 (c). It was found that both streamwise and spanwise Reynolds stresses are identical, the evidence indicating that a turbulence flow developed in wake region is isotropic. Similar to turbulence kinetic energy prediction, the dissipation of turbulence predicted by standard k-ε and RSM are identical and it is slightly higher than those computed using RNG model. Note that the wake flow pattern is unsymmetrical due to the body rotation effect.

Figs. 5 to 6 display the shock flow development when the incoming Mach numbers are 0.9, 1.02, and 4.5, respectively. The contour plot of Mach number for incoming Mach number 0.9 reveals the local shock (pocket) development at location of transitional location from the nose to cylinder parts. Further downstream, the local shock observed at the onset of aft part is developed. Predicting by RNG and SST k- ω shown Figures 5 (b) and (c), the local shock is stronger than that computed by Standard k-ε and RSM models, shown in Figures 5 (a) and (d). This leads to the overprediction of the coefficient using RNG and SST k- ω turbulence models (the percent disagreement is greater than 10%). For incoming Mach number 1.02, the bow shock is developed at the nose of projectile and the local shock at the cylinder and aft bodies are clearly observed. Once the Mach number reaches 4.5, the local shocks downstream disappear and the bow shock tends to attach the missile-like body.



(a) Standard *k*-*ε* model (b) RNG *k*-*ε* model

(c) *SST k*- *ω* model (d) RSM

Fig. 5 The shock development at Mach number of 0.9 using (a) standard *k*-*ε*, (b) RNG *k*-*ε*, (c) transitional *SST k*- *ω* mode, and (d) RSM.

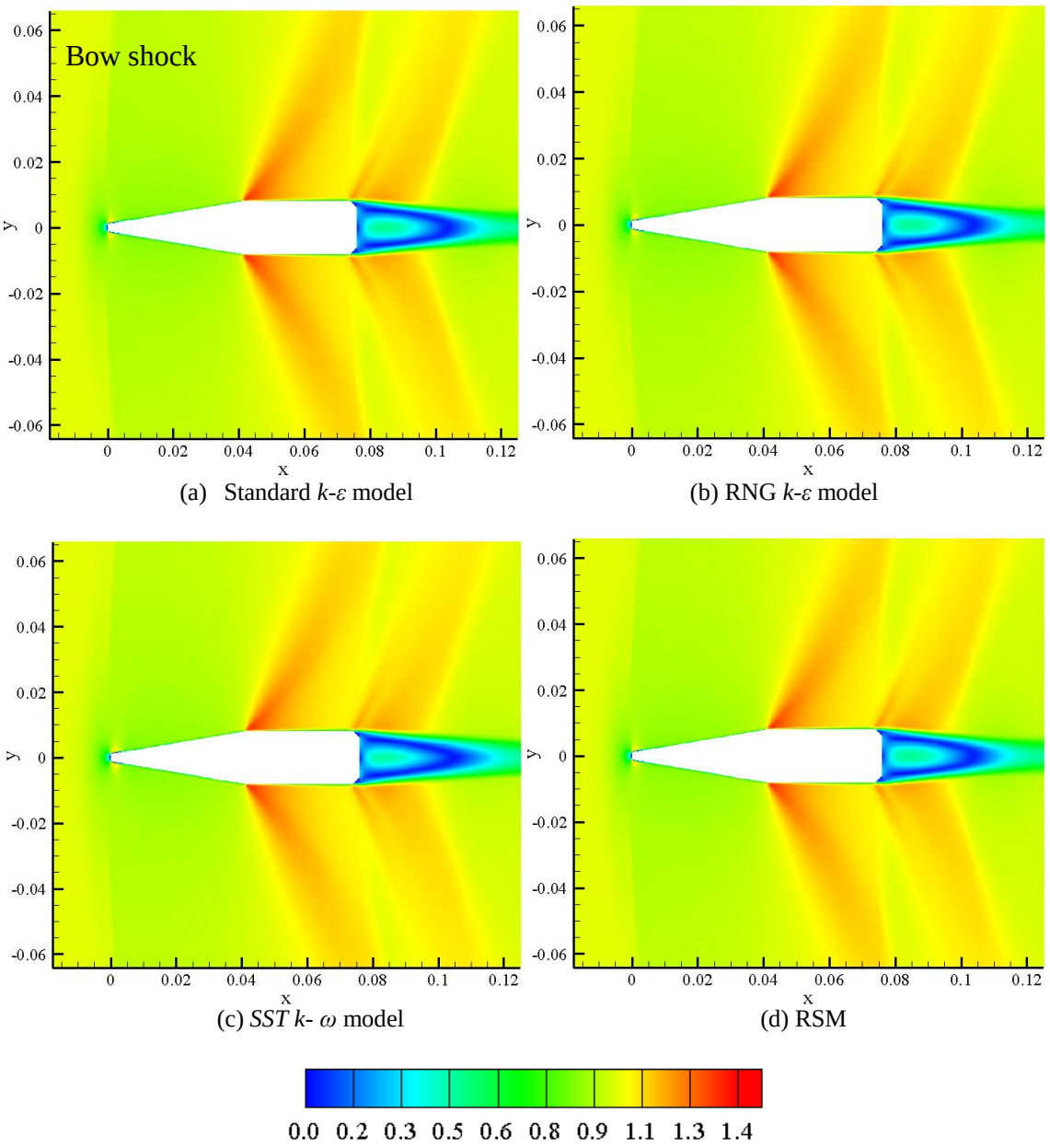


Fig. 6 The shock development at Mach number of 1.02 using (a) standard *k-ε*, (b) RNG *k-ε*, (c) transitional *SST k- ω* mode, and (d) RSM.

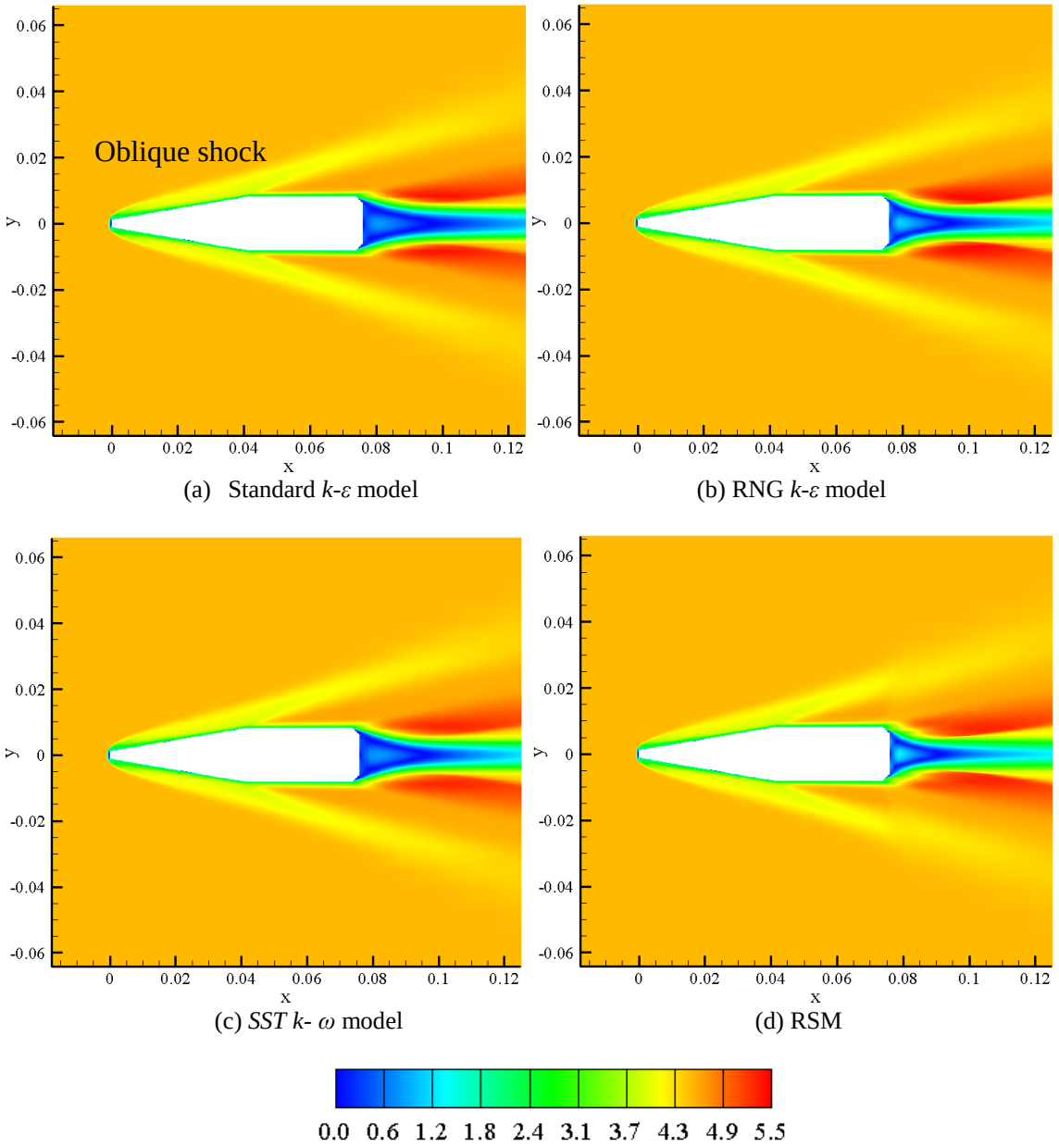


Fig. 7 The shock development at Mach number of 4.5 using (a) standard *k-ε*, (b) RNG *k-ε*, (c) transitional *SST k- ω* mode, and (d) RSM.

5 Conclusion to date

In order to use CFD codes with confidence as a tool for missile design ad optimization, it is of importance to validate and verify the codes against accepted benchmark experiments similar to the anticipated flow configuration. Of special interest in this exercise is learning the capability of various turbulence models to accurately simulate the anticipated complex high speed compressible flows.

In the present work, we began verifying and validating the FLUENT codes against the benchmark experiment of missile-like body, as experimentally and computationally reported by [3,4]. With four first-order turbulence models (standard k-ε and Renormalized Group (RNG k-ε), *SST* k- ω, and transitional *SST* k- ω models) and a second-order turbulence model (Reynolds Stress Model, RSM), we learned the following in summarized form

- that the axial coefficient trend can be simulated using both first-order and second-order turbulence models. We found that the transitional *SST* k- ω model give some of similar trends but was not found to compare well with experimental data.
- that both first-order and second-order turbulence models can simulate mostly the same "flow physics" such as shock and downstream wake flow. The RNG and *SST* k- ω model give some of similar trends but was not found to compare well with experimental data.
- that the standard k-ε and RSM models quantitatively give the coefficient results most similar to experimental coefficient. However at low speed Mach number, none of turbulence models can reproduce the experimental coefficients very well.
- that based on the Reynolds stress comparison, the turbulence flow around the missile-like body is isotropic.

The V&V exercise is as stated ongoing. We plan to report additional results at a future venue.

References

[1] Tennekes H. and Lumley J.L., *Introduction to turbulence*, MIT Press, Cambridge, MA USA, 1976

[2] Rodi W., *Turbulence models and their applications to hydraulics*, IAHR Monograph, Taylor Francis, 1993

[3] Plostins P., McCoy L. R. and Wagoner A. B., "Aeroballistic performance of the 25 mm M910 TPD-T range limited training projectile," U.S. Ballistic Laboratory, 1991.

[4] DeSpirito J. and Heavey R. K. 2006, "CFD computation of magnus moment and roll damping moment of a Spinning Projectile," U.S. Army Laboratory, 2006.

[5] Mendenhall M. R., *Tactical missile aerodynamics: prediction methodology*, American institute of aeronautics and astronautics, 1991, Chaps. 1.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

An attitude testbed for satellite formation flying research

A. Caon, M. Cesaro, E.C. Lorenzini
C.I.S.A.S. G. Colombo, University of Padova, Italy

Keywords: *satellite formation flying, spacecraft simulator.*

Abstract

Spacecraft formation flying technology is opening new scenarios for future space missions. It is a relatively young technology and it is a field of ongoing research.

This document describes the design features, and the design choices, of a testbed that allows to carry on experimental research on spacecraft formation flying in an earthbound laboratory. The work is part of the spacecraft formation flying related activities conducted at the University of Padova.

The facility developed is composed by two spacecraft, each representing a spacecraft of a simple formation, and it will allow to study the three dimensions attitude dynamics.

1 Introduction

This document summarizes the work done so far at the C.I.S.A.S. G. Colombo – University of Padova to develop a facility to study satellite formation flying in a ground laboratory environment.

The expression "satellite formation flying" indicates a group of satellites flying in a coordinate and cooperative way to reach a common task. Satellites flying in formation can stay very close each other or very far. In satellite formation flying knowledge of relative attitude and position of the satellites is an important parameter, this is the most distinctive feature between formations and constellations (such as GPS). Usually very accurate determination and very fine control of the relative position and of the relative attitude are fundamentals for the mission success.

EO-1 and Landsat 7 were the first satellites to accomplish a formation flight, it happened in 2000 when the two satellites flew following the same ground track with a 1 minute separation [1]. From then on other formation flying missions have been launched, for issue: GRACE (to map Earth's gravity field measuring the distance between two satellites), ATrain (a scientific mission composed by 4 satellites, at this moment, to study Earth's atmosphere), Prisma, Proba 3, some of them are scientific missions, others are for technological demonstration. GRACE task is to map Earth's gravity field measuring the distance between two satellites [2]; ATrain [3] is a scientific mission composed by four satellites, at this moment, to study Earth's atmosphere. Prisma, by the Swedish National Space Board, and Proba 3, by ESA, were launched to validate sensors, actuators and navigation strategies.

Other missions are planned for launch during the next years, as Terrestrial Planet Finder Interferometer and Darwin to look for extrasolar planets, LISA to measure gravitational waves.

Formation flying is considered an enabling technology for many future missions since it will allow to reach unprecedented performance levels, both for Earth observation and for outer space missions. Fields of application of formation flying will be: interferometry, gravity field mapping, high resolution imaging, data relaying, self-assembling satellites, on-orbit replenishment, Earth observation.

The advantages of formation flying are: increased accuracy, increased flexibility, redundancy, cost reduction, upgrade. The formation can be reconfigured in order to accomplish different tasks during the mission, the mission can continue even with a failure to one of the satellites, the formation can be upgraded adding new satellites. The main drawbacks are: increased control and communication needed and increased propellant consumption. Relative attitude and position must be controlled (in addition to the absolute ones), usually with high accuracy, this means to consume more propellant too.

Formation flying has not demonstrated all its capabilities yet and a lot of work must be done to improve it; according to Bauer [4] it is possible to identify the following fields for formation flying improvement:

- sensors development;
- actuators development;
- communications;
- control strategies;
- computing;
- testbeds.

The work described in this document deals with testbeds development for formation flying.

The next section describes various kinds of facilities, the testbed develop is described in chapter number 3.

2 Testbeds

Testbeds development represent focus research area for formation flying technology improvement. Testbeds are used to validate performance and capabilities of new technologies developed or navigation and attitude sensors, actuators, controls strategies, telecommunications and data handling.

Reliability and safety are key requirements for a space mission, since usually is not possible to fix issues that can arise in space. Careful experimental tests campaign need to be conducted to verify requirement satisfaction, experimental tests support numerical simulations and mathematical models allowing to overcome their restrictions.

An ideal testbed would be able to replicate all the features of the space environment that a spacecraft experiences during its operations, nevertheless testbeds can't be fully representative of the space environment, each testbed focuses on the replication of certain feature of the environment depending on what the testbed is used for. Testbeds to study control algorithms for satellite formation flying, for issue, don't focus on thermal effects.

Testbeds used to study orbital dynamic and attitude can be classified in:

- ground-based testbeds;
- parabolic flights;
- ISS testbeds;
- on-orbit testbeds.

In parabolic or 0-g flights a microgravity situation inside a plane is caused by suitable manoeuvres of the plane itself. Parabolic flights can simulate space microgravity and they allow the system to test to free float inside the plane, microgravity lasts about 20s and this could be a too short time for many research activities.

On-orbit testbeds consist of real space missions, as EO-1, Orion and Mustang missions, whose main target is not a scientific one but a technical one, for example testing sensors, actuators or control algorithms. On-orbit testbeds are obviously fully representative of the environment in which a satellite or a formation of satellites will operate, but they are really expensive too, they are much more expensive than the others testbeds.

In the International Space Station, payloads (such as SPHERES [5]) to test can free float inside the ISS, so it's possible to test them in a real, endless 0-g condition; however these tests are expensive because of the need to carry the

payloads on the Space Shuttle from Earth to the ISS.

Ground-based testbed, as the name implies, are placed at ground level and they can be divided in three categories:

- spacecraft simulators;
- water tanks;
- drop towers.

Drop towers are high buildings (Bremen drop tower, is 123m high) used to simulate microgravity. The system to study is loaded on the top of the building and it is released, at the bottom of the tower airbags or other devices are placed to arrest the experimental payload. Air inside the building is evacuated to create a low pressure environment in order to reduce aerodynamic drag. Microgravity condition last a few seconds, form 2s to 10s. Because of building cost and air evacuation cost tests using drop towers turn out to be enough expensive.

Water tanks testbeds enable microgravity conditions to last as long as needed; nevertheless the system to test must be water-resistant and the viscous forces and torques are not negligible. Water tanks are really useful in astronauts training activity.

Spacecraft simulators are systems that replicate the space environment in a earthbound laboratory. They are composite of mechanical and electrical hardware and software and they are use in space missions research and design. Simulators allow to realize "hardware-in-the-loop-simulation", that is to evaluate the behaviour of the real hardware (as sensors and actuators) that is going to be used on a satellite. They are mainly used to study orbital and attitude dynamic and control of a single spacecraft or of a formation.

Spacecraft simulators allow to simulate a microgravity environment with minimal external torques acting on the system, usually they don't take in account space environment vacuum and thermal cycles. Their primary concern is to represent spacecraft dynamics faithfully.

Compared to parabolic flights, water tanks and drop tower, spacecraft simulators don't enable payload to experience the same rotational and translational motion freedom but this is not a limitation in most cases. External torques and forces acting on a spacecraft simulator payload are lower than viscous forces and torques in water tanks.

Tab. 1 testbed comparison

Testbed	Pros	Cons
On-orbit testbeds	fully representative of operative conditions	very expensive
ISS testbeds	excellent microgravity simulation it's easy to interact with the payload long lasting microgravity	very expensive logistical problems
Parabolic flights	excellent microgravity simulation it's quite easy to interact with the payload	expensive microgravity last about 20s
Drop towers	good microgravity simulation	microgravity lasts few seconds expensive
Water tanks	good microgravity simulation long lasting microgravity condition	high viscous forces and torques need for a water resistant payload
Spacecraft simulators	good microgravity simulation it's very easy to interact with the payload	disturb forces and torques need to be mitigated partial motion of freedom

Test conducted using spacecraft simulators are cheaper than those conducted using drop towers

or parabolic flights, microgravity conditions last for a longer time and interaction with test system is easier.

Spacecraft simulators can be classified in: rotational systems, planar systems and combined systems depending on which degrees of freedom are free and which are bond. Rotational spacecraft simulators would ideally allow full rotational freedom of the payload without external torques action on it. Rotational freedom can be provided by spherical air bearings [6], ball-and-socked-joints or multiple gimbals. Planar systems allow payload to translate on a plane and to rotate along the axis perpendicular to the plane itself, that means two translational and one rotational degrees of freedom. Combined systems are a mixture of preceding systems allowing 4 or 5 degrees of freedom.

3 University of Padova research facility

The targets of the work are to design, realize and validate a facility to carry on research activities in the spacecraft formation flying field.

The testbed is depicted in Fig. 1: it is made up by two elements, representing the spacecrafts of a formation, each element has 3 rotational degrees of freedom (d.o.f.) in order to simulate attitude dynamics. The rotational motion freedom is ±41° for roll and pitch and yaw is unconstrained. For the time being the two elements have the rotational freedom, translational freedom is planned to be added soon: each element will translate on a plane using air cushion, allowing to study coupled attitude and position dynamics and control. Each of the two elements of the facility (Fig. 2) is a spacecraft simulator, so each of them will be called simply simulator from now on, testbed or facility will indicate the whole system of two simulators. A laptop is part of the facility too, its functions are to bootload the simulators, to send commands to them and to receive and store telemetry data sent by the simulators.

The main requirements that led the facility design are: fatefully representation of a satellite formation, easy configuration changes, low cost, reliability and safety.

It's possible to identify the following subsystems in each simulator:

- structure;
- attitude determination;
- propulsion;
- power;
- avionics;
- communication.

The next paragraphs describe the subsystems.

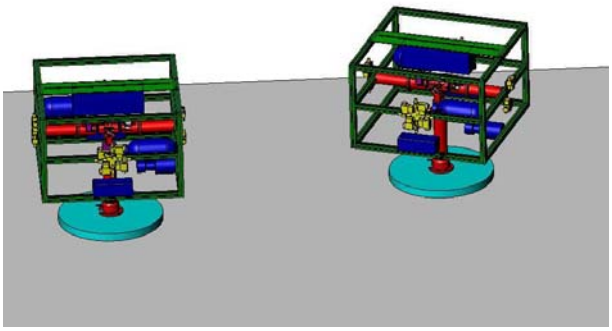


Fig. 1 facility CAD representation

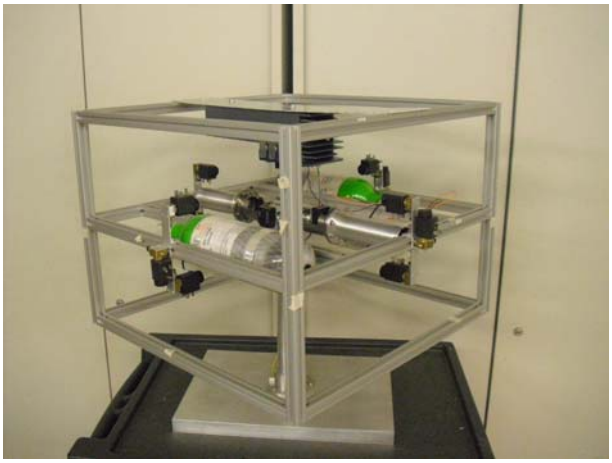


Fig. 2 one of the two spacecraft simulators

3.1 Structure

The functions of the structure subsystem are to support all the components of the simulator and to allow the rotational motion.

The structure is composed by an aluminium frame and a system of rotational joints. The aluminium frame configuration can be easily changed, this feature made easier the developing phase and it's useful during normal operation too. The dimensions of the frame are 50x50x30cm, almost all the simulator components are arranged on it.

The rotational motion is achieved using 3 consecutive rotational joints, each of them allows payload to rotate respect to one axis. This systems of joints is designed so that the first the first joint allows rotation about the vertical axis (coinciding with the yaw axis), the second joint allows rotation respect to the pitch axis and the third joint allows rotation about the roll axis (Fig. 3 and Fig. 4). This configuration allows to measure the attitude directly by measuring axes rotation (see next paragraph). The components of the joints are realized using Al7075 and Ti-8Mn to achieve low mass and stiffness at the same time. Spherical SKF air bearings were used to ensure low friction, all the bearings are preloaded.

Using rotational joints was preferred over an air bearing because they are much cheaper than an air bearing and because they are more suitable to a system that will translate on a plane.

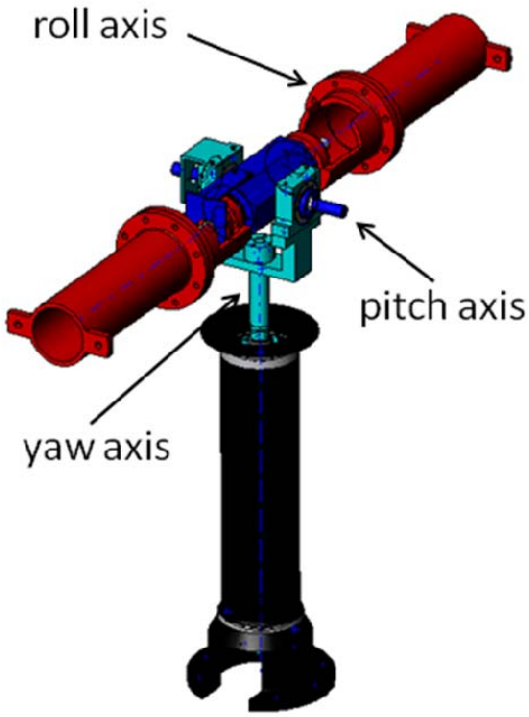


Fig. 3 gimbal system for attitude motion

3.2 Attitude determination

Each simulator is equipped with two kind of sensors to measure attitude: one inertial measurement unit (IMU) and three encoders.

The IMU is the Microstrain 3DM-GX1, it uses 3 accelerometers, 3 rate gyros and 3 magnetic field sensors to compute the attitude. An embedded microcontroller allows to compute the date measured by the sensors and to have in output directly the IMU orientation. All the data from the IMU sensors are available through a serial RS-232 port in digital format. Table 2 summarizes the IMU characteristics.

Tab. 2 IMU sensor properties

Range (°/sec)	+/- 300
In-Run stability, fixed temp. (°/sec)	0.1
In-Run stability, over temp. (°/sec)	0.7
Short term stability (°/sec)	0.02
Angle random walk, noise (°/√hour)	3.5
Scale Factor Error (%)	0.5
Nonlinearity (% FS)	0.2
Resolution (°/sec)	0.01
G-sensitivity (°/sec/*g*)	0.01
Alignment (°)	0.2
Bandwidth (Hz)	30
Range (*g*)	+/- 5
In-Run stability, over temp. (m*g*)	10
Short term stability (m*g*)	0.2
Noise (m*g*/√Hz rms)	0.4
Scale Factor Error (%)	0.5
Nonlinearity (% FS)	0.2
Resolution (m*g*)	0.2
Alignment (°)	0.2
Bandwidth (Hz)	50
Range (Gauss)	+/- 1.2
In-Run stability, over temp. (mGauss)	15
Noise (mGauss/√Hz)	TBD
Scale Factor (%)	0.7%
Nonlinearity (% FS)	0.4
Resolution (mGauss)	0.2
Alignment (°)	0.2
Bandwidth (Hz)	50

Three Avago Technologies HEDM 5500-B13 rotational optical encoders are used to measure the rotation of the joints (Fig. 4), as stated in the previous paragraph the joint axes coincide with the yaw, pitch and roll axes, in this way encoders provide directly the simulator attitude. Using quadrature encoding the resolution of the encoder measurement is 0.09°, and the accuracy is by far better than the one achievable with the IMU, so encoder

measurements are used as a reference for the IMU measurements.

Data from these sensors are acquired by the avionics described in the next paragraph.

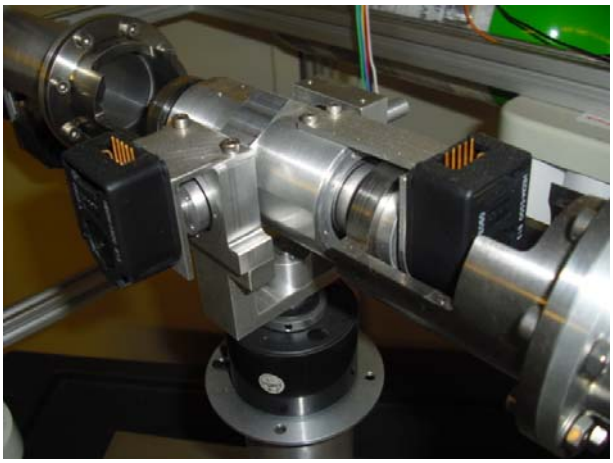


Fig. 4 joint detail and encoders

3.3 Avionics and communication

Avionics functions are: to acquire data from the sensors, to compute the acquired data, to generate suitable commands for the thrusters and to allow communication. Avionics is part both of the attitude determination subsystem and of the communication subsystem. Furthermore it interfaces with the propulsion subsystem and with the power one.

Avionics is based on the Digi International RabbitCore 4510w with a Rabbit 4000 29.49MHz microprocessor. The Digi Rabbit RIO is used to interface the microprocessor to the encoders, allowing for quadrature decoding.

Custom electronic was developed to power the avionic and the other electric components and to interface the avionics and the propulsion subsystem.

The communication subsystem allows wireless communication between the two simulators and between them and the laptop in a peer-to-peer network. Each simulator send messages containing its own state, communication are always in a broadcast mode, in this way the laptop can receive all data exchanged by the simulators. Communication is based on a ZigBee protocol, it was designed for low power, low cost, low data rate personal area network [8].



Fig. 5 RabbitCore 4510w

3.4 Propulsion

Propulsion subsystem allows the actuation of the commands from the control algorithms. Two possible solutions were considered for the propulsion system: one was to use little fans operated by electric motors, the other was to use pneumatic thrusters. The former would be simple and light, but preliminary studies [7] showed it causes unacceptable interferences on the IMU measurements, so a pneumatic system was used.

Air was preferred over helium as propellant gas, since it allows to operate the simulators for a longer time before to refill the tanks. Using helium would have meant an higher specific impulse, but its density is so low that too little gas would be stored in the tanks, so the operating time would be too short. Actual operating time is about 3minutes when four thrusters are always open.

12 thrusters are used to generate torque around three independent axes, each thrusters is made up of an UNIVER U1 electro-valve and a nozzle, every thrusters can generate 0.4N fixed thrust.

Propellant is stored in 2CFRP tanks, with a total capacity of 2liters, at a pressure of 200bars. Air flows from the tanks to a pressure reducing regulator with an outlet pressure of 10bars. A relief valve is used for overpressure protection. From the regulator air flows to the thrusters through the low pressure circuit. A PPE-P10A pressure gauge is used to check the correct operation of the propulsion system.

4 Conclusion

Research activities about spacecraft formation flying are being carried on at the

University of Padova. To allow experimental studies a suitable facility is being developed.

For the time being the facility is made up of two identical spacecraft simulators with fixed positions and three dimensional rotational freedom. This configuration allows to experimentally study attitude dynamics of a simple formation. The main use of the facility will be to allow to conduct experimental tests to validate the studies being conducted, regarding especially attitude determination, dynamics and control.

All the subsystems of the testbed were described, focusing on their features and explaining the design choices.

The facility will be improved providing to the simulators the capability to translate on a low friction table, in order to study coupled dynamics and control. Future activities will include obviously the use of the testbed for research activities, related to satellite formation flying or other distributed systems.

References

Example for Journals references

[1] Guinn, J., "Autonomous navigation flight demonstration results for the New Millenium Program EO-1 Mission", *AIAA/AAS Astrodynamics Conference Monterey, CA, USA*.

[2] Reigber C., Schmidt R., Flechtner F., König R., M. Ulrich et al., "An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S", *Journal of Geodynamics*, Vol. 39, No. 1, 2005, pp 1-10.

[3] Stephens G. L., Vane D. G., Boain R. J., Mace G. G., "The Cloudsat Mission And the A-Train", *Bulletin of the American Meteorological Society,* Dec. 2002, pp. 1771-1790.

[4] Bauer F. H., Hartman K., How J. P., Bristow J., Weidow D., Busse F., "Enabling Spacecraft Formation Flying through Spaceborne GPS and Enhanced Automation Technologies", *ION-GPS Conference, Nashville, TN, September 15, 1999.*

[5] Otero A.S., Chen A., Miller D. W., Hilstad M., "SPHERES: Development of an ISS Laboratory for Formation Flight and Docking Research", *IEEE Aerospace Conference*, 2002, #81.

[6] Schwartz J. L., Peck M. A., Hall C. D., "Historical Survey of Air-Bearing Spacecraft Simulators", *Journal of Guidance, Control, and Dynamics*, Vol. 26, July–August 2003, pp. 513–522.

[7] Caon A., Bettanini C., Lorenzini E.C. "Sviluppo e messa a punto di un simulatore di assetto per uso didattico", *Convegno Nazionale di Misure Meccaniche e Termiche,* Roma, 2010, C17.

[8] Gislason D., *ZigBee wireless networking*, Newnes, U.S.A., 2010, Chaps. 1.

Safety Aspects of Light Aircraft Spin Resistance Concept

R. Hankers, F. Pätzold, T. Rausch, S. Schönhals
Institute of Flight Guidance, Technische Universität Braunschweig, Germany

R. Kickert
Leichtwerk Ingenieurbüro Dr. Kickert, Braunschweig, Germany

M. Cremer, J. Troelsen
messWERK GmbH für Mess- und Anzeigetechnik, Braunschweig, Germany

Keywords: *General Aviation, Safety, Flight Mechanics, Flight Test, Certification*

Abstract

Federal Aviation Regulation FAR part 23 permits an applicant to decide whether the aircraft concerned shall be certified as recoverable from spins or as spin resistant. This study, funded by the European Aviation Safety Agency EASA, investigates safety aspects of the concept of spin resistance. It includes a literature research, an analysis of accident data, expert opinions gathered in interviews and during a workshop, and a survey among European manufacturers and authorities. Additionally, flight trials are performed.

The study reveales that, to prevent stall/spin-related accidents, spin resistance can only be one section within a range of possible measures, besides stall warning systems, envelope protection measures and concepts for pilot training.

Investigations show that the level of safety of the U.S. spin resistance regulation is not sufficient because operational accident situations related to spin resistant aircraft are not covered completely by the representative manoeuvres which have to be performed for certification. A requirement code containing essential additions is proposed.

1 Background

As background for the origination of the research project, EASA stated [1] that "traditionally, non-aerobatic light aircraft have been designed so that recovery can be made from spins. Up until 1991, both the Federal Aviation Administration FAA and European codes placed restrictions on the number of turns or the time it should take to recover from a spin, and ensuring that even if an incorrect recovery procedure was used, straight and level flight could still be recovered.

Studies in the USA (…) showed that in the vast majority of accidents attributed to spinning, the initial altitude was insufficient for recovery. It was concluded by the FAA that it would be safer to try to prevent spins rather than ensure recovery can be made once a spin had been entered. The FAA introduced an amendment to FAR 23.221 in 1991 [2], updated in 1996 [3], which allowed aircraft to be certified with spin resistance as an alternative to meeting the spin recovery requirements. No similar amendment has been introduced to the European Codes."

"It has been found that features in this design intended to make the aircraft spin resistant are detrimental to spin recovery, to the extent that aircraft do not meet the original requirements which only deal with spin recovery."

2 Review on the concept of spin resistance

2.1 Literature review on previous research on spin resistance

The concept of spin resistance is based on results of a research programme by NASA on spinning General Aviation aircraft. Literature reviews revealed that research activities were targeting at technical measures to prevent aircraft from entering spins [4]-[8]. The activities were initially dealing with configurations, wing modifications and stall deterrent systems.

Artificial stall/spin prevention was judged feasible and even "extremely effective" [4] for GA aircraft, but the questions of cost, failure modes, weight, performance penalties and pilot acceptance remained to be addressed. It was stated that the rapid evolution of microprocessors and advanced electromagnetic control actuators may help to accelerate the implementation of such systems [5].

Research activities then concentrated on wing modifications, and here on one technical measure: the modified outboard leading edges (see Fig.1 and Fig. 2) . The abilities of this measure were analysed by trying to enter a spin by performing a set of manoeuvres with varying parameters such as speed rate, aileron position or power setting. The flight trials revealed that the measure was mostly able to prevent the aircraft from entering a spin. But they also revealed that even aircraft with this modification may enter a spin if a high

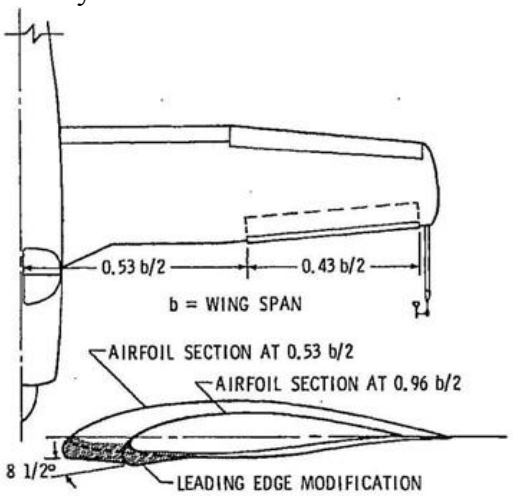




Fig. 1 PA-28X with outboard wing modification [6], [4]

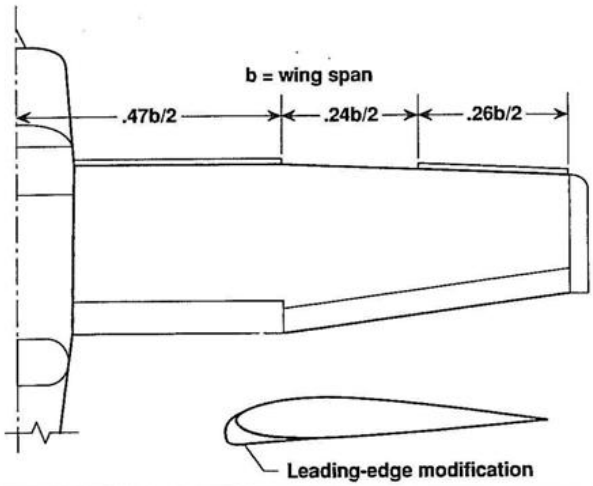




Fig. 2 Cessna C-172X with outboard wing modification [7], [8]

Manoeuvre		wings level zoom, steep climb rudder prior to stall						wings level zoom, steep climb rudder at stall						wings level 1 kn/s rudder at stall						30° banked turn 1 kn/s rudder at stall	
aileron position		neutral		against		with		neutral		against		with		neutral		against		with		neutral	
rudder		left	right	left	right	left	right	left	right	left	right	left	right	left	right	left	right	left	right	turn left	turn right
power	flaps																				
idle	0°																				
	40°																				
for level flight	0°																				
	40°																				
maximum	0°			spin						spin				spin	spin						spin
	40°																				

244 manoeuvres
9 loading conditions:

CG (%)	23,97	25,94	27,78	27,9	27,87	28,23	28,42	29,98	32,07
mass (lb)	2678	2678	2656	2690	2829	2677	2670	2690	2690

colour scheme:
inside certified loading envelope
outside certified loading envelope
covered by current FAR23.221(a)(2)

Fig. 3 Comparison of flight test results of PA-28X with wing modifications and FAR-requirement

power setting is applied, especially when combined with a high rate of speed reduction. As an example, Fig. 3 shows those combinations of manoeuvres and settings within the flight test matrix of NASA's modified PA-28X where spinning occurred, compared to the flight test matrix that has to be performed to comply with the attempted spin entry manoeuvre of FAR 23.221.

The aim was to significantly reduce the number of stall/spin-related accidents as a result of this measure being brought into widespread operation. To this end, a requirement for implementation into FAR23 was formulated based on the abilities of the modified outboard leading edges. To simplify the certification process, and because the modification was found to make recovery from spins more difficult, no proof of recoverability from spins is necessary if an applicant aspires certification as spin resistant. As the majority of stall/spin-related accidents starts at low height where the recoverability from spins is irrelevant, this can be interpreted as an "exchange deal": Preventing stall/spin-related accidents, especially those starting at low height, can justify that proof of recoverability is omitted, which, regarded separately, represents a decrease in the level of safety as regards stall/spin-related accidents starting from larger height.

2.2 Accident situations and statistics

Compared to the time period in which the NASA and FAA activities were carried out, it is now possible to analyse some years of operational experience with spin resistant aircraft, as two four-seat single engine non-aerobatic US aircraft are certified to this alternate FAA code (Cirrus SR20/22 and the Cessna 300/350 family). Two main results can be found when analysing accident data [9]. First, the proportion of stall/spin-related accidents among all fatal accidents is of the same order as, or even higher than with conventional aircraft. Second, there are no principal differences neither in the distribution of occurrences within the traffic pattern (cp. Fig. 4) nor in the sequence of events during the occurrences. It can be concluded that the aim to reduce stall/spin-related accidents is not achievable with the existing two spin resistant aircraft designs.

A general analysis of stall/spin-related accidents revealed that accidents occur distributed during both the descending and ascending phases of the traffic pattern, and revealed hints on steep pull ups and high bank angles in the sequence of events of many accidents.

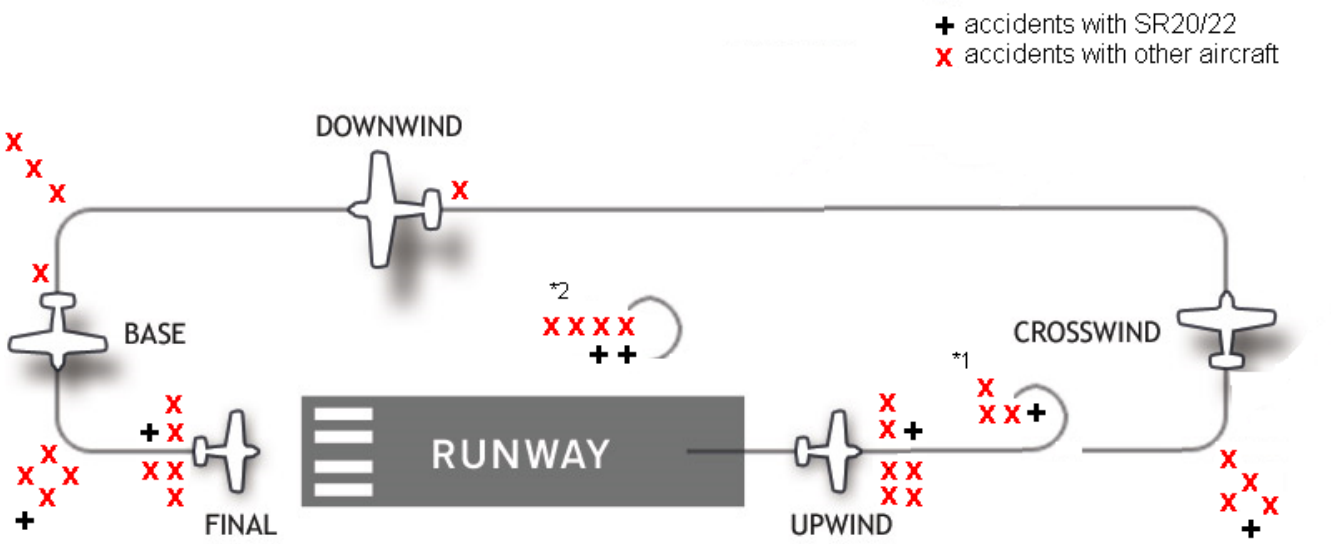


Fig. 4 Comparison of positions of stall/spin-related accidents within the traffic pattern
(depiction based on [10])

2.3 Questionnaire and interviews with experts

The topic was also discussed with manufacturer-independent experts in a workshop, in interviews, and with European manufacturers and authorities during a survey. The answers to the survey showed that the topic can and needs to be considered under a range of aspects. The survey also embedded the European manufacturers into the process of creating a European spin resistance requirement right from the beginning and helped to increase awareness for the topic of spin resistance and the intended rulemaking process.

During the interviews with the experts, a wide range of aspects was discussed. In summary, the main statements concentrated on operational accident situations, which are not covered by the current FAR spin resistance requirement, on the necessity for improved stall warning systems and envelope protection measures and on the necessity to combine technical measures with concepts for pilot training.

Based on these findings, it can be concluded that a wider range of possible measures should be considered with the aim of reducing stall/spin-related accidents, starting from flight mechanical measures to influence the behaviour of the aircraft at stall to improved stall warning systems, envelope protection measures and concepts for pilot training. The central aspects to be considered are the situational awareness of the pilot and the sequence of events during an operational accident. The concept of spin resistance can only be one issue within this range.

3 Flight Trials

To investigate the effects of parameter variations as regards control inputs on the behaviour of the aircraft during stall and the entry phase to a spin, flight trials were performed with a Cessna F 172N. Controllability of the aircraft at stall and the behaviour of the aircraft when performing manoeuvres for attempted spin entry were analysed. Based on the FAR spin resistance requirement and on manoeuvres proposed by EASA, parameters related to control inputs were also varied. 199 manoeuvres assorted to eight types of manoeuvres were analysed, containing:

- o Stalls with
  - Engine idle (representing descending flight)
  - Engine setting for level flight

- Engine at maximum power (representing climbing flight)
- Engine power raise from idle to 100% in less than 3 s (representing go around or the pilot's attempt to recover from stall)

o Stalls with 1kn/s and 5kn/s deceleration
o Bank angle up to 45°, left and right
o Stick at rear stop up to 7s
o Stalls with slip and skid
o Rudder at stall (dynamic rudder)
o Various aileron deflections

Results of the flight trials (a more detailed analysis can be found in the study report) show significant changes in the stall behaviour of aircraft when varying the entry parameters, e.g. a rate of speed reduction of 5 kn/s instead of 1 kn/s or a power setting of 100% instead of 75%. Any non-compliance with the conditions that had been defined for every type of manoeuvre has often been combined with an increased height loss.

The Cessna 172 does not comply with the spin resistance requirements. The analysis of accident statistics revealed that it has a lower rate of fatal mishaps per year per 1000 aircraft as well as a lower proportion of stall/spin-related accidents among all fatal accidents compared to the Cirrus SR20/22 which complies to the requirements. A power setting of 100% and a rate of speed reduction of 5 kn/s, being the emphasis of the non-compliances of the manoeuvre conditions in the flight tests, are not contained in this requirement. It can be concluded that pilots in operational accident situations presumably applied control inputs not contained in the requirement manoeuvres that have to be performed for certification.

4 Conclusion

Technical measures to influence the behaviour of aircraft at stall have been developed by NASA. Aircraft equipped with modified outboard leading edges have an enhanced lateral



Fig. 5 Sensor installation for flight testing on the Cessna F 172N

controllability at stall and are less likely to enter a spin than conventional aircraft. To enable a wide spread of this technology within General Aviation, a requirement conforming to the abilities of this technical measure was formulated and implemented into FAR23. Flight tests within this project as well as an analysis of NASA flight tests show that the behaviour of the aircraft is highly dependent on the flight state, control inputs and parameters before and at stall (cp. [11] additionally). Aircraft with modified outboard leading edges can enter a spin if a high rate of speed reduction or a high power setting is applied. This is despite the fact that these aircraft meet the FAA spin resistance requirements, as it is not required to demonstrate these manoeuvres.

Pilot behaviour
Speed rates in critical operational situations are often high, as a pilot under stress may apply control inputs rather abruptly. The awareness of the pilot for the stalled flight condition is a very central aspect (cp. [12]). To interrupt the sequence of events leading to a fatal accident, it is necessary to create this awareness in the pilot's mind. A mere extension of the time span until the aircraft enters a flight condition with decreased controllability does not create this awareness.

Operational experience with existing spin resistant designs
Analysing the operational experience with spin resistant aircraft, the conclusion can be drawn that the aim of preventing stall/spin-related accidents is not achievable with the existing spin resistant aircraft designs because decisive operational accident situations are not covered by the current FAR spin resistance requirement. In many of the accidents investigated, the pilots had presumably applied a very high power setting, whereas demonstration of manoeuvres with pro-spin control inputs at maximum available power is not required. It is also noted that, for aircraft with non-charged piston engines, testing for handling with power on is normally conducted at a safe altitude where the power is reduced relative to that at sea level, and therefore may not be representative. About 26% of the fatal accidents with aircraft certified as spin resistant according to the current FAR23.221(a)(2) are stall/spin-related.

This leads to an *evaluation of the current FAR spin resistance regulation*:
Pilots killed in stall/spin-related accidents with spin resistant aircraft had flown manoeuvres that are not covered by FAR23.221(2). Considering the manoeuvres in the regulation to be aiming at representing accident-prone operational situations, it must be ascertained that the representatives of operational situations in FAR23.221(2) are not complete. The set of manoeuvres must be extended.

Possible measures
In considering the aim of reducing the amount of stall/spin-related accidents from a more general point of view, a range of possible measures should be regarded and combined. This includes measures to enhance the stall characteristics of aircraft, but also improved stall warning systems, envelope protection measures and concepts for pilot training.

Proposed requirement
Concentrating on the aspect 'stall characteristics of aircraft', it can be concluded that a requirement must contain essential additional manoeuvres and parameters. A proposal for a requirement code for implementation into CS 23 has been worked out, additionally containing a higher rate of speed reduction (5 kn/s), maximum available power, banked turns with pro-spin control input, a maximum allowable height loss and a requirement for clearer information to the pilot. A flight test guide for the code has been worked out containing additional indications to the flight tests.

References

[1] EASA, "Specifications attached to the Invitation to Tender - EASA.2008.OP.03 - Safety Aspects of Light Aircraft Spin Resistance Concept", 05/19/2008.

[2] FAA, “FAR Final Rule”, amendment No. 23-42, 1991.
[3] FAA, “FAR Final Rule”, amendment No. 23-50, 1996.
[4] Stough, DiCarlo, “Spin Resistance Development for Small Airplanes – A Retrospective”, SAE Technical Paper 2000-01-1691, 2000.
[5] Chambers, Stough, “Summary of NASA stall/spin research for general aviation configurations”, AIAA paper 86-2597, 1986.
[6] DiCarlo, Glover, Stewart, Stough, “Discontinuous Wing Leading Edge to Enhance Spin Resistance”, Journal of Aircraft Vol. 22 No. 4, 1985.
[7] Langevin, G., “Spin Resistance”, NASA Concept to Reality, 2003.
[8] Manuel, DiCarlo, Stough, Brown, Stuever, “Investigations of Modifications to Improve the Spin-Resistance of a High-Wing, Single Engine, Light Airplane”, SAE Paper 891039, 1989.
[9] National Transportation Safety Board, Aviation Accident Database, Reports.
[10] Wikipedia, “Diagram illustrating the legs of a typical left-hand traffic pattern as flown at an airport”, 2005.
[11] Stinton, “Flying Qualities and Flight Testing of the Airplane”, Loughborough University of Technology, UK.
[12] Stowell, R., “The Light Airplane Pilot’s Guide to Stall / Spin Awareness”, Rich Stowell Consulting, Ventura CA 2007.

3rd CEAS Air&Space Conference
21st AIDAA Congress

A procedure for the cost modeling assessment of a launcher

Luca Visconti[(1)], Paolo Gaudenzi[(2)], Marco Bocciarelli[(3)]

[(1)]Research Engineer, "ESA-ESRIN",
Email: luca.visconti@esa.int
[(2)]Full Professor, University of Rome "La Sapienza",
Email: paolo.gaudenzi@uniroma1.it
[(3)]VERTA Flight Programme Manager, "ESA-ESRIN",
Email: marco.bocciarelli@esa.int

Keywords: *Cost engineering, launcher cost assessment*

Abstract

The following article deals with the methodology used to assess the production cost for a small launcher by means of a parametric cost modeling technique.
The methodology aims at providing the right approach in order to set the procedure in a proper way and to interface the technical available information with software capable to implement the methodology. The software performs a parametric cost analysis by means of appropriate cost estimating relationships and data provided by its own database.

1 Parametric Cost Analysis - Generality

Parametric cost analysis is one that uses Cost Estimating Relationships (CERs) and associated mathematical algorithms (or logic) to establish cost estimates.

The CER is the distinguishing feature of parametric estimating. A CER is a mathematical expression, which describes how the values of, or changes in, a "dependent" variable are partially determined, or "driven," by the values of, or changes in, one or more "independent" variables (cost drivers). The CER defines the relationship between the dependent and independent variables, and describes how it behaves. Since a parametric estimating method relies on the value of one or more input variables, or parameters, to estimate the value of another variable, a CER is actually the quintessential parametric estimating technique.

The dependent variable of a CER is cost or a closely related resource measure such as staff hours. The independent variable or variables are typically technical parameters that have been demonstrated to correlate with cost. In a cost-to-cost relationship, the independent variables are also costs.

In a non cost-to-cost relationship, the CER uses a characteristic of an item to predict its cost. Examples are CERs that estimate an item's manufacturing costs based on its weight (independent variable).

A CER is a valuable estimating tool and can be used at any time in the estimating process. For example, CERs may be used in the program concept or validation phase to estimate costs when there is insufficient system definition for more detailed approaches, such as the classical "grass roots" or "bottoms-up" methods. CERs can also be used in a later phase of a program as primary estimates or as crosschecks of non-parametric estimates. CERs may also form the primary Basis of Estimate for

proposals submitted to the Government or higher-tier contractors. They are also used extensively by government agencies to develop Independent Cost Estimates for major elements of future programs. In practice, a reasonable estimate of a program can sometimes be made parametrically with as little as one simple CER consisting of a single independent variable. As the program definition is fleshed out, additional parameters become available for use in cost estimation. Parametric cost models comprised of several CERs can then give estimates at lower levels of definition. Before developing complex parametric models, analysts typically create simple CERs which demonstrate the utility and validity of the basic parametric modeling approach to company and Government representatives [1].

The proper development and application of CERs depends on the collection and preparation of data on historical programs and on applying appropriate mathematical and statistical techniques. Figure1 is a schematic of the CER development process.

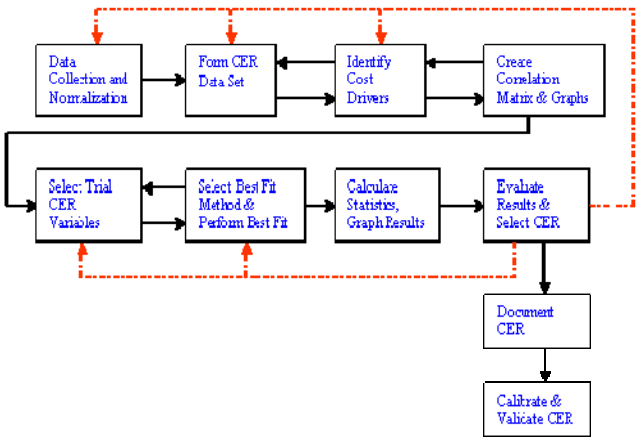


Figure 1 CER Development process [1].

1.1 Data Collection and Normalization

Parametric techniques require the collection of historical cost data and the associated non-cost information and factors that describe and strongly influence those costs. Data should be collected and maintained in a manner that provides a complete audit trail with expenditure dates so that costs can be adjusted for inflation. Non-recurring and recurring costs should be separately identified. While there are many formats for collecting data, one commonly used by industry is the Work Breakdown Structure (WBS), which provides for the uniform definition and collection of cost and certain technical information [1].

The company's management information system is the collection point for cost data which in most instances contains the general ledger and other accounting data.

Technical non-cost data describe the physical, performance, and engineering characteristics of a system, sub-system or individual item.

Also programmatic information (e.g. tools and skills of the project team, working environment, compression of schedule etc…) should be collected.

Cost reports, historical database, technical database, contracts and cost proposals can be identified as the main sources of useful data.

Cost data must be adjusted to eliminate any bias or "unevenness" which other factors may cause in it. This is called normalization and is intended to make the data set homogeneous, or consistent. The analyst needs to examine every data set to ensure it is free from the effects of:

- the changing value of the dollar over time
- cost improvement as the organization improves its efficiency
- various production quantities and rates during the period from which the data were collected

Non-recurring and recurring costs are also segregated as part of the normalization process. Non-recurring costs include all the efforts required to develop and qualify a given item, such as requirements definition, design, analysis, development, and qualification and verification.

Recurring costs cover all the efforts required to produce end-item hardware, including manufacturing and test, engineering support for production, and spare units or parts.

Several more complex adjustments must be taken into account:

- Adjustment for consistent scope (necessary to correct for differences in program or product scope between the historical data and the estimate being made)
- Adjustment for anomalies (necessary when it is not reasonable to expect the under estimated project to contain extra-costs given by these unusual events)

1.2 CER Development

CERs are usually expressed in the form of linear or curvilinear statistical regression equations that predict cost as a function of one or more "cost drivers" ("independent" variables). At the top of a cost-estimating task lies the work-breakdown structure (WBS), a list of everything that has to be paid for in order to bring a system to its full operational capability. A WBS includes "high level" categories such as research, development, and testing, production and operations, maintenance, and support. At "lower" levels of the WBS, costs of software modules, electronic boxes and other components are accounted for. A CER can be established for any WBS element at any level, as long as data are available to support its derivation.

The more common kind of CER expresses the cost of an item (system, subsystem, component, or activity) in terms of some intrinsic characteristics of the item. For example, CERs express the cost (dependent variable) of an item in terms of independent variables such as its weight, the number engineering drawings required for its design, some measure of its heritage from previously produced items of the same kind, or the difference between its current technology-readiness level (TRL) and the TRL required for its final production and deployment [2].

CERs can be divided in 4 main algebraic categories showed in the following table:

Category	Cost	Cost Driver	Coefficients	Basic Algebraic Form
Factor	y	x	a	$y = ax$
Linear	y	x	a,b	$y = a + bx$
Power	y	x	a,b	$y = ax^b$
Triad	y	x	a,b,c	$y = a + bx^c$

Table 1 CER algebraic Forms and Categories [1].

Obviously in many situations the cost driver x can be a vector.

In such cases, the resulting *multivariate* CER is typically a combination of the four basic CER forms listed in the previous Table.

The *multiple linear* CER form is a combination of the univariate (i.e., one cost driver) factor and linear CERs. Its generic algebraic form is:

$$y = a + bx + cw + dz + \ldots, \quad \textbf{(1)}$$

where x, w, and z are cost-driver values and a, b, c, and d are coefficients.

The *multiple power* CER form is a combination of the univariate factor and power CERs. Its algebraic form is:

$$y = ax^b w^c z^d \ldots, \quad \textbf{(2)}$$

where again x, w, and z are cost-driver values and a, b, c, and d are coefficients.

The *multiple triad* CER form is:

$$y = a + bx^c + dw^e + fz^g + \ldots, \quad \textbf{(3)}$$

where again x, w, and z are cost-driver values and a, b, c, d, e, f, and g are coefficients [1].

In order to apply good sense in the use of CERs, it is needed to know their strengths:

1. the CERs based on historical data can claim objectivity
2. CERs reveal sensitivity of cost to physical and performance attributes
3. Use of valid CERs can reduce proposal preparation, evaluation, negotiation costs and cycle time.

4. Most CERs can be used with a small amount of top level information about the product or service being estimated
5. the estimation process by using CERs is a quick and easy process

and their weaknesses:

1. If detailed information are available CERs may be too simple to be used to estimate costs; in this case a detailed estimate could be more reliable than one based on a CER.
2. Historical data are difficult to obtain and properly normalize
3. CERs cannot estimate costs of Beyond State of the art technology

1.3 PRICE System

The standard tool used in space applications to perform parametric cost analysis is PRICE (Parametric Review of Information for Costing and Evaluation).

PRICE allows to perform risk and planning analysis for hardware and software product. PRICE algorithm uses a group of CERs with the scope to get a more accuracy in the result [5].

The most important input is the Estimating Breakdown Structure (EBS) that is a graphical tree structure depicting the hardware or blocks of software code hierarchy of the system being estimated [3].

The EBS for a hardware product is composed by assembly, structural, electronic, integration & test, furnished and purchase elements, while for software is composed by a computer software configuration item, computers software components, purchased or furnished elements.

Each of these elements needs to a number of input variables (cost drivers).

The main input variables for the hardware cost estimation are the mass, the volume, the manufacturing complexity, the operating environment, the difficulty in the integration of the single component, the skill and the experience of the team, the percentage of newel in the design, programmatic information etc….

An example of a hardware EBS is showed in the following figure. The system 1 and 2 represent the assembly elements composed by electro-mechanical elements, structural elements, purchased elements etc…

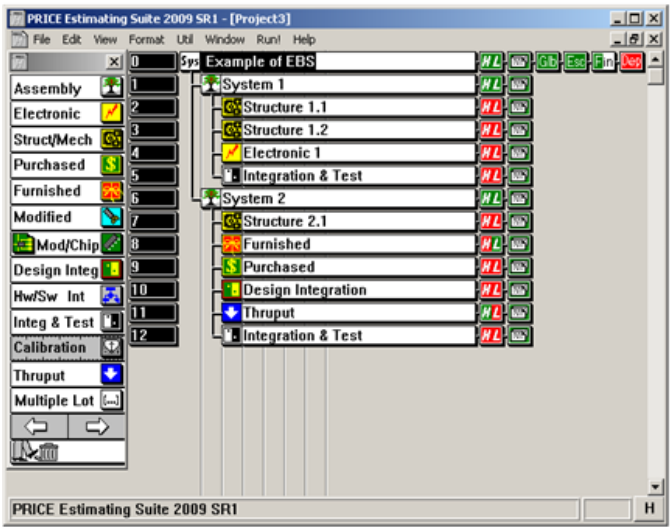


Figure 2 Example of Estimating Breakdown Structure for a hardware system (EBS).

An EBS for the software evaluation is mainly composed by Computer Software Configuration Item (CSCI) representing first level software module in the system and by Computer Software Component (CSC) representing the lower level always subordinated to the developed CSCIs. CSC must have one or more subordinated programming language elements.

In order to evaluate the cost of the software the main input parameters (cost drivers) are the number of the line of code, the kind of programming language, the operating environment, the skill and the experience of the team [3].

The figure 3 represents the EBS of a software product. Every block CSCI is composed by CSC elements along with the own programming language:

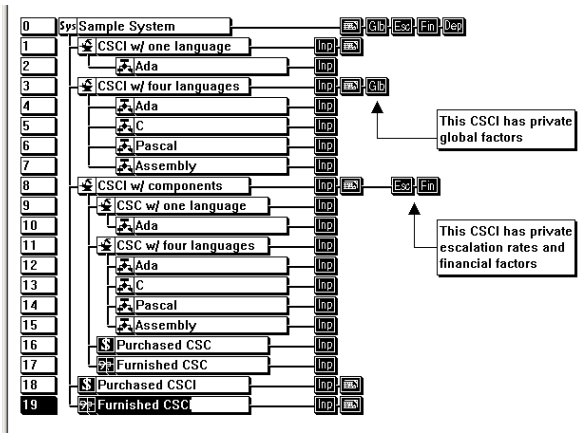


Figure 3 Example of Estimating Breakdown Structure for a software system (EBS) [3].

2 Model Development Methodology

The present chapter deals with the methodology developed to assess a cost for a small launcher, by means of a parametric cost modeling software. The methodology aims at providing the right approach in order to set the software in a proper way and to interface the technical available information with the tool.

Starting from functional, technical and programmatic requirements it is possible to classify the launcher and its subsystems in terms of operating environment, manufacturing complexity and difficulty in the integration at lower and higher level.

The main important step is the acquisition of the knowledge about the product to estimate [6]. For a space-craft, the mission objectives definition is the starting point requested in order to define the mission requirements.

The sub-system and afterwards the Work Breakdown Structure (WBS) and the product tree can be defined once the requirements are frozen.

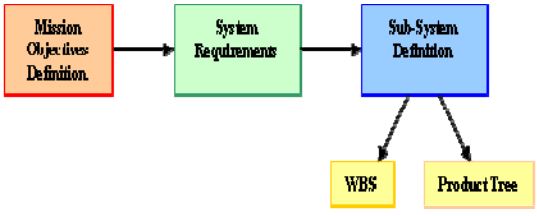


Figure 4 Model Development Process.

For each element composing the product tree it is possible to find the main cost drivers and to identify the more suitable estimating CER.

In order to set-up each element with the proper technical and programmatic parameters from which the cost depends on, it is necessary to collect data.

Concerning the hardware estimation, the mass budget is a useful tool in order to verify and to detail the Estimating Breakdown Structure (EBS) and to assign the mass to every element. In addition to the mass others technical parameters can be found and set with the right numerical value.

For this purpose can come to the aid the filling of the following data-sheet, representative of the product being estimated in terms of technical and programmatic characteristics:

Equipment	Structural element 1	Structural element 2	Purchased element	Electro-mechanical element 1	Electro-mechanical element 2	Integration Element
Total Quantity						
Quantity at next higher assembly						X
Scope						
Operating environment						X
Duration of cold life						X
Duration of operating life						X
Number of cycle						X
Total Weight [Kg]						X
Type of electronic components	X	X				X
Electronic Components Density	X	X				X
Volume [dm^3]						X
Material				X	X	X
Precision				X	X	X
Manufacturing technology						X
Number of parts				X	X	X
Difficulty of integration						
Start of production or Integration						
End of Production or Integration						

Table 2 Data Sheet format for the technical definition of the product being estimated.

Referring to the table above, it is mandatory to precise that the total quantity means the size of the batch of the product to produce, the quantity at next higher level means the number of units required to be integrated at the next higher assembly level and the difficulty of integration defines the level of effort required for the integration activity:

- Low level (L) means a simple interface and a routine operation
- Medium level (M) means new and moderately difficult interface (it requires Precision alignment, special tools and fixtures.
- High level (H) means complex operation and difficult interface (it requires precision alignment using special jigs and tools possibly monolithic).

The filling of the previous table is a useful instrument to set a cost estimating tool in a proper way.

Concerning the software estimation, the adopted approach is similar to the previous one performed for the hardware evaluation.

In general, one of the first steps in any estimate is to understand and define the system to be assessed. However software is intangible, invisible, and intractable. It follows the enormous difficulty to assess a cost. The software WBS is an excellent tool for visualizing the software product.

The software WBS should not be complex or highly detailed. A simple product tree line drawing is often adequate for initial software estimates. The hardware WBS can be a useful tool in developing the initial WBS for software.
Each of the major software functional units can be modeled as a Computer Software Configuration Item (CSCI) associated with each hardware Line Replaceable Unit (LRU). Lower level WBS elements can be modeled as a component.

As the program evolves, the initial or draft WBS should include all software associated (developed, furnished, or purchased).
The software estimation process includes:

- estimating the size of the software product to be produced
- determining which functions the software product must perform
- estimating the effort required
- developing preliminary project schedules
- estimating overall cost of the project

A scheme of the main blocks-code can be useful to identify the EBS and to evaluate the software size expressed in Lines of Code (LOC) along with other important parameters:

Input Parameter	Description
Requirement Start Date	Represents the beginning of software requirement analysis and definition.
Design Start Date	Represents the beginning of the detailed software design.
Operating Environment	Represents the variable that describes the customer's requirements stemming from the planned operating environment.
External Integration	Represents the level of difficulty of integrating the CSCIs to the system level.
Internal Integration	Represents the level of difficulty of integrating and testing to the CSC level.
Productivity Factor	This parameter is representative of the skill, experience, productivity and efficiency level of the team involved in the software development.
Language	Represents the programming language.
Size	Represents the number of units (source lines of code) needed to create a software element and to effect a software estimate.
Application	Represents a parameter that summarizes the functional mix of instructions (math, real time, online communications and so on).

Table 3 Main Software Cost Drivers.

Once a broad EBS (for hardware or software) is built and every cost drivers set up, it would be necessary to test the model or to revise it by means extensive analysis of design and production hypotheses with technical chief.

Where the information has not been sufficient, a risk analysis has been performed so that each possible case (defined in a suitable range) could be covered.

For instance, the reference mass value of each launcher component is referred to the Best

Engineering Estimate (BEE); a percentage number (D) representative of the dispersion is also associated to the BEE and so it is possible to consider a triangular distribution where the BEE is the most likely value and the two extreme values are given by:

$$Ext = BEE\left(1 \pm D/100\right) \quad \textbf{(4)}$$

Finally, the Monte Carlo method is used to compute the possible cost associated to each component.

In order to perform a comparison with the supplier's offer the results need to be harmonized with

Eventually it can be useful to use a learning curve in order to compute the recurring cost starting from an average cost of the batch or from the cost of the first item.

3 Case Study

The goal of the work is to estimate a cost for a launcher designed to inject small payloads, from 300 up to 2,000 kg for scientific and earth observation missions into polar and low earth orbits.

The reference mission is a polar orbit bringing a spacecraft of 1,500 kilograms to an altitude of 700 km [4].

The mission is defined and the launcher requirements can be evaluated:

- Geometry and mass of the launcher depend on structural requirements, on the maximum volume able to accommodate the satellite inside the fairing and on its weight and on the amount of propellant necessary to inject the payload into the final orbit with the appropriate accuracy
- The launcher will have to withstand the mechanical stress given by the working environment (acceleration, vibration, etc…)
- The launcher propulsion is entrusted to three solid rocket motors (1st, 2nd and 3rd stage) and to a liquid propulsion system (4th stage) that guarantees at least five ignitions, giving the launcher the capacity to perform a broad range of missions.

Starting from the requirements it has been possible to identify the main components of the launcher and the relative subsystems: mass, length and the average diameter; technical parameters of the solid rocket motor and of the liquid propulsion system; requirements of the batteries and of the avionic system and so on.

1st STAGE	
Length	11.2 m
External diameter	3 m
Weight without propellant	7416 kg
Propellant weight	88380 kg
Thrust	3000 kN
Specific impulse	280 s
Time of combustion	106.8 s
2nd STAGE	
Length	8.39 m
External diameter	1.904 m
Weight without propellant	1877 kg
Propellant weight	23906 kg
Thrust	1196 kN
Specific impulse	289 s
Time of combustion	71.7 s
3rd STAGE	
Length	4.12 m
External diameter	1.904 m
Weight without propellant	833 kg
Propellant weight	10155 kg
Thrust	280 kN
Specific impulse	295 s
Time of combustion	117 s
4th STAGE	
Length	2.04 m
External diameter	2.18 m
Weight without propellant	418 kg
UDMH weight	183 kg
NTO weight	367 kg
Helium weight	3.5 kg
Thrust	2.45 kN
Specific impulse	315.5 s

Table 4 Launcher technical characteristics [4].

On the basis of the previous consideration a product tree and the associated activities related to each item have been defined.

The product tree is fundamental to get an EBS for the PRICE H model. Every assembly elements can be expanded reaching a very fine level of detail.

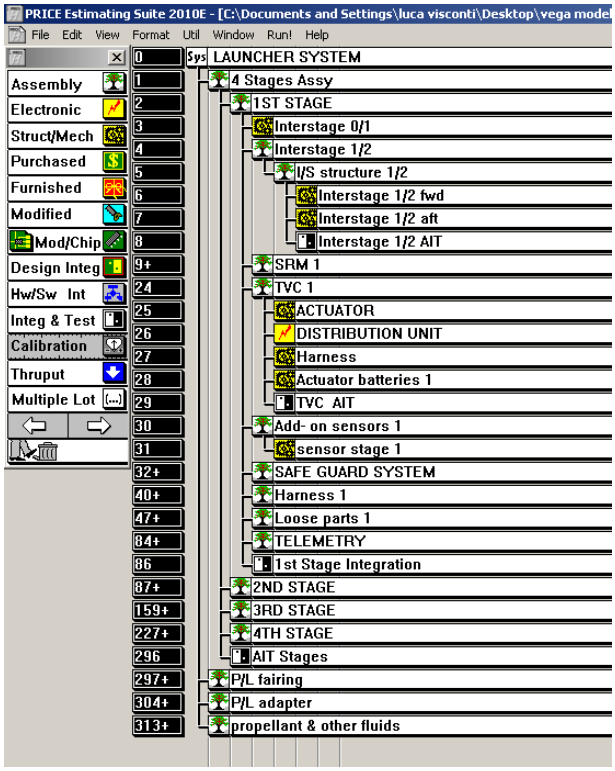


Figure 5 Launcher EBS composed by a total of 319 elements.

The recurring-cost model developed is based on the following hypotheses:

- Production batch size: 10
- Cadence: 4 launchers / year
- Recurring production costs only
- Transportation and operations are not included (but they are known)

The main cost drivers have been identified [3]:

- PLTFM (space unmanned operating environment)
- CPLXS (manufacturing complexity for the structure components)
- MCPLXE (manufacturing complexity for the electronic components)
- MASS
- INTEGS (difficulty in the integration of the assembly)

The pie chart shows the percent of the cost for each item with respect to the overall launcher:

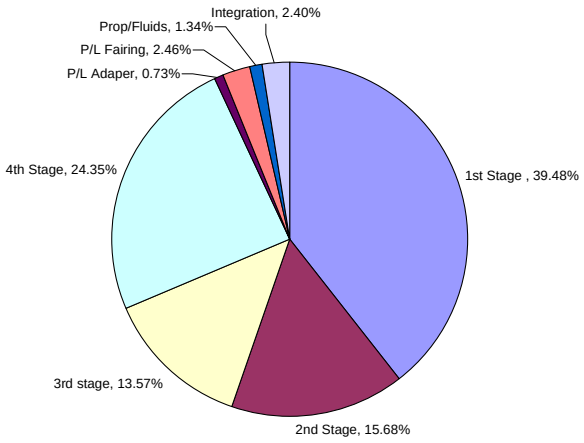


Figure 6 Percentage results with respect to the overall launcher average cost.

The model does not estimate a cost for the operation on the launch site (payload integration), transportation and the post launch analysis activities.

A learning coefficient equal 0.85 (by NASA website learning curve calculator) has been used to get in output the cost of the first item from the average cost of the batch.

The model is set in order to consider also a triangular distribution of the masses of every component. The reference value is the BEE and the extreme values are computed by the mass contingency reported in the launcher mass budget.

So the masses are random variables of the cost model and are randomly sampled to produce the model's output by means of a Monte Carlo simulation. In this way the output is also a statistical distribution.

Concerning the Flight Program Software (FPS) a similar approach can be used. It is necessary to precise that a parametric cost model has sense only for estimating the

Invariant Flight Program Software (IFPS), the one that does not change from mission to mission. The IFPS is composed only by data structure and includes three main blocks [4].
Block1 is composed by:

- Launch Vehicle Data (representative of the hardware of the launcher)
- Guidance, Navigation and Control data
- Guidance, Navigation and Control algorithms (to guarantee the predefined orbital accuracy)
- Flight Management and Sequential (which executes the GNC instructions in the right sequence)

Block2 is the layer, which comprises the following functions:

- Scheduler (to manage the activation of cyclic and sporadic functions)
- Frame manager (to control frame sequencing)
- Input/Output manager (to provide interface services necessary during the mission to manage and configure the equipment)

Block3 is the layer, which services are mainly:

- Time management
- I/O management

The software EBS is showed in the following figure inspired to the description above:

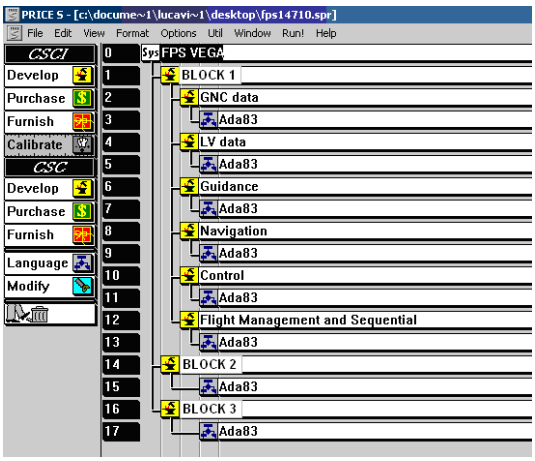


Figure 7 Launcher Flight Program Software EBS.

In this case the reported output is the schedule of the software development showed in the following Gantt diagram:

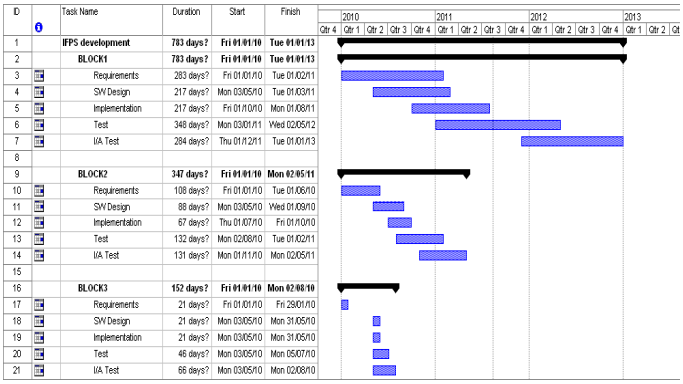


Figure 8 IFPS schedule development.

The previous Gantt Diagram guesses as a starting software development date the first of January 2010 and shows that the overall IFPS development lasts three years

4 Conclusion

The main reasons for creating a parametric cost model are:

- Generating an independent cost estimate
- Obtaining terms of comparison for industry Cost documentation and quoted price
- Identification of critical cost items
- Performing a sensitivity analysis (e.g. cadence, batch size)

In other words a parametric cost model can be considered a powerful instrument for the negotiation during every phase of a project.

The limits are essentially due to the estimation of the non- recurring activities like the ones changing from mission to mission.

References

[1] ISPA (International Society of Parametric Analysts), “Parametric Estimating Handbook”, fourth edition April 2008..
[2] Nguyen, P., N. Lozzi, W. Galang, et al., “Unmanned Space Vehicle Cost Model, Seventh Edition”, U.S. Air Force Space and Missike Systems Center (SMC/FMC), 2430 E. El Segundo Blvd., Suite 2010, Los Angeles AFB, CA 90245-4687, August 1994, xvi+452 pages.
[3] PRICE Estimating System user manual.
[4] VEGA User Manual Issue 3 Revision 0, March 2006, Approved and Issued by ArianeSpace.
[5] G. Cifani, P. Gaudenzi, C. Portelli, G. Belvedere “Analisi Tecnico Economica di un rover lunare” A.A. 2009-2010.
[6] G. Belvedere, E. Lamboglia, C. Portelli, P. Gaudenzi, V. Tota, “The Effect of the amount and level of detail of technical information upon the cost estimate process in a concurrent engineering environment”, SECESA 2010.
[7] G. Filippazzo, “Complexity Based Cost Estimating Relationships for Space Systems”, 2004 IEEE Aerospace Conference Proceedings.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Conceptual Design Of An Aeroelastic Fly-By-Wire Aircraft With Manned Flight Simulation In The Optimization Loop

F. Bocola, P. Masarati and G. Quaranta
Politecnico di Milano, Dipartimento di Ingegneria Aerospaziale
Via La Masa 34, 20156 Milano Italy
bocola@aero.polimi.it, masarati@aero.polimi.it, quaranta@aero.polimi.it

Keywords: aeroelastic flight simulator, active flight control systems, handling qualities evaluation

Abstract

Objective of this work is the creation of an interdisciplinary simulation environment for the optimization of new generation aircraft with Fly-by-Wire (FWB) control systems. Such tool must be able to merge flight dynamics and structural dynamics, allowing handling qualities evaluations through simulations with pilot-in-the-loop and the inclusion of such results in the conceptual design loop.
This engineering approach should reduce time to market and development costs of new aircraft and exploit the potential increase of performances that can be delivered by Fly-By-Wire Flight Control System (FCS), accounting also for aeroservoelastic phenomena from early stages of design.

1. Introduction

The classic approach to aircraft Preliminary Design develops throughout a sequential and iterative method, where every output from each single design department circles in the loop until all design constraints are satisfied (Fig. 1).

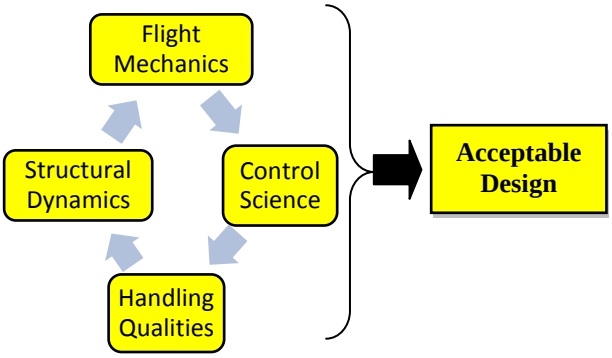


Fig. 1: Classic design approach

This approach ends in a low number of iterative cycles as long as the considered airplane has high structural stiffness, conventional flight control systems architecture and low level of automation for flight path control (basic autopilot).

The increasing application of Active Feedback technology and Fly-by-Wire control systems for envelope protection and path control has indeed helped aviation industry to easily correct undesired dynamic behaviour (in terms of certification requirements), also introducing a

potential chance to control aeroelastic phenomena.

However, the implementation of this technology has introduced new aeroservoelastic problems that require a reassessment of the traditional design approach [1].

Additionally, the need to contain within an acceptable level the development costs of new aircraft, led the research community and industries to explore innovative integrated approaches among the subjects involved in the design, typically:

a) Flight Mechanics
b) Structural Dynamics and Aeroelasticity
c) Control Science
d) Certification requirements and Handling Qualities

The long term target in this research field is to consolidate optimization procedures in the conceptual design phase, in order to integrate all the collected parameters about mass, structural geometry and control laws, thus extending the potential applications of conventional airplane configurations by exploiting their mutual interaction.

In this new integrated design process (Fig. 2), elements are captured in a concurrent design environment, thus implementing structural elasticity from the very beginning of FBW control law design process [2, 3, 13].

An essential part of this process is the development of an Aeroelastic Flight Simulator for pilot-in-the-loop Flight Tests, for early validation of control law design and handling qualities evaluation (Cooper-Harper rating, absence of PIO-Pilot Induced Oscillation), within the ability of the simulator to reproduce reliable accelerations to the pilot.

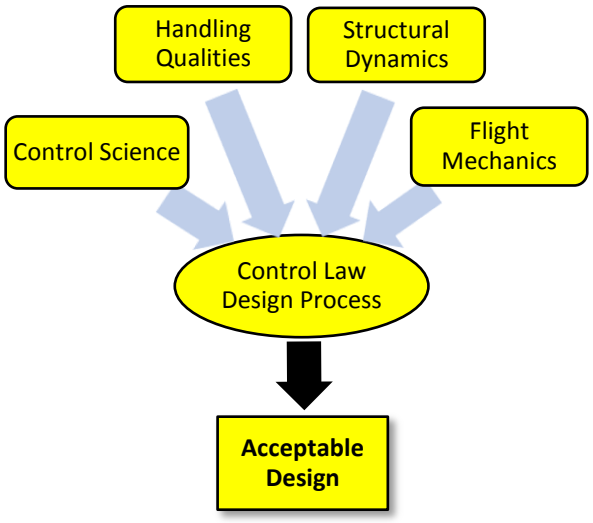


Fig. 2: Integrated design approach

The selected aircraft configuration is a Three-Surface Regional Jet with forward swept wing and T-tail. The potential of this architecture has been exploited for drag minimization of the main wing, as investigated at Politecnico di Milano by means of an aeroelastically scaled wind tunnel model denominated XDIA [5]. An upscaled variant of the XDIA has been used in this research.

Scientific debate on three surface aircraft is still ongoing, having many Authors as Kendall [6, 7, 8] who claim for three surface superiority due to the theoretical ability to trim the aircraft with the minimum induced drag possible, wherever the center of gravity is, provided both the canard and horizontal tail surfaces are independently operated.
Although a synergic cooperation between structure and aerodynamics towards weight saving and drag reduction is expected, drawbacks must be considered in terms of additional weight, complexity and interference drag [9].

2. Work plan

This research is articulated in the following steps:

a) Realization of Flight Simulator

b) Aerodynamic characterization of a three-surfaces aircraft (rigid and flexible) and its implementation in the simulator for preliminary pilot-in-the-loop handling qualities evaluation

c) Implementation of active flight control system (FBW)

d) Handling qualities evaluation with pilot-in-the-loop and FBW control laws tuning

3. Realization of flight simulator

Beside traditional use for crew training, flight simulation is a powerful tool for scientific research in flight mechanics to achieve good handling qualities of new aircraft design before first flight.

They can either be static (with or without virtual reality visualization) or full motion (6 DOF with visual), as reported in Fig. 3a.



Fig. 3a: Full Flight Simulator (6 Degrees of Freedom)

Full flight simulator has the great advantage to allow evaluation of Aircraft-Pilot Coupling (APC) due to cockpit accelerations induced by structural elasticity [1], if the flight simulator has enough bandwidth to reproduce aeroelastic frequencies with an adequate fidelity.
When this is the case, the pilot is provided with strap-down accelerometers (Fig. 3b).



Fig. 3b: Aircraft/Pilot coupling evaluation in a helicopter full flight simulator

However, at Conceptual Design stage a Static flight simulator with visual is considered a good trade-off between realization costs and ability to produce useful information for preliminary tuning of FCS Control Laws [3]: this is the most adopted solution in the Aviation Industries [3].

In this case, because of the absence of simulator motion, Flying Qualities Evaluation must be mainly based on the ability to satisfy Instrument Flying tasks: for this reason, natural scale visualization of Primary Flight Display (Flight Instruments) is needed.

The chosen Flight Simulator architecture is illustrated in Appendix 1, where integration of two different codes has been realized in a way that each one accounts for:

- **SIMULINK** [14]: Flight Control inputs (Joystick), Aircraft Characterization (through Aerodynamic, Aeroservoelastic and Control Derivatives), Flight Dynamic computation, Navigation, implementation of Flight Control Laws (Fly-By-Wire), weather environment (wind, turbulence).

- **FLIGHT GEAR** [15]: 3D animation Scenarios and Cockpit instruments.

The approach of using SIMULINK Block Diagrams for the Simulator engine, allows faster modification capability during the design process, where aircraft characterization is totally

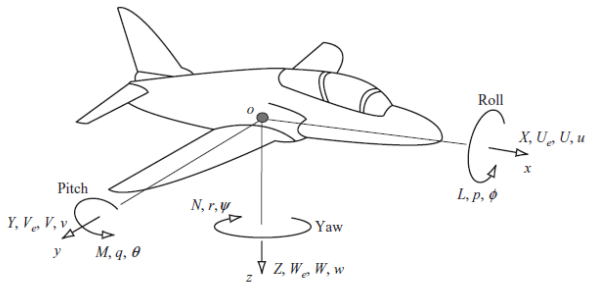
contained in the Aerodynamic Forces/Moments Block.

Other features, such as Navigation on spherical earth and Attitude Parameterization with Quaternions, allow this Simulator to be suitable not only for short Handling Qualities evaluation flights, but also for long-haul and sub-orbital flight assessment.

Involved Reference Systems are:

- BODY Frame
- NORTH EAST DOWN (NED) Frame (local horizon)
- EARTH CENTERED EARTH FIXED (ECEF) Frame (navigation)

Assuming a spherical Earth, they are illustrated in Fig. 4 and Appendix 2.



Motion variables notation.

Summary of motion variables

	Trimmed equilibrium			*Perturbed*		
Aircraft axis	*ox*	*oy*	*oz*	*ox*	*oy*	*oz*
Force	0	0	0	*X*	*Y*	*Z*
Moment	0	0	0	*L*	*M*	*N*
Linear velocity	U_e	V_e	W_e	*U*	*V*	*W*
Angular velocity	0	0	0	*p*	*q*	*r*
Attitude	0	θ_e	0	ϕ	θ	ψ

Fig. 4: BODY Reference System [4]

To improve simulation environment, Flight Dynamics Module is based on the following non-linear rigid-body differential equations of motion (Eq. 1 and 2), expressed in the Body Reference System:

Translations

$$\begin{cases} m\left(\dot{U} - V\,r + W\,q\right) = X \\ m\left(\dot{V} + U\,r - W\,p\right) = Y \\ m\left(\dot{W} - U\,q + V\,p\right) = Z \end{cases} \qquad (1)$$

Rotations

$$\begin{cases} I_X\,\dot{p} - I_{XZ}\,\dot{r} + q\,r\,(I_Z - I_Y) - I_{XZ}\,p\,q = L \\ I_Y\,\dot{q} + r\,p\,(I_X - I_Z) + I_{XZ}\,(p^2 - r^2) = M \\ -I_{XZ}\,\dot{p} + I_Z\,\dot{r} + p\,q\,(I_Y - I_X) + I_{XZ}\,q\,r = N \end{cases} \qquad (2)$$

where:

m = aircraft mass
I_X, I_Z, I_{XZ} = inertia moments and product
U, V, W = linear velocity
p, q, r = angular rate
X, Y, Z = external force components
L, M, N = external moment components

In this way, the code allows a pilot to fly inside a wide flight envelope, not limited to small perturbations about a trimmed condition, even if the plane is longitudinally unstable, provided an adequate aerodynamic and control derivatives database is available.

As mentioned before, Attitude kinematics are based on Quaternion parameters to avoid "gimbal lock" (matrix singularity) during aircraft maneuvers, typical of Euler Matrix parameterization when Pitch attitude reaches ±90°.

A dedicated algorithm is implemented to continuously update local horizon (NED Ref Sys), for a correct representation of aircraft Attitude (Pitch and Roll angles) on spherical Earth.

To allow stability evaluation about trimmed flight conditions, a state-space form linearized dynamic model (Appendix 3) is obtained from the non-linear simulator using specific MATLAB commands: this procedure is used to either assess open-loop and closed-loop handling qualities for control law analysis.

Figure 5 illustrates a snapshot taken during a pilot-in-the-loop flight for Simulink/FlightGear interface tuning.



Fig. 5: FLIGHT GEAR 3D visualization environment and cockpit instruments

4. Aerodynamic modeling of three-surface aircraft

As mentioned before, the selected aircraft is a Three-Surface Regional Jet with forward swept wing and T-tail. It was selected in order to require a non trivial FBW implementation, for both flight mechanics (due to relaxed static stability caused by the canard configuration) and aeroelasticity (due to negative swept wing).

As mentioned before, the applied external forces and moments are computed using Aerodynamic and Control Derivatives.

To obtain an aeroelastic model, the adopted procedure is illustrated in fig. 6, where a panel code (AVL, Athena Vortex Lattice) [10] has been used to compute rigid-body Derivatives and a structural code (NASTRAN) [11] to define corrections for aircraft flexibility.

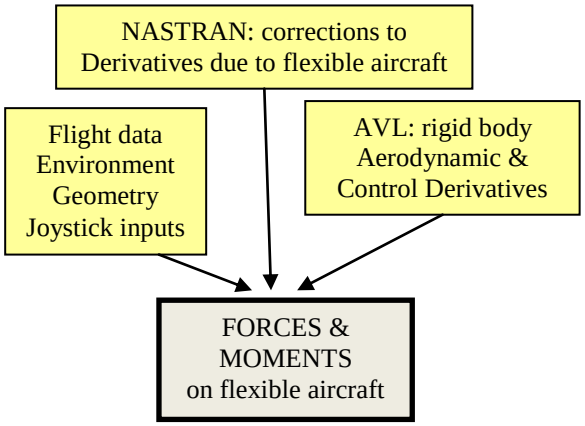


Fig. 6: Forces and Moments on flexible aircraft

A basic draft of aerodynamic mesh for the aircraft model is represented in fig. 7.

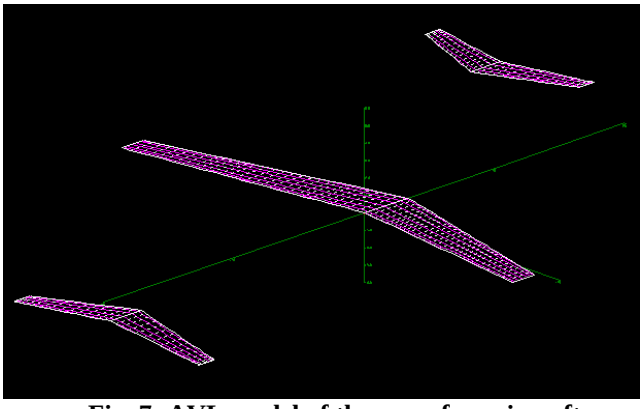


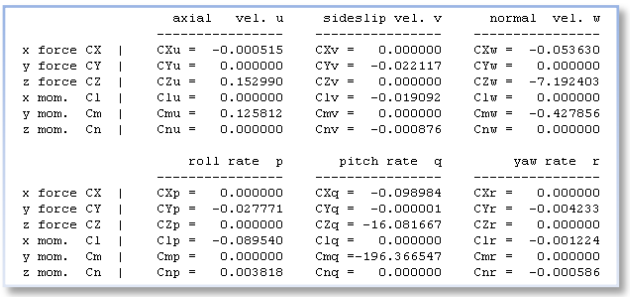
Fig. 7: AVL model of three-surface aircraft

As mentioned before, due to Flying Qualities simulation requirements, a wide flight envelope data (in terms of Altitude and Mach number) around evaluation point is required, due to initial difficulties to maintain the selected flight condition.

For this reason, a Rigid Body Aerodynamic and Control derivatives database has been created, able to provide derivative coefficients for a range of Mach numbers, as a function of:

- Angle of Attack α
- Sideslip Angle β
- Flight Control Surfaces deflection (Elevator and Canard, Ailerons and Rudder).

An example of typical AVL outputs for a selected flight condition is illustrated in fig. 8.



	axial vel. u	sideslip vel. v	normal vel. w
x force CX	CXu = -0.000515	CXv = 0.000000	CXw = -0.053630
y force CY	CYu = 0.000000	CYv = -0.022117	CYw = 0.000000
z force CZ	CZu = 0.152990	CZv = 0.000000	CZw = -7.192403
x mom. Cl	Clu = 0.000000	Clv = -0.019092	Clw = 0.000000
y mom. Cm	Cmu = 0.125812	Cmv = 0.000000	Cmw = -0.427856
z mom. Cn	Cnu = 0.000000	Cnv = -0.000876	Cnw = 0.000000

	roll rate p	pitch rate q	yaw rate r
x force CX	CXp = 0.000000	CXq = -0.098984	CXr = 0.000000
y force CY	CYp = -0.027771	CYq = -0.000001	CYr = -0.004233
z force CZ	CZp = 0.000000	CZq = -16.081667	CZr = 0.000000
x mom. Cl	Clp = -0.089540	Clq = 0.000000	Clr = -0.001224
y mom. Cm	Cmp = 0.000000	Cmq = -196.366547	Cmr = 0.000000
z mom. Cn	Cnp = 0.003818	Cnq = 0.000000	Cnr = -0.000586

Fig. 8: Typical dimensionless AVL Aerodynamic Derivatives output

These non-dimensional coefficients are then used in Eq. 3a, 3b, 4a, 4b, to compute the applied total Forces and Moments.

$$Forces = \frac{1}{2}\rho V^2 S C_i \qquad (3a)$$

$$Moments = \frac{1}{2}\rho V^2 S(\text{"c" or "b"}) C_i \qquad (3b)$$

Where C_i are the non-dimensional Aerodynamic derivatives; similar expression are used for Control derivatives [4].

Longitudinal Motion:

$$\begin{aligned} &\text{Normal Force} = Z(\alpha, M) + \Delta Z(\alpha, M, \delta_{elev}) \\ &\text{Axial Force} = X(\alpha, M) + \Delta X(\alpha, M, \delta_{elev}) \\ &\text{Pitch Moment} = M(\alpha, M) + \Delta M(\alpha, M, \delta_{elev}) \end{aligned} \qquad (4a)$$

Lateral/Directional Motion:

$$\begin{aligned} &\text{Lateral Force} = Y(\alpha, M, \beta) + \Delta Y(\alpha, M, \delta_{rud}) \\ &\text{Yaw Moment} = N(\alpha, M, \beta) + \Delta N(\alpha, M, p, \delta_{rud}) \\ &\text{Roll Moment} = L(\alpha, M, \beta) + \Delta L(\alpha, M, r, \delta_{ail}) \end{aligned} \qquad (4b)$$

5. Principles of aeroelastic modeling

The selected approach to aircraft aeroelastic modeling is based on the computation of aerodynamic and control forces and moments as a contribution of rigid body and structural elasticity components.

The governing structural dynamics equations, referred to "mean axes", are reported in Eq. (5).

Advantages of the adopted reference frame are:

- rigid and elastic modes inertially decoupled
- equations only coupled through aerodynamics

$$\begin{bmatrix}[M]_{rr} & [0] \\ [0] & [M]_{ee}\end{bmatrix}\begin{Bmatrix}\{\ddot{x}_r\} \\ \{\ddot{x}_e\}\end{Bmatrix} + \begin{bmatrix}[0] & [0] \\ [0] & [D]_{ee}\end{bmatrix}\begin{Bmatrix}\{\dot{x}_r\} \\ \{\dot{x}_e\}\end{Bmatrix} + \begin{bmatrix}[0] & [0] \\ [0] & [K]_{ee}\end{bmatrix}\begin{Bmatrix}\{x_r\} \\ \{x_e\}\end{Bmatrix} = \begin{Bmatrix}\{f_r\} \\ \{f_e\}\end{Bmatrix} \qquad (5)$$

where aerodynamic forces and moments, linearized with respect to $\{x_e\}$ when elastic modal coordinates are null ($\{x_e\} = 0$), are reported in Eq. (6):

$$\begin{Bmatrix}\{f_r\} \\ \{f_e\}\end{Bmatrix} = q\left(\begin{Bmatrix}f_{0r} \\ f_{0e}\end{Bmatrix} + \begin{bmatrix}[\partial\{f_r\}/\partial\{x_e\}] \\ [\partial\{f_e\}/\partial\{x_e\}]\end{bmatrix}\{x_e\}\right) \qquad (6)$$

where "q" is the dynamic pressure.
Applying a static aeroelastic approximation, where $\{\ddot{x}_e\} \simeq 0$ and $\{\dot{x}_e\} \simeq 0$, from Eq. (5) we obtain:

$$\begin{cases}[M]_{rr}\{\ddot{x}_r\} = \{f_r\} \\ [K]_{ee}\{x_e\} = \{f_e\}\end{cases} \qquad (7)$$

Substituting $\{f_e\}$ from the second of Eqs. (6) into the second of Eqs. (7):

$$[K]_{ee}\{x_e\} = q\left(\{f_{0e}\} + \left[\frac{\partial\{f_e\}}{\partial\{x_e\}}\right]\{x_e\}\right) \qquad (8)$$

we obtain $\{x_e\}$:

$$\{x_e\} = \left([K]_{ee} - q\left[\frac{\partial\{f_e\}}{\partial\{x_e\}}\right]\right)^{-1} q\{f_{0e}\} \qquad (9)$$

By substituting $\{f_r\}$ from the first of Eqs. (6) (where $\{x_e\}$ results from Eq. (9)) in the first of Eqs. (7), we obtain:

$$[M]_{rr}\{\ddot{x}_r\} = q\left(\underbrace{f_{0r}}_{\textbf{rigid}} + \underbrace{\left[\frac{\partial\{f_r\}}{\partial\{x_e\}}\right]\left([K]_{ee} - q\left[\frac{\partial\{f_e\}}{\partial\{x_e\}}\right]\right)^{-1} q\{f_{0e}\}}_{\textbf{elastic}}\right) \qquad (10)$$

This is a simplified rigid-body equation of motion (centrifugal terms are omitted here for simplicity, but they are retained in the analysis), where aerodynamic "rigid" and "elastic" contributions to forces and moments have been highlighted.
This approach neglects the dynamics of the elastic part, considering only the quasi-static

behavior at low frequency, assuming a clear frequency separation from the rigid body dynamics. Detailed modeling of the elastic part is obtained with a specific code MASST [12].

As mentioned before, both effects are accounted for in the "Aerodynamic Characterization" block of the Simulator (Appendix 1).

6. Active Flight Control System (FBW)

The final step of the project is the implementation of a module that simulates the Fly-By-Wire flight control system, to better exploit aircraft configurations that provide increased aerodynamic efficiency. A possible evolution consists in the control or exploitation of aeroelastic phenomena.

The selection of the FBW architecture and control laws is still in progress. The final decision will also depend on results of preliminary Pilot-in-the-Loop Handling Qualities evaluation.

7. Conclusions

The target of this work is to create an aeroelastic flight simulator to be used in the Conceptual Design process of new aircraft.

The capabilities of this tool can be exploited in the integrated design of FBW aircraft, where validation of control laws and preliminary handling qualities evaluation require to consider aeroservoelastic phenomena from the early stages of design.

Future work will address the implementation of the aeroelastic model and the active flight control system (FBW) in the flight simulator, and handling qualities evaluation with pilot-in-the-loop.

References

[1] Rommel, A., "*Flight Simulation within the frame of multidisciplinary optimization of large flexible aircraft*", RTO AVT Specialists' Meeting on Structural Aspects of Flexible Aircraft Cotrol, Ottawa Canada, 18-20 October 1999

[2] Gangsaas, D. and Hodgkinson, J. and Harden, C. and Saeed, N. and Chen, K., "*Multidisciplinary Control Law Design and Flight Test Demonstration on a Business Jet*", AIAA Guidance, Navigation and Control Conference and Exhibit 18-21 August 2008 AIAA-2008-6489

[3] Shin, J. and Harden, C. and Saeed, N. and Hartley, R. and Chen, K., "*Pitch Control Law Analysis and Flight Test Demonstration of an Executive Transport Aircraft*", AIAA Guidance, Navigation and Control Conference 10-13 August 2009 AIAA-2009-6120

[4] Cook, M, "*Flight Dynamics Principles*", Butterworth-Heinemann, ed. 2007

[5] Ricci S. Terraneo M. "*Conceptual Design of an adaptive wing for a three surface airplane"* Proceeding of the 46th AIAA Structures, Structural Dynamics and Material Conference, AIAA-2005-1959. Austin, Texas, 18-21 April 2005

[6] Kendall, E. "*The aerodynamic of three-surface airplanes*" Proceedings of the AIAA, AHS, ASEE, Aircraft Design Systems and Operations Meeting, San Diego, CA, October 31-November 2 1984. AIAA-1984-2508

[7] Kendall, E. "*The minimum induced drag, longitudinal trim and static longitudinal stability of two surface and three surface aircraft*", (AIAA-1984-2164), Seattle, Wa, Aug. 21-23, 1984

[8] Kendall, E. "*Performance trades of two surface and three surface configuration*", (AIAA-1984-2221), Hampton, Va, Jul 10-12, 1984

[9] Raymer, D.P. "*Aircraft design: a conceptual approach*" AIAA educational series, 2006

[10] See: "*http://web.mit.edu/drela/Public/web/avl*"

[11] "*MSC/NASTRAN Encyclopedia V 70.5*", 1998

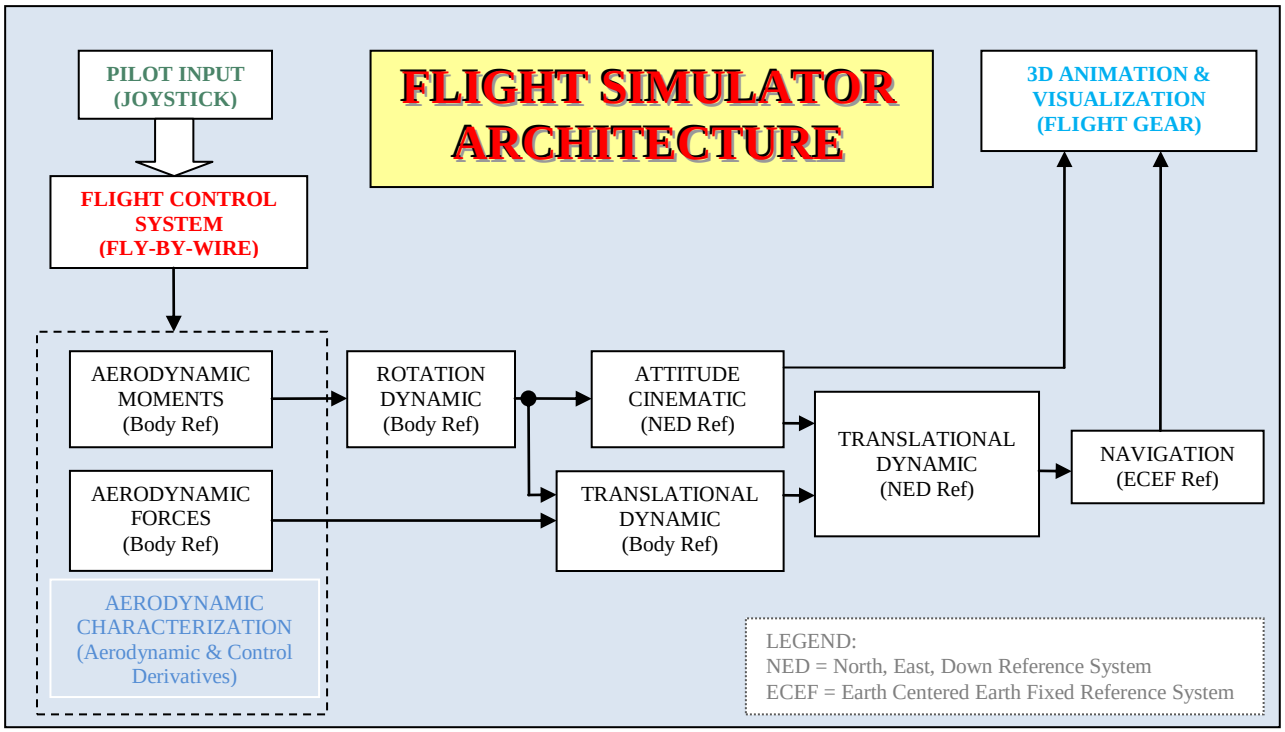
[12] Masarati, P., Muscarello, V., Quaranta, G., "*Linearized Aeroservoelastic Analysis of Rotary Wing Aircraft* ", 36th European Rotorcraft Forum, Paris, France, 7-9 September 2010

[13] Becker, B., Caldwell and Vaccaro, "*The interaction of flight control systems and aircraft structures*", In RTO AVT Specialists Meeting on Structural Aspects of Flexible Aircraft Control, Ottawa, Canada, 18-20 October 1999

[14] See: "*www.mathworks.it"*

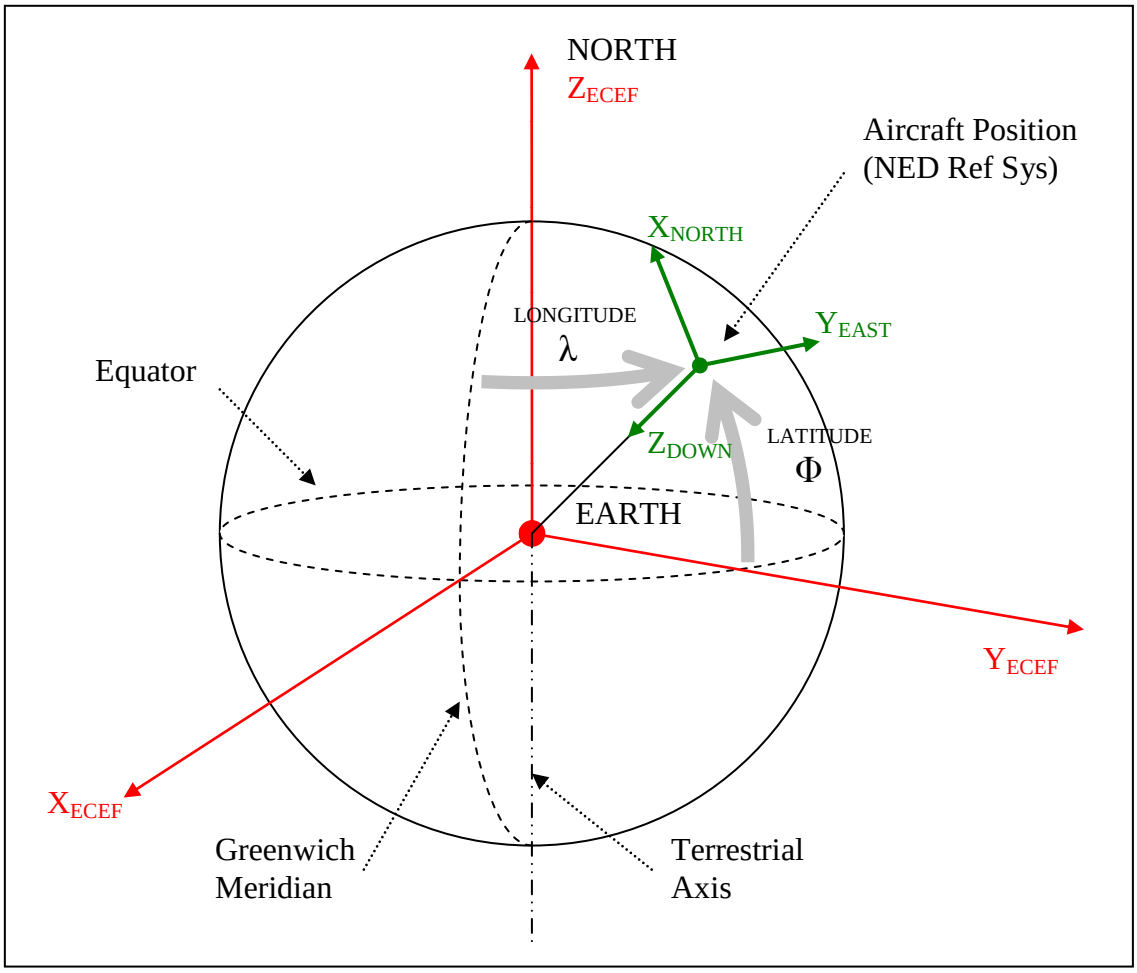
[15] See: "*www.flightgear.org"*

Appendix 1



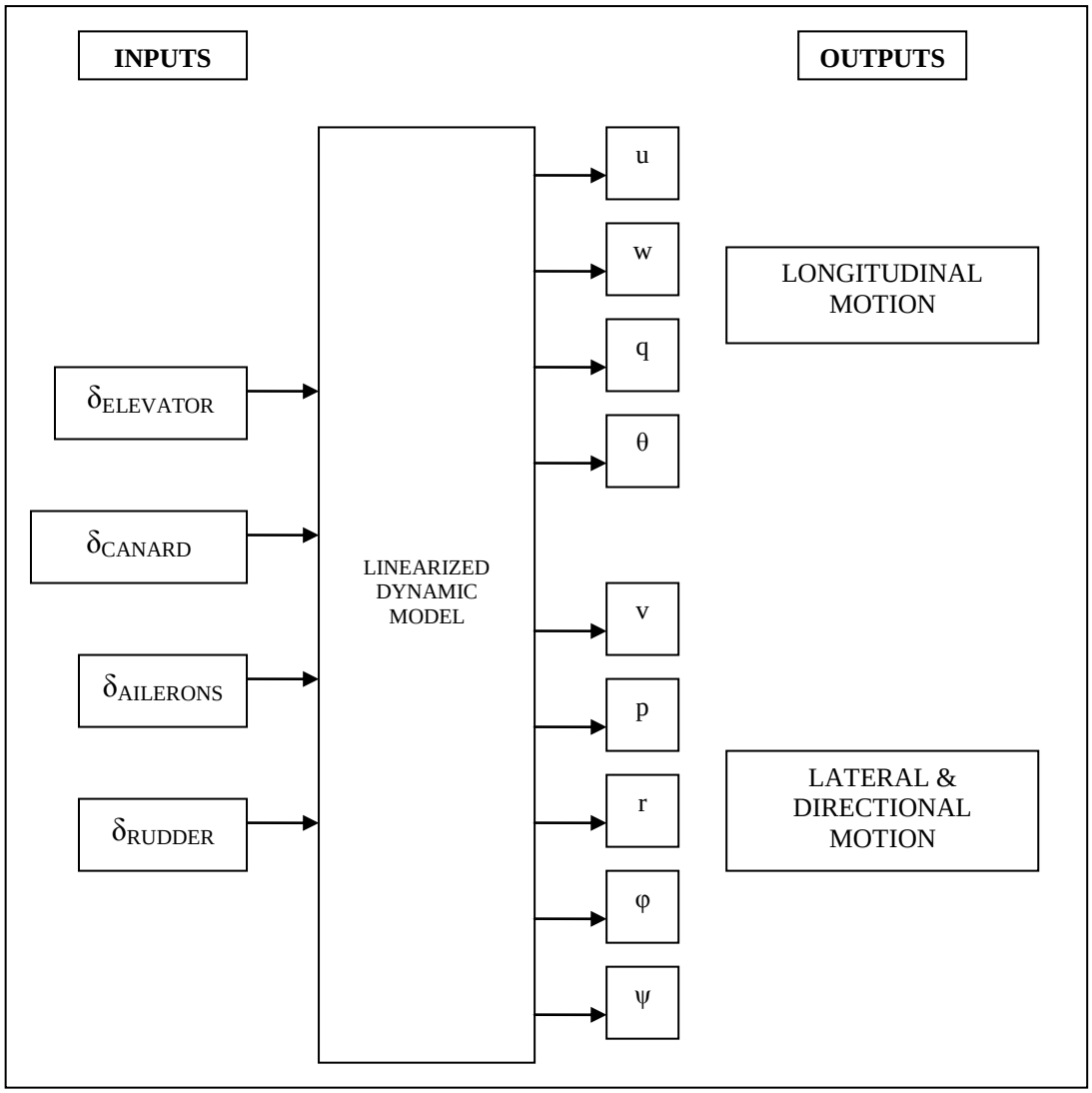
Flight Simulator Architecture: integration of two different codes, SIMULINK and FLIGHT GEAR

Appendix 2



NED and ECEF Reference Systems

Appendix 3



Aircraft linearized dynamic model

3rd CEAS Air&Space Conference
21st AIDAA Congress

Testing of multi layer insulators in martian-like environment

Diego Scaccabarozzi, Bortolino Saggin, Marco Tarabini
Politecnico di Milano, Department of Mechanics, Lecco's Campus, via M. D'Oggiono 18/A, 23900 Lecco, Italy

Keywords: *Multi layer insulator, degraded vacuum, cryogenic testing, hot guarded ring apparatus.*

Abstract

Low density materials are generally used as passive thermal insulators in space applications thanks to their thermal properties and lightness. We describe here a testing activity performed on a multi layer insulator (MLI) candidate as thermal barrier on MIMA, a space borne interferometer for the ExoMars Pasteur mission. The specimen was tested in degraded vacuum condition, carbon dioxide at various pressures and with temperature extremes within -100÷60 °C. MLI heat transfer properties were measured by means of guarded hot plate apparatus similar to the set-up described on ISO 8302 standard. Parameters derived by the testing activity, i.e. MLI equivalent conductivity, internal specific conductance and emissivity, are of primary importance for the thermal design of the instrument and the modeling of the system in operative condition.

1 Introduction

The performances of an MLI depend on its manufacturing characteristics, i.e. foils and spacers materials, number of layers, but also on the environment where it operates, in particular the type of atmosphere and its temperature and pressure. The normal application for this kind of insulators is mostly characterized by high vacuum where they achieve their best performances however, there are terrestrial applications where the vacuum may degrade or space conditions where the vacuum may not be granted for the whole mission phases as for the case of the exploration of celestial bodies having an atmosphere. Heat transfer in MLI with interstitial gas is a rather complex problem because gas thermal conductivity is often in a pressure region where pressure itself strongly influence thermal conductivity and the interlayer distance may ply a critical role in the relationship. So, in order to model its behaviour testing in the expected working conditions is generally required. Several studies on MLI thermal properties are present in literature. In reference [1] different models describing heat-flux through MLI material are presented. Another theoretical approach is proposed in reference [2], with an empirical description of MLI conductivity as function of temperature and pressure. Reference [3] shows results of MLI testing varying many influencing parameters, but no sufficient information was found about MLI performances at low temperature, pressure, and with CO_2 as

interstitial gas, all features that distinguish application in Martian environment. Besides, Aerogels® have been preferred to MLIs in the recent MER missions [4] but their application inside optical instruments poses non trivial problems about cleanliness while the MLIs have a long heritage on this aspect. Being the actual gas thermal conductivity at low pressures depending on the type of surfaces and distances different [5] types of MLI can exhibit quite different behaviour. To assess the performances of the candidate material a testing activity was performed and the results will be reported in the following. The paper is organized as follows: heat-flux modelling is described in chapter 2 while paragraph 3 shows in detail the testing facility developed for the study. Chapter 4 depicts the experimental results and the paper is eventually concluded in paragraph 5.

2 Modeling of the system.

Heat flux through the MLI specimen is in general the combination of conductive, convective and radiative processes. The specific heat-flux can be modeled as follows:

$$Q_i = (T_a^4 - T_b^4)\,\varepsilon_i\,\sigma + \Lambda_i (T_a - T_b) \quad (1)$$

where T_a, T_b are the temperatures at the MLI inner and outer layers, σ is the Stefan-Boltzmann constant, $\varepsilon_{i(N,p)}$ and $\Lambda_{i(N,p,T)}$ are respectively the internal emissivity and conductance, functions of the number of layers N and interstitial gas pressure p. The above is the common model in vacuum conditions because the two parameters are constant for an MLI and allows an easy and accurate modeling in a wide range of temperatures. Nevertheless, when a gas is filling the system the gas conductivity has to be considered and being gas conductivity function of the temperature, the internal conductivity becomes dependent on it too. The use of two parameters when one of them is temperature dependent is doubtful and a representation for instance with an equivalent specific conductance function of temperature may become preferable.

Considering the thermal system realized with the hot guarded ring apparatus, shown in figure 1, the relationship between the measured power, temperatures and the internal parameters of equation (1) can be easily determined.

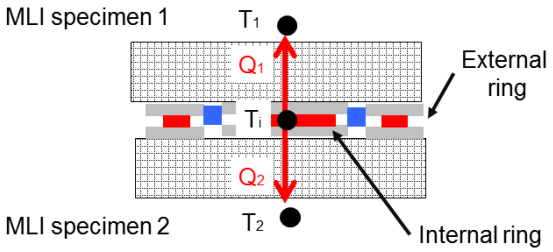


Fig. 1 Scheme of the heat-flux through MLI specimens with hot guarded ring apparatus.

The concept of the hot guard system is quite simple: an heater in the middle of the specimen warms up the internal side of the MLI up to the temperature *Ti* while, depending on the environmental conditions and MLI properties different temperatures T_1 and T_2 are reached on the external layers of the specimens. This is the setting for the double specimen configuration of the hot guarded ring apparatus described in standard [6]. The external ring is thermally controlled to the same temperature of the internal one to avoid any heat flux in radial direction. If temperatures and power are measured, and no heat flows radially, considering equation (1) in vacuum i.e. with parameters constant, one deals with four unknowns (ε_i, Λ_i,*Q1*,*Q2)*, while only three equations are available from the thermal balances. The parameters can be determined by considering two different sets of measurements with different temperatures i.e. by solving the following system of equations.

$$\begin{cases} Q = Q_1 + Q_2 \\ Q_1 = \Lambda_i (T_i - T_1) + (T_i^4 - T_1^4)\,\varepsilon_i\,\sigma \\ Q_2 = \Lambda_i (T_i - T_2) + (T_i^4 - T_2^4)\,\varepsilon_i\,\sigma \\ Q_0 = Q_{1,0} + Q_{2,0} \\ Q_{1,0} = \Lambda_i (T_{i,0} - T_{1,0}) + (T_{i,0}^4 - T_{1,0}^4)\,\varepsilon_i\,\sigma \\ Q_{2,0} = \Lambda_i (T_{i,0} - T_{2,0}) + (T_{i,0}^4 - T_{2,0}^4)\,\varepsilon_i\,\sigma \end{cases} \quad (2)$$

One can guess that if the two conditions are "similar" the system can be no more solved, so conditions quite different for instance in terms of total power should be achieved. In order to

assess the uncertainty of the proposed method and find an optimal configuration which minimizes the overall uncertainty on the extracted parameters, Monte Carlo simulations were performed using as reference values 0.016 and 0.022 $Wm^{-2}K^{-1}$ for the MLI internal emissivity and specific conductance. These values were derived from a previous testing activity on a similar MLI though differing in for the number of layers. A favorable condition for the testing would be that of an equal MLI external temperatures for both specimens, this would warrant a symmetrical heat flux through the specimens and only one condition would allow determining the MLI parameters; unfortunately, our set-up did not allow achieving such condition for all external temperatures. In the simulation, extreme temperatures were varied within the range of interest, shown in table 1. The uncertainty for the temperature measurement was computed in compliance to the tolerance defined by EN-60751 standard, being the measurement made with Pt100, class A sensors, while as a worst case, the relative uncertainty of the overall heat flux was estimated to be 1%. Simulations underlined the presence of a minimum, with 7.3% and 2.1% respectively for the MLI conductance and emissivity with T_i, T_1, $T_{i,0}$ and $T_{1,0}$ equal to 303K, 173K, 373 K, 293K i.e the extremes of the temperature ranges.

Parameter	minimum [K]	maximum [K]
Ti	293	373
T_1 T_2	173	293
$T_{i,0}$	293	373
$T_{1,0}$ $T_{2,0}$	173	293

Table 1 Temperature ranges used in the uncertainty propagation process.

The above modeling, as anticipated is not very convenient in case there is a residual gas inside the MLI. A common approach often used for MLI modeling is the one based on the equivalent conductivity, under the hypothesis that the heat flux exchange inside MLI specimens is driven by conductance. With the assumption of having equal T_1 and T_2 temperatures, equations (2) can be rewritten as:

$$Q = \frac{\lambda_e}{s}(T_i - T_1) \quad (3)$$

where λ_e is the equivalent conductivity and s is the MLI thickness. This parameter is widely used because it allows a direct comparison of the MLI performances with that of more common thermal insulators, for instance of the foams family however, one should take into account that it depends on the average temperature even for applications in vacuum.

3 Experimental Setup description

A sketch of the test setup is shown in figure 2.

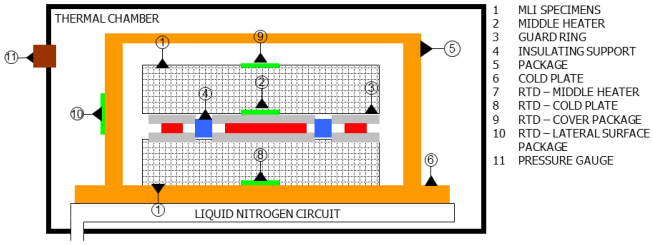


Fig. 2 Sketch of the setup for measuring thermal properties of MLIs.

Two heaters were packaged into aluminum disks: the central one provides an adjustable heat flux through the MLI specimens depending on the desired hot temperature *Ti* while the external one is powered to follow the temperature of the internal one by a variable voltage power supply driven by a closed loop control strategy. The temperature difference between the disks was measured by means of a copper/constantan thermopile (shown in figure 3) with 200 $\mu V°C^{-1}$ sensitivity and whose signal was exploited to feed the thermal control system of the guard ring. Moreover, in order to reduce the power measurement error, any thermal shunt between the central disk and other parts had to be minimized. The overall system was packaged into a cryostat eventually mounted inside a thermal chamber whose temperature was controlled by a cooling system based on liquid nitrogen circuit. Pictures of the test set up are provided in figure 4. At each measurement, several parameters were acquired:

- Central disk temperature T_i;

- cold temperature T_1 at the external side of the MLI specimen 1;
- cold temperature T_2 at the external side of the MLI specimen 2;
- middle heater supply voltage V_0 and current I_0;
- thermopile voltage V_{Th};
- pressure p_0 inside thermal chamber;

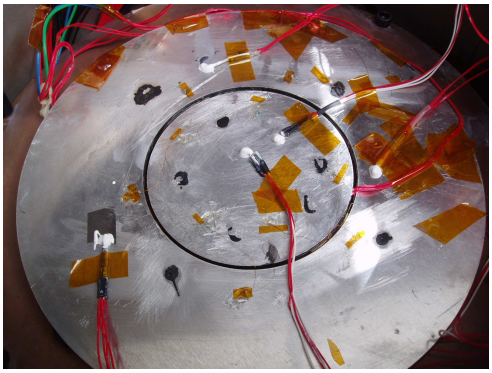


Fig. 3 Disks and thermopile during calibration process.

The pressure was monitored by means of one Pirani and one cold cathode gauges and was changed within the $10^{-1} \div 10^{5}$ Pa range in order to measure the parameters around the expected condition on Mars (10^{3}Pa).

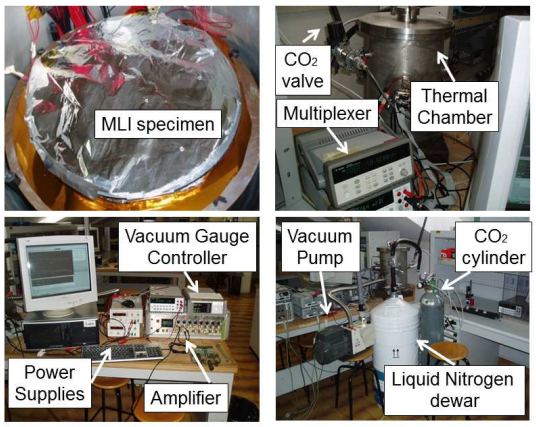


Fig. 4 Description of the testing setup.

To achieve the wanted pressure once vacuum condition was reached carbon dioxide was injected inside the chamber until the desired pressure was established. While keeping constant the pressure, middle disk temperature could be varied by changing the power dissipation. Temperature and power were recorded every 10 s but only the data collected with a constant Ti and minimum V_{Th} were used to perform the parameter extraction. An example of the transient condition up to the steady state is shown in figure 5.

4 Results

Despite, as mentioned above, the representation of the MLI with the internal emissivity and conductance is questionable in presence of internal gas, this approach has been used at first and it effectiveness then evaluated by using the parameters extracted one condition to others measured situations. Extraction of the MLI internal parameters was performed taking five sets of data for each pressure condition. This was done in order to combine tests with different average temperatures, like those shown in table 2 concerning the measured parameters at a pressure of 10^{3} Pa.

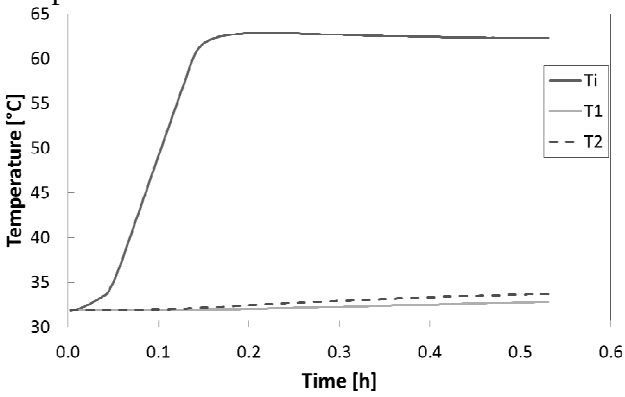


Fig. 5 Temperature measurements with 0.1 Pa pressure and uncooled chamber.

Ti [K]	T_1 [K]	T_2 [K]	Q [W]	$T_{av.\,1}$ [K]	$T_{av.\,2}$ [K]
311.3	248.4	271.4	2.26	279.9	291.4
318	242.7	276.5	2.50	280.4	297.3
310.8	255.4	281.2	1.88	283.1	296
346.1	279	302.8	2.71	312.5	324.5
337.4	305.8	314.9	1.36	321.7	326.2

Table 2 Data set for 10^{3} Pa and CO^{2} presence.

Figure 6 shows MLI internal specific conductance and emissivity trends computed with different pressure conditions, while their values are reported in table 3. On the last column of this table the relative average error in the overall heat power estimation obtained by

using the extracted parameters for all the tested conditions at the same pressure is reported to allow evaluating the effectiveness of the modeling. Positive trends with pressure increasing are evident for both parameters with the exception of the internal emissivity that, at the highest pressure, decreases. The relative error on the estimated powers based on the extracted parameters anyway remains lower than 2% in most of the tested configurations indicating that this type of modeling is still valid.

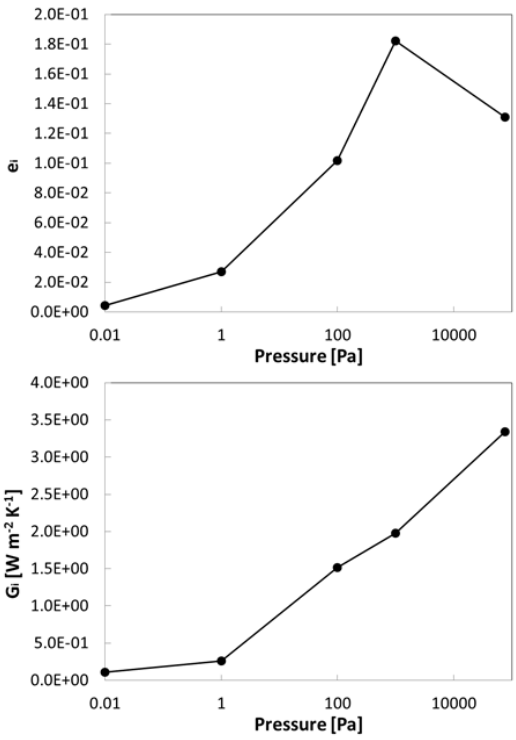


Fig. 6 MLI internal emissivity and conductance.

Pressure [Pa]	G_i	e_i	% average error
7.6 10^4	3.340	0.131	1
10^3	1.979	0.182	1
10^2	1.517	0.102	2.1
1	0.262	0.027	10
0.1	0.108	0.0045	7

Table 3. Summary of the MLI parameters and percentage error for the overall power estimation at different pressure conditions.

The error increasing at the lowest pressures where the modeling should become more accurate instead depends on the fact that these conditions of pressure where difficult to achieve and strong differences are expected in the actual MLI internal pressure and gas composition between various measurements. These pressures where achieved by regulating the vacuum pumping after the filling of the system with carbon dioxide but a stable condition could not be maintained so the measurements at these pressures (or better the actual pressure during these measurements) have only a qualitative meaning. According to equation (3) MLI equivalent conductivity was computed for each pressure and with different average temperatures. An interesting condition is the Martian "hot" case with cold and hot MLI sides respectively at -20°C 40°C. Equivalent conductivity as function of pressure is shown in figure 6.

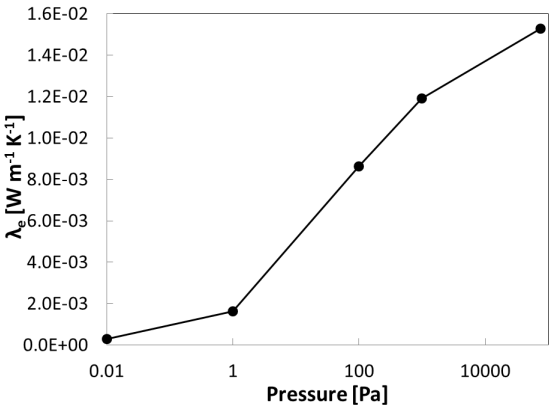


Fig. 7 Equivalent conductivity at average temperature 10 °C.

The equivalent conductivity strongly increases with the pressure. Moreover, at the highest pressure the measured conductivity is close to the expected value of the carbon dioxide in standard conditions. The most interesting result however is that the equivalent thermal conductivity at 1kPa is very close to values reported for the aerogels i.e. in the range 10-15 mW/(°Cm) in similar pressure and temperature conditions.

5 Conclusions

Aim of this work was the characterization of an MLI insulator selected as candidate thermal

barrier for an instrument bound to operate in Martian atmosphere. A characterization at low temperatures and with different pressures was performed. MLI internal and equivalent parameters were computed, The modeling with the internal emissivity and conductance has proved effective even in presence of gas though these parameters becomes dependent on the gas pressure. The thermal conductivity for pressures above 10 kPa, as expected, becomes close to that of the carbon dioxide at ambient pressure. The measured equivalent thermal conductivity in conditions similar to those expected on Mars have evidenced that it can be a valid alternative to Aerogels at least in cleanliness sensitive items such as the optical instruments.

Acknowledgment

This work has been carried-out in the framework of the development of the MIMA instrument. The authors gratefully acknowledge the financial support to the MIMA/Exomars instrument provided by ASI, the Italian Space Agency.

References

[1] C.K. Krishnaprakas, K. Badari Narayana, Pradip Dutta, "Heat transfer correlations for multilayer insulation system", Cryogenics 40 (2000) 431-435.

[2] A. Hofmann, "The thermal conductivity of cryogenic insulation materials and its temperature dependence", Cryogenics 46 (2006) 815-824.

[3] P.J. Sun, J.Y. Wu, P. Zhang, L. Xu, M.L. Jiang, "Experimental study of the influences of degraded vacuum on multilayer insulation blankets", Cryogenics 49 (2009), 719-726.

[4] S.M. Jones Aerogel: Space exploration applications", J Sol-Gel Sci Techn (2006) vol 40 pp. 351–357

[5] Dushman-Laferty "Scientific foundations of vacuum technique" John Wiley & Sons N.Y 1962.

[6] International standard, ISO 8302.

Parameterized Automated Generic Model for Aircraft Wing Structural Design and Mesh Generation for Finite Element Analysis

M. Sohaib
Master's Student in Mechanical Engineering
Linköping University, Sweden

Keywords: *Modelling and Simulation, Innovative design and configurations*

Abstract

This research paper presents the development of a parameterized automated generic model for the structural design of an aircraft wing. Furthermore, in order to perform finite element analysis on the aircraft wing geometry, the process of finite element mesh generation is automated. *[1]*

The generic model that is developed is able to automate the process of creation and modification of the aircraft wing geometry based on a series of parameters which define the geometrical characteristics of wing. Two different approaches are used for the creation of generic model of an aircraft wing which is "Knowledge Pattern" and "PowerCopy with Visual Basic Scripting" using CATIA V5 software. A performance comparison of the generic wing model based on these two approaches is also performed.

The process of finite element mesh generation for generic wing model is automated. The finite element mesh is generated for the wing panels, wing spars and wing ribs.

1 Introduction

During the initial conceptual stage of the aircraft design process, a large number of alternative aircraft configurations have to be studied and analyzed. A major portion of these studies goes towards the aircraft wing design. However, at this stage, high-end CAD software's are not used because they are thought to be too complex or demanding *[2]*. Therefore, the promising aircraft configurations have to be remodeled again later in the detailed design process. A generic model is required in this regard, that would automatically generate CAD models to speed up the design process of the analyzing different aircraft wing configurations especially from the structural design perspective.

Software's like PRADO (Preliminary Aircraft Design and Optimization Program) are able to create simple FE mesh but since they are not CAD based, it is not possible for the internal wing structure to be modeled. Modern CAD software like CATIA v5 offers an alternative in this regard which can be used for the preliminary aircraft design process. This requires that the designer doesn't spend time on modeling the wing geometry but rather on

analyzing the results. A creation of an automated generic model is thus desirable.

The aim of this study was to create an automated parameterized generic aircraft wing model using a CAD system that can be employed very early in the aircraft design process. The generic wing model is made up of trapezoidal panels, whose number and shape can be controlled by different parameters. A large number of parameters e.g. root and tip chord length, wing span, dihedral, leading edge sweep angle, twist, skin thickness and airfoil among others are defined so that the designer has the capability to design any type of wing configuration. Furthermore, the wing model also contains spars and ribs whose number, position, angle and thickness can also be controlled by different parameters. Moreover, an automated routine has been built up so as to generate a FEM compatible mesh of the entire wing geometry with its structural components. The resulting wing structural model can be analyzed using an FE analysis tool.

1.1 Tools and Methods

In the CATIA V5 software, there are a number of workbenches available that can be used for geometry generation and automation. In order to introduce knowledge in the generic model, and for the automation and management of the geometry, the tools available in the "Knowledgeware Workbench" and "product knowledge template" are used. There are two approaches that are adopted in this work, which is knowledge pattern and powercopy with VBA scripting. Both of these are implemented by using this workbench and also by using Microsoft Visual Basic IDE environment for VBA scripting and debugging of code.

Furthermore, since, the model requires sufficient surface modelling; a specialist workbench for surface modelling, "Generative Shape Design" is used. The generative shape design offers a wide variety of features and tools for advance surface modelling.

Thus, the generic model of the wing is completely developed in CATIA V5 including wing spars and wing ribs. Both the surface model and the solid model, which are part of the generic wing model, are integrated together. A number of parameters define the shape of the wing panels, wing spars and wing ribs.

Since, the automated mesh generation of the generic wing model geometry is also required; the tools available in the "Advanced Meshing Tools" workbench are used to integrate the wing model with the meshing program. As the generic wing model changes, the mesh attached with the generic wing model will change also. The tools available in the "Generative Structural Analysis" workbench can then be used to perform the structural analysis on the aircraft wing model.

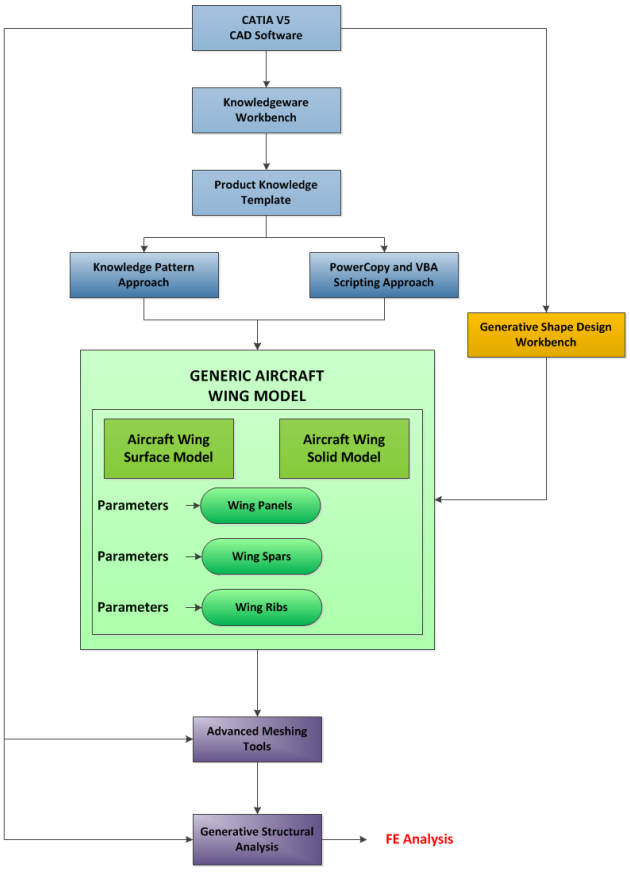


Fig. 1 Tools and methods used for the generic wing model

2 Generic Aircraft Wing Design Concept

A generic aircraft wing model should constitute an external surface of the wing, and should also be comprised of structural elements of the wing which are spars and ribs. In order to define accurately the external shape of the wing, a number of parameters are required to define the geometry of the wing in detail. The generic

aircraft wing model should comprise of both a surface model and a solid model of the geometry of the wing. Both of these models should be integrated together, which means that the parameters should change both the surface and the solid model of the aircraft wing simultaneously. The generic wing model should be able to transform into any wing planform and shape based on the changes in the wing geometry parameters.

2.1 Surface and Solid Model Integration

In the surface model, all the features of the wing (wing panels, spars and ribs) will be represented by a series of surfaces. As, the number of wing panels, spars or ribs are changed, the surfaces for each wing panels, spars and ribs will be connected together. The surface model is made because of its use in carrying out the structural analysis of the wing where, the surface model will be utilized to do the mesh generation of the entire aircraft wing geometry.

The solid model will have thicknesses attach to each wing panel, spars and ribs. The solid model can be used for the design purposes, but it is important that both the surface and the solid geometries are properly integrated inside the same wing model.

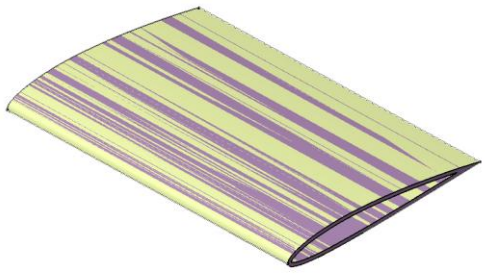


Fig. 2 Surface and solid model integration in generic wing in model

2.2 Concept for Wing Panels

The entire wing panels of the aircraft wing should be properly connected, which means that all the wing panels should be connected in such a way so as to yield a continuous surface of the wing. Furthermore, the geometric features associated with the tip of the wing panel should be the same as the root of the next wing panel, so as the geometric features e.g. tip chord length or tip airfoil associated with the tip of the wing panel are changed, the geometric features of the root of the next wing panel should updated accordingly.

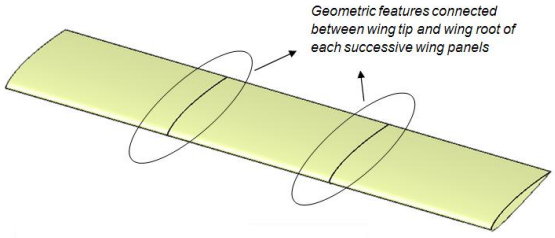


Fig. 3 Geometrical features connected in generic wing model

2.3 Concept for Wing Spars

For the wing spars, the spar position will be defined by “point values on curve” along the root and the tip of the wing. The point values on curve will range between 0 and 1. 0 means that the spar position will start at the leading edge while, 1 means that the spar position will start at the trailing edge. There are two approaches for the construction of wing spars inside the wing model. One is that the spars run continuously throughout the wing across all the wing panels while the other is that the wing spars are placed along each wing panel separately and then they are joined together to each other. In either approach, it is important that it is not possible for two spars to intersect each other in any way. Furthermore, it is important that, all new spar positions along the root and the tip of the wing should be modified based on the position of the old spars.

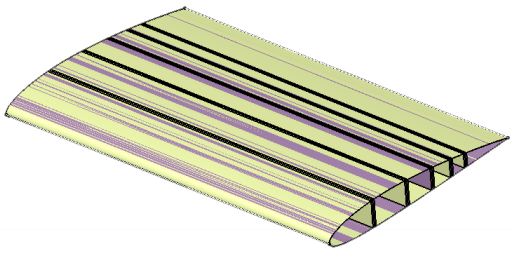


Fig. 4 Wing spars in generic wing model

The wing spars will have a thickness associated with them, however, this thickness should not protrude inside the thickness of the wing panel skin.

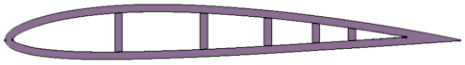


Fig. 5 Wing spars thickness not protruding in the wing thickness

2.4 Concept for Wing Ribs

Similar to the case of the spars, the wing rib positions will be defined by "point values on curve" along the span of the wing (front and aft curve of the wing). For the surface model, the rib will be placed perpendicularly or an angle and will just be composed of surfaces, while, for the case of solid model, it is important that the ribs are divided based on the number of spars present in the generic wing model. For example, if there is one spar present in the wing, the wing rib should be divided into two. If there are two spars present in the geometry, the rib is divided into three and so on. It is also important that no ribs should intersect each other. Similar to the case of spars, the ribs will have a thickness associated with them. It is important that the rib thickness doesn't protrude the spar thickness neither should it protrude the wing thickness associated with the skin of the wing panels.

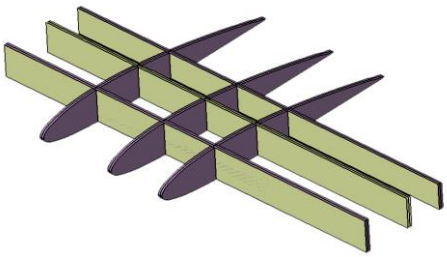


Fig. 6 Wing Ribs in the generic wing model

It is also to be ensured that a wing rib can be placed across multiple wing panels, as an example, in the figure below; a wing rib is placed across two wing panels. The wing ribs can also be placed at any angle inside the generic wing model and no wing ribs should intersect each other.

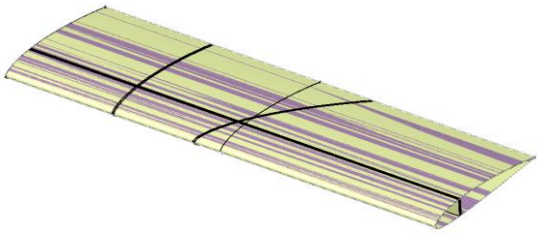


Fig. 7 Wing rib across multiple wing panels

3 Methodology for Generic Wing Model

The methodology adopted for the creation of the generic wing model for both the "knowledge pattern" and "powercopy with VBA scripting" approaches is shown in the figure below,

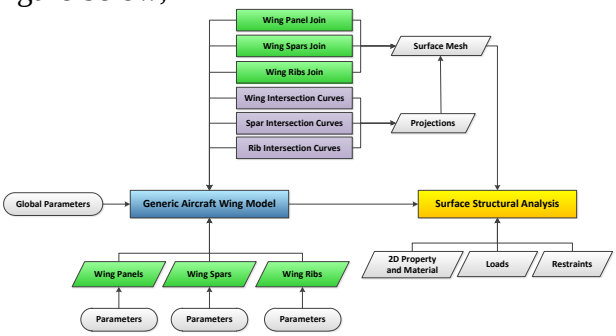


Fig. 8 Methodology for the generic wing model

The generic aircraft wing model is composed of both the surface and the solid model for wing panels, wing spars and wing ribs. Each wing panel, wing spar and wing ribs also have individual parameters that define the geometry and shape of each element, furthermore, there are also global parameters which control the number of wing panels, spars and ribs as well as the mesh characteristics. Whenever, a new wing panel, a wing spar or a wing rib is added into the model, a join which already exists in the model is updated with the new geometry. These joins are connected to each individual surface mesh for wing panels, wing spars and wing ribs. In order to ensure, that all the mesh elements are properly connecting at the nodes of the aircraft model, projections are required for generating the mesh in the correct fashion. This is done by defining three intersections between the join surfaces. These are Intersection between Wing and Spars, intersection between wing and ribs and intersection between spars and ribs. These projections are then used to correctly define the

surface mesh for wing panels, spars and ribs. In order to perform the structural analysis, 2D properties, materials, loads and restraints are required. A 2D property adds thickness to the surface. Different types of loads and materials can be applied to the wing structure. The 2D property for each surface is linked with a material as well.

4 Generic Aircraft Wing Model

The generic aircraft structural model is developed by knowledge pattern using EKL (Engineering Knowledge Language), Visual Basic scripting and geometry automation tools and features that are available in CATIA V5 CAD software. The generative structural analysis and Advanced meshing tools workbenches available in CATIA V5 software are used to perform the structural analysis and automated mesh generation of complete aircraft wing geometry. The aircraft structural wing model is made up of different elements whose number, shape and geometry can be changed by a range of different parameters. The aircraft wing model includes,

1. Wing Panels
2. Wing Spars
3. Wing Ribs

A model of an aircraft wing with (e.g. 3 panels) is shown in the fig. 9 below. The surface shown below the wing is the projected wing surface area calculated based on the wing geometry. Furthermore, the mean aerodynamic chord (MAC) of each individual wing panel (blue line) and lines showing the 25% and 50% chord positions are also shown.

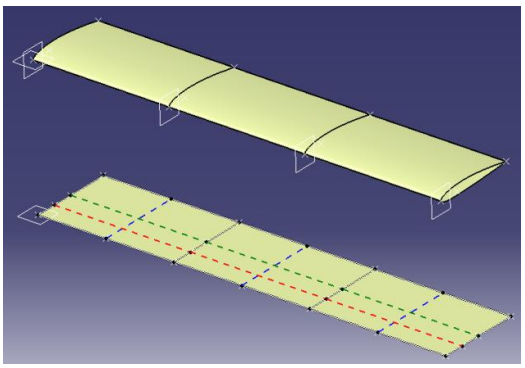


Fig. 9 Generic aircraft wing model with e.g. 3 wing panels

The generic wing model can be used to make any type of wing geometry as an example; an aircraft configuration including spars and ribs is shown in the fig. 10 below,

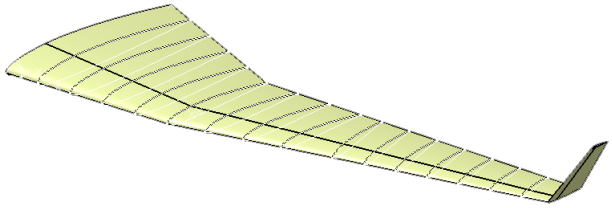


Fig. 10 An example of aircraft wing geometry created by using generic model

5 Automated Mesh Generation

The automated mesh generation is done in such a way, that the aircraft wing model can be modified and changed and the mesh will be updated accordingly. The surface model is used for creating the finite element mesh of the entire aircraft wing. This is done so because in the conceptual design phase, a detailed structural analysis is not required. At this stage, an estimate of the structural characteristic of an aircraft wing is beneficial for which a 2D structural analysis is more suitable in terms of computational time and resources. Furthermore, it enables fast analysis of different configurations of the aircraft wing as compared to a full 3D structural analysis.

The methodology used for the mesh generation of the generic wing model is shown in the fig. 11 below,

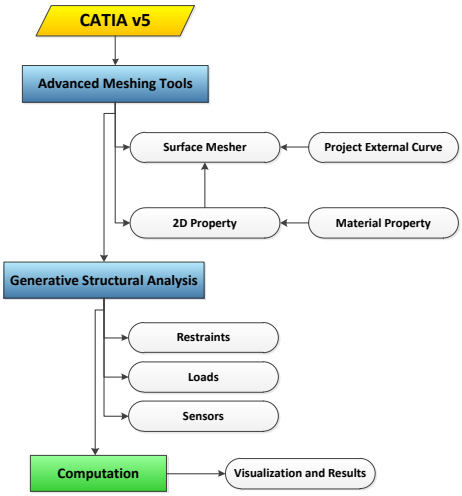


Fig. 11 Methodology for mesh generation of generic wing model

The surfaces of the generic wing model and the intersection curves are used to generate the mesh of the generic wing model.

5.1 Mesh Criteria

The finite element mesh of the aircraft wing should be of good quality. This means that there are not any uneven or large angles in the mesh nodes. Furthermore, it is essential that the all the mesh elements of the wings, spars and ribs are properly connected at the nodes. This means that the growth points of the mesh be properly controlled so that the mesh elements take into account the position and orientation of the ribs, spars and the wing panels.

5.2 Finite Element Mesh Generation Examples

Some examples of the finite element mesh generated for the generic aircraft wing model are shown in the fig. 12 and fig. 13.

The mesh is dynamically updated, so that any change in the wing geometry will automatically update the mesh for the new geometry.

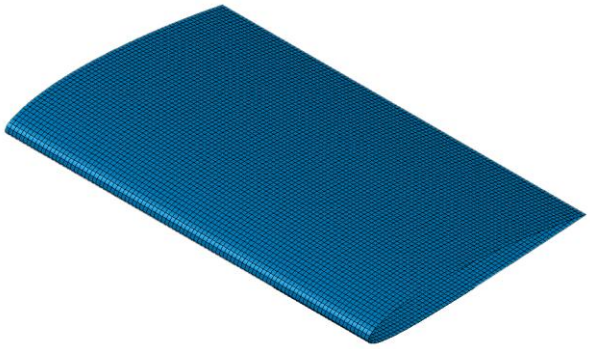


Fig. 12 Finite element mesh generated for the wing panel

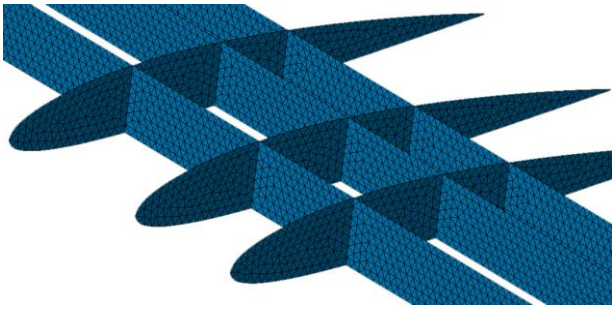


Fig. 13 Finite element mesh connectivity between wing spars and wing ribs

6 Performance Comparison

A comparison between the knowledge pattern approach and the power copy with VBA scripting approach is performed for the generic wing model.

The time to instantiation and deletion of the generic wing geometry is measured for the two approaches. In order for the results to be as accurate as possible, the test is performed on the same computer with only CATIA v5 CAD software running in the foreground and all other applications close with minimum number of processes of windows operating system running in the background. The computer that was used to perform the test is,

Computer: Intel Pentium Dual CPU T2390 @ 1.83 GHz with 1.99 GB of RAM

Operating System: Windows XP Professional, Service Pack 3

6.1 Test # 1: Time to Instantiate Wing Panels

A comparison between the times to instantiate wing panels in both approaches is shown in the table 1 below,

Table 1 Time to instantiate wing panels

Number of Wing Panels	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3	11.9 sec	45.2 sec
5	16.4 sec	1m 45 sec
10	37.8 sec	6m 50 sec

6.2 Test # 2: Time to Delete Wing Panels

A comparison between the times to delete the wing panels in both approaches is shown in the table 2 below,

Table 2 Time to delete wing panels

Number of Wing Panels	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3 to 1	2.8 sec	57.2 sec
5 to 1	4.1 sec	2m 5 sec
10 to 1	6.9 sec	6m 37 sec

Test # 3: Time to Instantiate Wing Spars

A comparison between the times to instantiate wing spars in both approaches is shown in the table 3 below. In this case, the number of wing panels is set to one for both approaches.

Table 3 Time to instantiate wing spars

Number of Wing Spars	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3	3.1 sec	12 sec
5	3.5 sec	22.3 sec
10	6.1 sec	40.9 sec

Test # 4: Time to delete Wing Spars

A comparison between the times to delete wing spars in both approaches is shown in the table 4 below. In this case, the number of wing panels is set to one for both approaches.

Table 4 Time to delete wing spars

Number of Wing Spars	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3 to 1	1.3 sec	11 sec
5 to 1	1.5 sec	18.2 sec
10 to 1	1.9 sec	33.6 sec

Test # 5: Time to instantiate Wing Ribs

A comparison between the times to instantiate wing ribs in both approaches is shown in the table 5 below. In this case, the number of wing panels is set to one for both approaches.

Table 5 Time to instantiate wing ribs

Number of Wing Ribs	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3	2.5 sec	11 sec
5	3.5 sec	20 sec
10	6.8 sec	44 sec

Test # 6: Time to Delete Wing Ribs

A comparison between the times to delete wing ribs in both approaches is shown in the table 6 below. In this case, the number of wing panels is set to one for both approaches.

Table 6 Time to delete wing ribs

Number of Wing Ribs	*Knowledge Pattern Approach*	*Power Copy with VBA Scripting Approach*
3 to 1	1.2 sec	5.3 sec
5 to 1	1.4 sec	9 sec
10 to 1	1.6 sec	18.9 sec

It is observed that in both the time to instantiate and the time to deletion for all the six tests on the generic aircraft wing model, the knowledge pattern approach is *significantly faster* than the power copy with visual basic scripting approach. The difference in time is

sufficiently clear in both the time to instantiate and time to delete the wing panels, while the difference is less in the time to instantiate and time to deletion of the wing spars and wing ribs. Since, a lot of geometric data has to be added and deleted from the generic model in powercopy with VBA scripting approach, the addition and deletion takes more time than the equivalent knowledge pattern approach.

7 Conclusions

The following conclusions are made from this research.

1. Both the knowledge pattern approach and the powercopy with VBA scripting approach can be used successfully in creating a generic wing model as shown in this thesis work.

2. If the time for modification, update and deletion of the geometry of the generic model is critical, then the knowledge pattern model is superior to the powercopy with VBA scripting model. The knowledge pattern based generic model was faster in all the tests compared to the powercopy with VBA scripting based generic model.

3. Not all features and tools in CATIA can be accessed by the knowledge pattern, whereas, a large number of features and tools can be accessed using VBA.

4. Not all geometrical entities of the generic model are accessible in the knowledge pattern based model while, all the geometrical entities of the generic model are accessible in the powercopy with VBA scripting based generic model.

5. The amount of code required to be written for the knowledge pattern based generic model is significantly smaller than the powercopy with VBA scripting based model.

6. Debugging and error checking is much easier in powercopy with VBA scripting than in the knowledge pattern due to the availability of the debugging tools in VBA environment.

7. Automatic finite element mesh generation and modification of the generic wing model is possible and efficient.

8 Aircraft Wing Design Examples

A rectangular wing planform along with 3 wing spars and 3 wing ribs generated from the generic wing model is shown in the fig. 14 below.

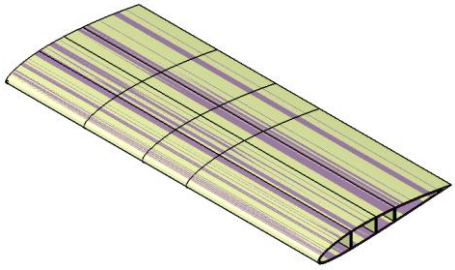


Figure 14 Rectangular Wing Design

A tapered wing planform design generated from the generic wing model is shown in the fig. 15 below.

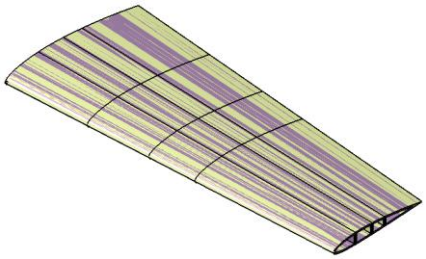


Figure 15 Tapered Wing Design

References

[1] Sohaib M., “Parameterized automated generic model for aircraft wing design and mesh generation for finite element analysis,” Master Thesis, Linköping University, Linköping, 2011.

[2] Master Thesis Proposal, “Aircraft Wing Design for Initial Design Space Exploration” IEI, Linköping University, Sweden

SMAT – An Advanced Land and Environment Monitoring System

S. Maddaluno, B. Tranchero and M. Fornaiolo
Alenia Aeronautica, Italy

Keywords: *Unmanned Air Systems, Environment monitoring, Flight Test, MALE.*

Abstract

The SMAT-F1 project is the first phase of a broader initiative aimed at defining, designing and developing an advanced land and environment monitoring system based on Unmanned Air Systems. The whole initiative is split in four phases and its name is: "Sistema di Monitoraggio Avanzato del Territorio" (SMAT). The project working group is a well balanced team with relevant participation of large industries, universities and Research Centres and Small Medium Enterprises. The coordination of the project is assured by Alenia Aeronautica. The innovative architecture designed during the project will be tested in flight. The final demonstration flight is planned to be performed in Fall 2011 and will be performed in Piemonte, in an area located in the south of the region.

1 Introduction

SMAT-F1 is a research project funded by the Regione Piemonte (Italy), and co-funded by the European fund for regional development (F.E.S.R.) within the Regional Operative Program 2007/2013 [1].

SMAT-F1 started on January 2009 and the project working group is composed of three large industries: Alenia Aeronautica, Selex Galileo, Altec; two universities: Politecnico di Torino, Università di Torino; one Research Centre: Istituto Superiore Mario Boella and eleven Small Medium Enterprises (SME) operating in the Regione Piemonte: Auconel, Axis, Blue Engineering, Carcerano, DigiSky, Envisens Techonologies, Nautilus, Nimbus, SEPA, SPAIC, Synarea.

In the industrial field it's commonly agreed that innovation can spring from two different causes, either as the answer to a new demand or thanks to a technological breakthrough that leads to novel applications, while allowing new solutions to existing problems.

The SMAT project springs from new requirements of cost reduction, capabilities and performances that are maturing in the latest years in the wide domain of surveillance, but it also belongs to the latter of the two causes, as it is bound to the evolution of technologies of unmanned aircraft that are assuming an important role in aeronautics and that are leading to a new class of highly reliable products.

SMAT represents the point of contact between the increasing need of Governmental Agencies

and private organizations of a land and environment monitoring system and the technology evolution in Unmanned Air Systems (UAS). As a matter of fact, Unmanned Air Vehicles (UAV) can been employed in so called DDD missions (Dull, Dirty, and Dangerous) and in routine missions, that require a long flight endurance. Ordinary and extraordinary land monitoring are mission typologies with these peculiarities and in addition UAS life-cycle costs are smaller than the costs of comparable manned aircraft.

One of the main point of innovation concerns the fact that SMAT is based on the use of several unmanned platforms that are able to cooperate through a ground coordination and control station. The flight in non-segregated airspace is still a main issue because of the lack of a common consensus and agreement on the regulation of UAS and for this reason a specific attention is devoted to this topic in the project.

The system will lead to the integration of several resources that already exists on the territory.

The first phase of the project consists is focused on the demonstration of the integration of the flying platforms and the realization of the Station for Coordination and Supervision (SSC). The objective of the next phases will be the functional demonstration of the whole surveillance system, that will be followed by the design and realization of the system of system prototype and then will be concluded by the industrialization.

2 The SMAT architecture

The SMAT architecture is organized into three segments: aerial, ground and communications.

The Aerial Segment [2] is composed of innovative Unmanned Air Vehicles; three different typologies of UAV will form the aerial segment:

- A Medium Altitude Long Endurance (MALE) UAV, developed by Alenia Aeronautica
- The Falco: a Medium Altitude – Medium Endurance (MAME) UAV, produced by Selex Galileo
- The D–Fly: a Low Altitude - Short/Medium Endurance (LASME) UAV, developed by Nimbus.

The Ground Segment is formed by:

- The Ground Control Station of each UAV
- A Supervision and Coordination Station (SSC), which will coordinate the mission, will elaborate the data received in real time by the UASs and will disseminate it to the Operative Centres on the territory.

The Communication Segment will assure a reliable, secure, efficient data transmission among the different components of the system, in particular:

- from the A/C to the GCS, through data links,
- from the GCSs to the SSC and from the SSC to the Operative Centres, through wireless and/or landlines networks.

The SMAT architecture will be integrated with the existing surveillance network. For the first phase of SMAT, i.e. SMAT-F1, the MALE system is represented by the SKY-Y of Alenia Aeronautica while the Low Altitude - Short/Medium Endurance UAS will be the C-Fly, designed by Nimbus.



Fig. 1 – Alenia Aeronautica Sky-Y

The resulting, complex, operational picture can be seen in the next figure.

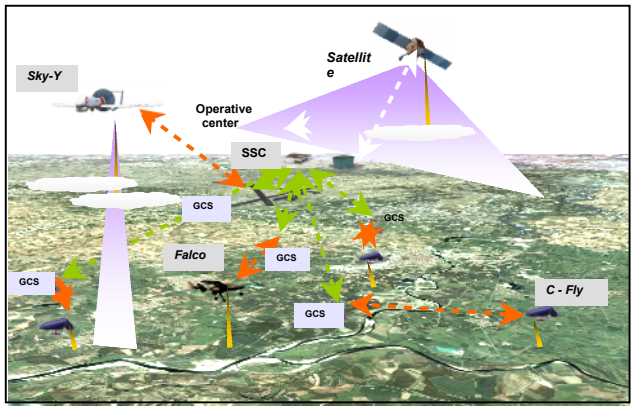


Fig.2 – SMAT Operational View

3 SMAT applications

The purpose of SMAT project is to study and demonstrate a surveillance system capable to support prevention and control of a wide range of events [3]. Different categories of users and needs have been identified, ranging from the highly sophisticated and real time operators down to simple users that only need post-processed data to be delivered offline on a pre-planned schedule [4].The potential SMAT System applications are:

- Monitoring of specific areas aimed at prevention,
  - – Surveillance of power transmission lines (power lines, gas and oil pipelines)
  - – Surveillance of areas subject to fire hazard.
- Territory surveillance aimed at planning:
  - – Monitoring of rural districts,
  - – Monitoring of rivers/watercourses,
  - – Data gathering on urban areas for town planning,
  - – Traffic monitoring, both urban and suburban.
  - – Surveillance of areas subject to natural disasters (landslides, floods, earthquakes, fires)
  - – Continuous monitoring of areas hit by natural disasters,
  - – Support to the police forces with the possibility of providing a back-up communication facility to the damaged area.
- Border patrol and critical infrastructure surveillance
  - – Border and coastline patrolling against illegal activities and illicit immigration
  - – Surveillance of critical infrastructures (industrial plants, dams etc)
  - – Search and rescue support
- Surveillance of areas with high population density
  - – Areas with high industrialization and high pollution risk
  - – Surveillance of major events.

.

The target of SMAT-F1 is to demonstrate, by performing a demonstration flight, the use of three integrated UASs inside of a primary scenario relevant for the purposes of Regione Piemonte.

4 Technologies

During the development of the SMAT-F1 project a number of technologies particularly relevant for obtaining a competitive product, were investigated [5].

A not exhaustive list of the topics investigated during the projecy is:

- Integration of a complex and distributed system
- Autonomous Flight
- High altitude and long endurance flight
- Diesel/Hybrid Propulsion, power generation system
- Innovative materials
- Advanced SW and HW systems to control the mission and the flight
- Navigation systems based on EGNOS/Galileo
- Remote sensing Sensors
- Communication Network.

Some of these technologies, in particular the ones that reached a proper technology readiness level during the project, will be tested during the final demonstration.
At the time being the definition of systems and subsystem is completed, and most of the subsystem and system integration is completed as well.
In particular the UASs have been properly modified to integrate remote sensing sensors adequate to support monitoring and surveillance tasks, such as electro-optica/infrared (EO/IR) sensors.
In the following pictures there are some examples of sensors data captured by the UAV sensors in different missions.



Fig. 3 – EO sensor data: coastline patrol



Fig. 4 – IR sensor data: Surveillance of critical infrastructure



Fig. 5 – EO sensor data: Surveillance of areas subject to natural disasters (fire)

The interfaces between UAS's Ground Control Stations and Supervision and Coordination Station have also been developed and tested [6]. The SSC is one of the main output of the project and it is really the core of the system, able to elaborate the data received in real time and to disseminate information to the final users. The following figure provides an overview of the SSC HW configuration [7]. It refers mainly to the part that supports the hosting and execution of the specific SMAT-F1 software application.

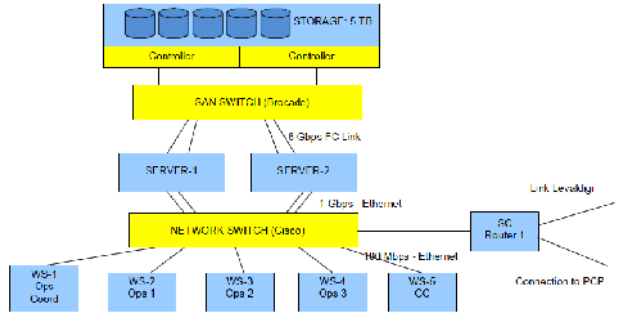


Fig. 6 – SSC HW configuration

The SSC development required a huge effort by the project team.
For each UAS involved in the SMAT-F1 mission, the SSC, before the start of the mission, provides Ground Segment with a mission request, specifying:

- Area Of Operation,
- Position, type and size of the target to monitor,
- Time to start the operations,
- Duration of the operations,
- Mission constraints,
- Allocation of communication frequencies,
- Dissemination plan, including time slot for the transmission of the collected data.

During the mission the SSC receives from each Ground Segment the data collected by the Air Segment and:

- performs the second level exploitation of the data,
- monitors the mission progress through the visualisation of the UAS Air Segment on a
- tactical map,
- presents the received data on screen.



Fig. 7 – SMAT SSC

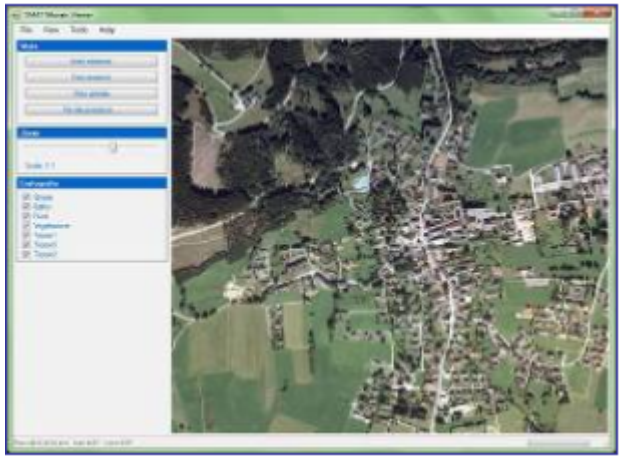


Fig. 8 – SSC User Interface

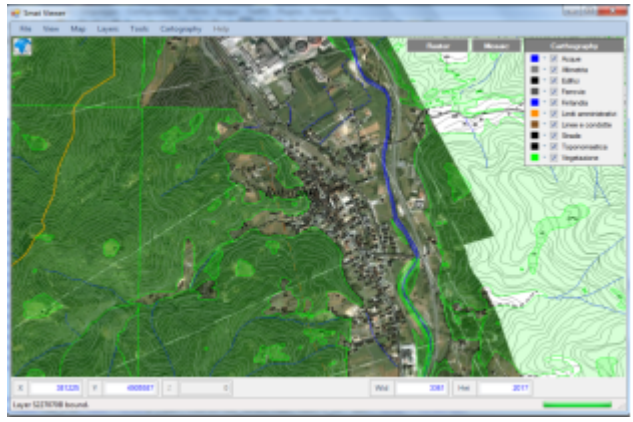


Fig. 9 – SSC User Interface

5 Flight test

The demonstration flight is planned to be performed in Fall 2011 and will be performed in Piemonte, in particular in an area located in the south of the region. Appropriate flight corridors and flight zones were identified in the surrounding of Cuneo starting from the Levaldigi civil airport.
The selected flight test area presents a number of features particularly interesting for the scope of the demonstration as:

- Mountains
- Hills
- Plain terrain
- Tanaro Basin
- Agricultural Areas
- Industrial Areas
- Freeways
- Railways.

A strong collaboration with Italian Certification Authorities (ENAC) is on going to perform all the activities necessary to obtain the Permit to Flight in the selected area.

The usual process for obtaining the Permit to Flight in this case is more complex because the test activities concerns a civilian airport and for the first time three different UAS are supposed to operate in the same area at the same time.

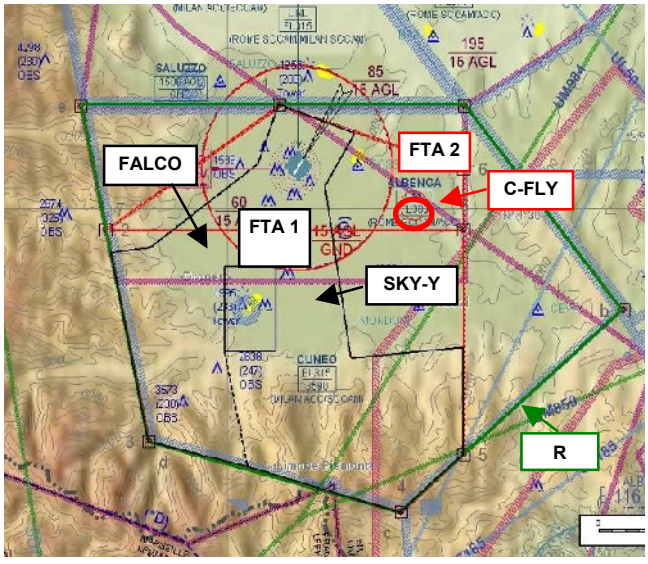


Fig. 10 – Flight test area

6 Conclusion

The SMAT-F1 project is the first phase of a broader initiative aimed at defining, designing and developing an advanced land and environment monitoring system based on Unmanned Air Systems.
The research program is the result of the cooperation of many stakeholders such as Piemonte Region, Department of Protezione Civile and all the project team, large industries, academia and research centers, small medium enterprises.
The system will be an important asset for emergency management and for supporting prevention and control of a wide range of events.
The final demonstration phase is planned to be performed in Fall 2011 and will be performed in Piemonte (south of the region). The area is highly populated with several cities, towns, and extensive road network, communication and power lines and bridges; many agricultural fields are present in the territory as well as industrial infrastructures.

Acknowledgments

The SMAT research program is the result of the work of a very large number of persons each one with indispensable specific skills and competences, so it is not possible to explicitely mention all of them. The authors thank all the people that directly or indirectly have contributed to the success of the project.

References

[1] Tranchero B. et aa., “Sistema di Monitoraggio Avanzato del Territorio – Fase 1”, Alenia Aeronautica, 2008.

[2] P. Nurra., “D3.2 – 01 SMAT-F1 Aerial Segment System Requirements,” Alenia Aeronautica, 2010.

[3] Mascolo D. et al., “Definition of SMAT-F1 Concept of Operations,” Alenia Aeronautica, 2010.

[4] Mascolo D. et al., “D1.1-01 SMAT Reference Scenarios and Operational requirements,” Alenia Aeronautica, 2009.

[5] Task 1.3 Team, “D1.3-02 Aggiornamento Rapporto descrizione stato dell’arte,” Nimbus, 2009.

[6] Altec Team, “D3.1-01 SMAT-F1 SC&SSC Design Specification,” Altec, 2009

[7] Altec Team, “D4.21-01 SC&SSC HW Functional Description”, Altec, 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Analysis of drop nose concept with 1D fem models.

F. Miglioretti[a]*, E. Carrera[a]
[a]Department of Aeronautic and Space Engineering, Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 Torino, Italy.

Abstract

Morphing structures are structures that change shape or state in order to change their operating characteristics or as a response to changes in the environment conditions. It is of special interest to achieve a camber variation with one structural system providing a smooth contour having no additional gaps. In this paper a new beam models able to evaluate the local deformation of a smart wing is presented. A refined beam finite element (FE) formulation is employed which permits us to obtain up-to-fourth-order expansions for the three displacement components over the section domain. Classical (Euler-Bernoulli and Timoshenko) beam theories are considered as particular cases. Comparison with 3D solid FE analyses are provided.

1 Introduction

Shape morphing has been used to identify those aircraft that undergo certain geometrical changes to enhance or adapt to their mission profile. In spite of there is not a clear definition for an enabling "shape morphing", it is a general agreement that the conventional hinged control surfaces or high lift device, such as flap or slat that provide discrete geometry changes cannot be considered as "morphing". No matter how one would choose to define the wing morphing, Reich and Sanders [1] listed the major challenges of shape morphing aircraft design to be: the requirement for distributed high-power density actuation, structural mechanization, flexible skins, and control law development. An example of shape-morphing complexity it can possible find in the case of the classical shape adaptation problem: the realization of airfoils with changeable camber for the purpose of regulating lift. A substantial conflict occurs between the deformability requirements, defining the change in camber which must be allowed by the airfoil, and the stiffness requirement, which limit the shape changes under load. It is of special interest to achieve a camber variation with one structural system providing a smooth contour having no additional gaps. Slender bodies such us airplane wings, helicopter blades, bridges, and frames are mainly analyzed through beam theories. The use of one-dimensional 1D models is 'historically' preferred to the introduction of more cumbersome two-dimensional 2-D (plate and shell theories) and threedimensional 3-D analyses. Classical 1-D models for beams made of isotropic materials are based on Euler-Bernoulli [2], de Saint-Venant [3, 4] and Timoshenko [5, 6]. The first two do not account for transverse shear deformations. Timoshenko foresees a uniform shear distribution along the cross-section of the beam. Mucichescu [7] made a comparison between Eulero-Bernoulli and Timoshenko. None of these are able to detect non-classical effects such as warping, out- and in-plane deformations, torsion-bending coupling and localized boundary conditions, both geometrical and mechani-

*Author for correspondence: federico.miglioretti@polito.it

cal. These kinds of effects are usually due to small slenderness ratios, thin-walled structures, and the anisotropy of the materials. Moreover in the field of smart wing this kind of effects are wanted in order to improve the wing efficiency. Many higher-order theories have been introduced to include non-classical effects which are based on enhanced displacement fields over the beam cross-section. Some considerations on higher order beam elements were made by Washizu [8]. An advanced model was proposed by Kanok-Nukulchai and Shik Shin [9]; they improved classical finite beam elements introducing new degrees of freedom to describe cross-section behavior. The above literature overview clearly shows the interest on further developments on refined theories for beams. Due to that interest, Carrera and co-authors have recently proposed refined 1D theories with only generalized displacement variables for the analysis of compact and thin-walled sections/airfoils. Higher-order models are obtained in the framework of the Carrera Unified Formulation, CUF. This formulation has been developed over the last decade for plate/shell models [10, 11, 12, 13] and it has recently been extended to beam static and dynamic modeling [14, 15, 16, 17, 18, 19]. CUF is a hierarchical formulation which considers the order of the model as a free-parameter (i.e. an input) of the analysis. In other words, refined models are obtained with no need for *ad hoc* formulations. Beam theories are obtained on the basis of Taylor-type expansions. Euler-Bernoulli and Timoshenko beam theories are obtained as particular cases of the first-order expansion. The finite element method is used to handle arbitrary geometries as well as geometrical and loading conditions. In this paper the finite element formulation of refined beam models based on Taylor polynomials are used to made a preliminary study about the problem of drop nose. A simple test-case has been used in order to verify the models capabilities ad is useful in the field of smart wing.

2 Beam model and FEM formulation

The coordinate reference is presented in Fig. 1, x and z are the section coordinates while y beam longitudinal axes. The displacement vector is:

$$\mathbf{u}(x,y,z) = \left\{ \begin{array}{ccc} u_x & u_y & u_z \end{array} \right\}^T \tag{1}$$

The superscript T represent the transposition

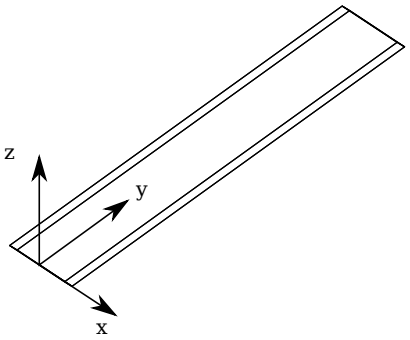


Figure 1: Beam

operator. Stress, σ, and strain, ϵ, components are grouped as follows :

$$\begin{array}{ll} \boldsymbol{\sigma}_p = \{\sigma_{zz}\ \sigma_{xx}\ \sigma_{zx}\}^T & \epsilon_p = \{\epsilon_{zz}\ \epsilon_{xx}\ \epsilon_{zx}\}^T \\ \boldsymbol{\sigma}_n = \{\sigma_{zy}\ \sigma_{xy}\ \sigma_{yy}\}^T & \epsilon_p = \{\epsilon_{zy}\ \epsilon_{xy}\ \epsilon_{yy}\}^T \end{array} \tag{2}$$

The subscript "n" stands for terms laying on the cross-section, while "p" stands for terms laying on planes which are orthogonal to Ω. Linear strain-displacement relations are used:

$$\begin{array}{l} \epsilon_p = \boldsymbol{D}_p \boldsymbol{u} \\ \epsilon_n = \boldsymbol{D}_n \boldsymbol{u} = \left(\boldsymbol{D}_{n\Omega} + \boldsymbol{D}_{ny}\right) \boldsymbol{u} \end{array} \tag{3}$$

whit

$$\boldsymbol{D}_p = \begin{bmatrix} 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} & 0 & 0 \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial z} \end{bmatrix} \quad \boldsymbol{D}_{n\Omega} = \begin{bmatrix} 0 & \frac{\partial}{\partial z} & 0 \\ 0 & \frac{\partial}{\partial x} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$\boldsymbol{D}_{nz} = \begin{bmatrix} 0 & 0 & \frac{\partial}{\partial y} \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \end{bmatrix} \tag{4}$$

Hooke's law is used to compute stress components:

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\epsilon} \tag{5}$$

According to the Eq. 2, the previous equation becomes:

$$\begin{aligned}\boldsymbol{\sigma}_p &= \boldsymbol{C}_{pp}\boldsymbol{\epsilon}_p + \boldsymbol{C}_{pn}\boldsymbol{\epsilon}_n \\ \sigma_n &= \boldsymbol{C}_{np}\boldsymbol{\epsilon}_p + \boldsymbol{C}_{nn}\boldsymbol{\epsilon}_n\end{aligned} \tag{6}$$

Where the material matrices are grouped as follows:

$$\begin{aligned}\boldsymbol{C}_{pp} &= \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & C_{66} \end{bmatrix} \\ \boldsymbol{C}_{nn} &= \begin{bmatrix} C_{55} & 0 & 0 \\ 0 & C_{44} & 0 \\ 0 & 0 & C_{33} \end{bmatrix} \\ \boldsymbol{C}_{pn} = \boldsymbol{C}_{np}{}^T &= \begin{bmatrix} 0 & 0 & C_{13} \\ 0 & 0 & C_{23} \\ 0 & 0 & 0 \end{bmatrix}\end{aligned} \tag{7}$$

The C_{ij} components are:

$$\begin{aligned}C_{11} = C_{22} = C_{33} &= \lambda + 2\mu \\ C_{12} = C_{23} = C_{23} &= \lambda \\ C_{44} = C_{55} = C_{66} &= \mu\end{aligned} \tag{8}$$

Where

$$\begin{aligned}\mu \equiv G &= \frac{E}{2(1+\nu)} \\ \lambda &= \frac{\nu E}{(1+\nu)(1-2\nu)}\end{aligned} \tag{9}$$

The Eq. 6 is used in the numerical analysis reported in this paper. In the framework of the Carrera Unified Formulation (CUF) ([11], [12], [14], [17], [15]), the displacement field is assumed as an expansion in terms of generic function, F_τ:

$$\boldsymbol{u} = F_\tau(x,z)\boldsymbol{u}_\tau(y) \qquad \tau = 0, 1, ..., M \tag{10}$$

where F_τ are functions of coordinates x and z on the cross-section. u_τ is the displacement vector and M stands for the number of terms of the expansion of order N. According to the Einstein notation, the repeated subscript τ indicates summation. Equation 10 consists of a Maclaurin expansion that uses as base the 2D polynomials $x^i\ z^j$, where i and j are positive integers. The maximum expansion order, N, is supposed to be 10. For example, the second-order displacement field is:

$$\begin{aligned}u_x &= u_{x1} + xu_{x2} + zu_{x3} + x^2u_{x4} + xzu_{x5} + z^2u_{x6} \\ u_y &= u_{y1} + xu_{y2} + zu_{y3} + x^2u_{y4} + xzu_{y5} + z^2u_{y6} \\ u_z &= u_{z1} + xu_{z2} + zu_{z3} + x^2u_{z4} + xzu_{z5} + z^2u_{z6}\end{aligned} \tag{11}$$

were u_{x_τ}, u_{y_τ}, and u_{z_τ} are y-dependent.

By introduction the shape function N_i, and the nodal displacement vector, $\boldsymbol{q}_{\tau i}$

$$\boldsymbol{q}_{\tau i} = \{q_{u_{x_{\tau i}}} q_{u_{y_{\tau i}}} q_{u_{x_{\tau i}}}\}^T \tag{12}$$

the displacement vector becomes:

$$\boldsymbol{u}_\tau(x,y,z) = N_i(y)F_\tau(x,z)\boldsymbol{q}_{\tau i} \tag{13}$$

For sake of brevity, the shape function are not reported here. They can be found in [20]. Elements with 2, 3, and 4 nodes are formulated. These elements are addressed as B2, B3, and B4 respectively. The stiffness and the mass matrices of the elements and the external loadings that are coherent to the models are obtained via the Principle of Virtual Displacements :

$$\partial L_{int} = \int_V (\partial\epsilon_p^T\sigma_p + \partial\epsilon_n^T\sigma_n)dV = \partial L_{ext} - \partial L_{ine} \tag{14}$$

where the L_{int} stands for the strain energy, L_{ext} is the work of the external loadings, and L_ine is the work of the inertial loadings. The virtual variation of the strain energy is rewritten using Eq.s 3, 6, and 15, in a compact form becomes:

$$\partial L_{int} = \partial q_{\tau i}^T K^{ij\tau s} q_{sj} \tag{15}$$

Where $K^{ij\tau s}$ is the stiffness matrix in the form of the fundamental nucleus.

Its is components are:

$$
\begin{aligned}
K_{xx}^{ij\tau s} = {} & \tilde{C}_{22}\int_\Omega F_{\tau,x}F_{s,x}d\Omega\int_l N_iN_jdy + \tilde{C}_{66}\int_\Omega F_{\tau,z}F_{s,z}d\Omega\int_l N_iN_jdy + \\
& \tilde{C}_{44}\int_\Omega F_\tau F_s d\Omega\int_l N_{i,y}N_{j,y}dy \\
K_{xy}^{ij\tau s} = {} & \tilde{C}_{23}\int_\Omega F_{\tau,x}F_s d\Omega\int_l N_iN_{j,y}dy + \tilde{C}_{44}\int_\Omega F_\tau F_{s,x}d\Omega\int_l N_{i,y}N_j dy \\
K_{xz}^{ij\tau s} = {} & \tilde{C}_{12}\int_\Omega F_{\tau,x}F_{s,z}d\Omega\int_l N_iN_jdy + \tilde{C}_{66}\int_\Omega F_{\tau,z}F_{s,x}d\Omega\int_l N_iN_jdy \\
K_{yx}^{ij\tau s} = {} & \tilde{C}_{44}\int_\Omega F_{\tau,x}F_s d\Omega\int_l N_iN_{j,y}dy + \tilde{C}_{23}\int_\Omega F_\tau F_{s,x}d\Omega\int_l N_{i,y}N_j dy \\
K_{yy}^{ij\tau s} = {} & \tilde{C}_{55}\int_\Omega F_{\tau,z}F_{s,z}d\Omega\int_l N_iN_jdy + \tilde{C}_{44}\int_\Omega F_{\tau,x}F_{s,x}d\Omega\int_l N_iN_jdy + \\
& \tilde{C}_{33}\int_\Omega F_\tau F_s d\Omega\int_l N_{i,y}N_{j,y}dy \\
K_{yz}^{ij\tau s} = {} & \tilde{C}_{55}\int_\Omega F_{\tau,z}F_s d\Omega\int_l N_iN_{j,y}dy + \tilde{C}_{13}\int_\Omega F_\tau F_{s,z}d\Omega\int_l N_{i,y}N_j dy \\
K_{zx}^{ij\tau s} = {} & \tilde{C}_{12}\int_\Omega F_{\tau,z}F_{s,x}d\Omega\int_l N_iN_jdy + \tilde{C}_{66}\int_\Omega F_{\tau,x}F_{s,z}d\Omega\int_l N_iN_jdy \\
K_{zy}^{ij\tau s} = {} & \tilde{C}_{13}\int_\Omega F_{\tau,z}F_s d\Omega\int_l N_iN_{j,y}dy + \tilde{C}_{55}\int_\Omega F_\tau F_{s,z}d\Omega\int_l N_{i,y}N_j dy \\
K_{zz}^{ij\tau s} = {} & \tilde{C}_{11}\int_\Omega F_{\tau,z}F_{s,z}d\Omega\int_l N_iN_jdy + \tilde{C}_{66}\int_\Omega F_{\tau,x}F_{s,x}d\Omega\int_l N_iN_jdy + \\
& \tilde{C}_{55}\int_\Omega F_\tau F_s d\Omega\int_l N_{i,y}N_{j,y}dy
\end{aligned}
\tag{16}
$$

It should be noted that no assumptions on the approximation order have been done. It is therefore possible to obtain refined beam models without changing the formal expression of the nucleus components. This is the key-point of CUF which permits, with only nine FORTRAN statements, to implement any-order beam theories. The shear locking is corrected through the selective integration, see [20]. For the sake of brevity, the derivation of the loadings vector variationally coherent to the model is not reported, it can be found in [15].

3 Results and Discussion

Several static analysis of different beam model are presented in thi section. The preliminary analysis are two assessments where the results are compared with results form bibliography. In the rest of work a wing configuration is analyzed.

3.1 Compact section

A rettangular cantilever beam is considered. The geometry is shown in figure 2. The height h is

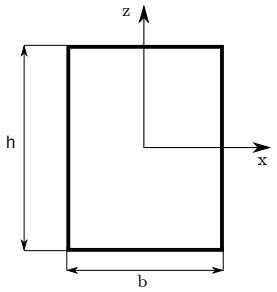


Figure 2: Rectangular cantilever beam section

assumed to be equal to 1[m], the ratio b/h is assumed to be equal to 0.5, the aspect ratio L/h is assumed to be 6, where L is the length of the beam. A concentrated load in z-direction, Fz is considered in [0, L, 0] and are equal to 1 [N]. An orthotropic material is considered, where E_{11} = 206.80 GPa, $E_{22} = E_{33}$ = 5,17 GPa, $G_{12} = G_{13}$ = 3,10 GPa, G_{23} = 2,55 GPa and $\nu_{12} = \nu_{13} = \nu_{23}$ = 0.25. A comparison with a 3D-FEM solution is given. Fig. 3 shows the value of the normal tension σ_{yy} in the area close to the constraint for a beam subject to bending load. The diagram show that the results from a third-order beam model agree with those from 3D FE model.

3.2 Thin Walled Section

A cantilevered C-shaped cross-section beam is considered to evaluate the capabilities of the present beam model in the thin-walled structures analysis (see Carrera et.al. [21]). The geometric characteristics have been taken from [22]. Fig. 4 shows the cross-section geometry, the height h is considered equal to 5 [m], b is equal to 3.5 [m] and the thickness, t, is assumed to be 0.2 [m], the length of the beam, L, is 18 [m]. Young's

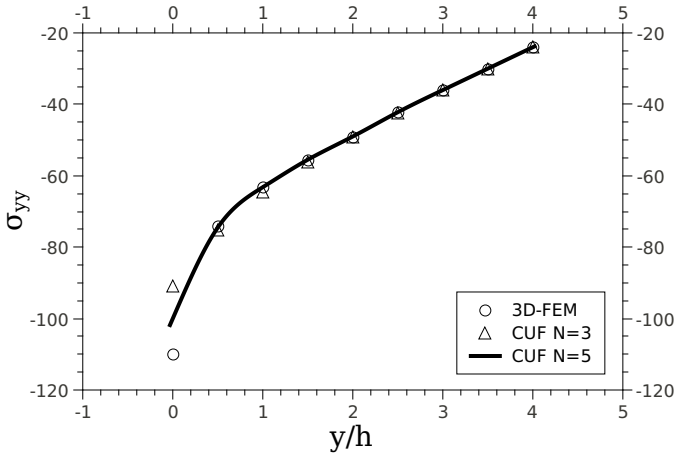


Figure 3: End-effects due to the boundary conditions on a cantilevered beam with rectangular compact section. Comparison of the results from CUF beam theories and 3D FEM mode.

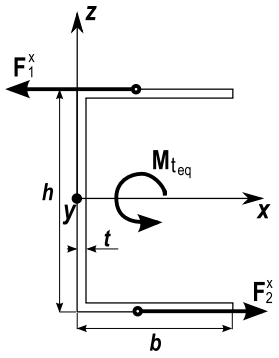


Figure 4: C-shaped cross-section geometry and loading.

modulus, E, is equal to 30 [GPa], and the Poisson ratio, ν is equal to 0.33. Two forces, F_1^x and F_2^x (see Fig. 4), equal to 200 [kN] are posed in [1.095, $\pm h/2$, L] acting in $\pm x$-direction in order to reproduce an equivalent torsional moment, $M_{t_{eq}}$, equal to 1000 [kNm]. Ten elements B4 are used in the axial discretization. Figure 5 shows the deformed cross-section at beam tip evaluated by means of different structural model, the displacements have been scaled by a $\times$100 factor to make the figure clearer. The beam model is able to capture the section deformation with an accuracy close to the solid model.

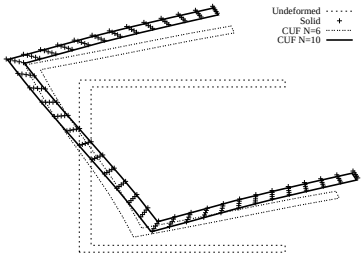


Figure 5: Deformation of C-shaped cross-section beam at y=L for different structural models, displacements amplification x100.

3.3 Airfoil



Figure 6: Cross-sections used for the wing configurations.

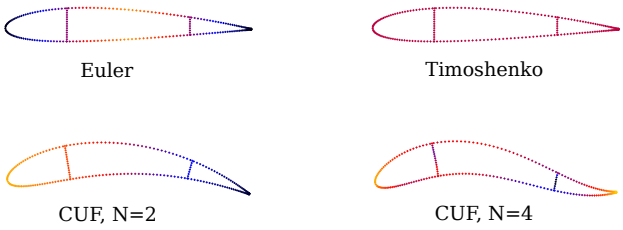


Figure 7: Cross-sections deformation for different beam model.

A conventional wing configuration is analyzed and the results are presented here. The wing is subject to internal forces in order to simulate an actuation system able to modify the airfoil camber. A concentrated load in z-direction, Fz is considered in [0, L, 0] in order to simulate the lift effect. Cantilever boundary conditions is accounted for. An aluminium wing is considered (Young's modulus $E = 69$ [GPa] and Poisson's ratio $\nu = 0.33$) The wing shape examined in this work are summarized in Table 1 The beam-like structures are considered to have an airfoil-shaped section. The NACA 2415 airfoil is adopted as cross-section profile, which is subdivided into three cells. The cells are obtained by inserting two spars along the span-wise direction at 25 % and 75 % of the chord length, see Fig. 6. Their thicknesses are, respectively, 10% and 7% of the maximum airfoil thickness, whereas the percentage is about 4% for the skin. In Fig. 7 are reported the cross section deformation for different beam models. It is possible to see that the conventional beam model, Euler and Timoshenko, are not able to capture the cross section deformation. Using high order beam model it is possible to simulate the cross section deformation.

Table 1: Wing configuration adopted to discuss the results.

Name	Value
Λ	0
Γ	0
λ	1
c_{root} $[m]$	1
c_{tip} $[m]$	1
$\bar{c}$ $[m]$	1
L $[m]$	5
$l/\bar{c}$	5
b $[m]$	5
S_w $[m^2]$	10
AR	10

4 Conclusion

In the field of aeronautics, shape morphing has been used to identify those aircraft that undergo certain geometrical changes to enhance or adapt to their mission profile. Much of the recent work in the area of reconfigurable airfoil structures has focused on locally deforming a portion of an airfoil to achieve some specific single design point goal. Typical applications include the actuation of leading and trailing edge devices as well as small-scale surface deflections to increase performance in a specific flow regime. In this work a 1-D FEM model able to simulate local deformation of portion of an airfoil has been presented. From results it is possible to see that the presented beam element is able to capture local effects that can't be simulated with classical beam model. The shell-like capabilities of the presented model allow to use beam model for the study of smart wing.

The present work has been performed through a close cooperation with Alenia Aeronautica S.p.A and through the use of Confidential Information and Data, property of Alenia Aeronautica S.p.A. which remains the sole owner of all such relevant IP rights. The results of the present work shall be, therefore, property of Alenia Aeronautica S.p.A. Authors wish to thank G.M.Carossa, E.Baldassin

References

[1] G. Reich, B. Sanders, Introduction to morphing aircraft research., Journal of Aircraft 44 (4).

[2] L. Euler, De curvis elasticis, Lausanne and Geneva: Bousquet.

[3] A. de Saint-Venant, Mémoire sur la flexion des prismes, sur les glissements transversaux et longitudinaux qui l accompagnent lorsquelle ne sopére pas uniformément ou en arc de cercle, et sur la forme courbe affectée alors par leurs sections transversales primitivement planes, Journal de Mathmatiques pures et appliques 1 (1856) 89–189.

[4] A. de Saint-Venant, Mémoire sur la torsion des prismes, avec des considérations sur leur flexion, ainsi que sur léquilibre interieur des solides élastiques en général, et des formules pratiques pour le calcul de leur résistance divers efforts sexercant simultanément, Académie des Sciences de l'Institut Impérial de Frances 14 (1856) 233–560.

[5] S. P. Timoshenko, On the corrections for shear of the differential equation for transverse vibrations of prismatic bars, Philosophical Magazine 41 (1921) 744–746.

[6] S. P. Timoshenko, On the transverse vibrations of bars of uniform cross section., Philosophical Magazine 43 (1922) 125–131.

[7] D. T. Mucichescu, Bounds for stiffness of prismatic beams, Journal of Structural Engineering 110 (1984) 1410–1414.

[8] K. Washizu, Variational methods in elasticity and plasticity., Oxford: Pergamon Press.

[9] W. Kanok-Nukulchai, Y. S. Shin., Versatile and improved higher-order beam elements., Journal of Structural Engineering 110 (1984) 2234–2249.

[10] E. Carrera, A class of two dimensional theories for multilayered plates analysis, Atti Accademia delle Scienze di Torino, Memorie Scienze Fisiche 19-20 (1995) 49–87.

[11] E. Carrera, Theories and finite elements for multilayered, anisotropic, composite plates and shells, Archives of Computational Methods in Engineering 9 (2) (2002) 87–140.

[12] E. Carrera, Theories and finite elements for multilayered plates and shells: a unified compact formulation with numerical assessment and benchmarking, Archives of Computational Methods in Engineering 10 (3) (2003) 216–296.

[13] E. Carrera, S. Brischetto, A. Robaldo, Variable kinematic model for the analysis of functionally graded material plates, AIAA Journal 46 (2008) 194–203.

[14] E. Carrera, G. Giunta, Refined beam theories based on a unified formulation, International Journal of Applied Mechanics 2 (1) (2010) 117–143.

[15] E. Carrera, G. Giunta, P. Nali, M. Petrolo, Refined beam elements with arbitrary cross-section geometries, Computers and Structures 88 (5–6) (2010) 283–293.

[16] E. Carrera, M. Petrolo, P. Nali, Unified formulation applied to free vibrations finite element analysis of beams with arbitrary section, Shock and Vibration 18 (3) (2011) 485–502.

[17] E. Carrera, M. Petrolo, A. Varello, Advanced beam formulations for free vibration analysis of conventional and joined wings, Journal of Aerospace EngineeringIn press. DOI: 10.1061/(ASCE)AS.1943-5525.0000130.

[18] E. Carrera, M. Petrolo, Analysis of slender, thin walled, composite made structures with refined 1d theories, in: 52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA Paper 2011-2168, Denver, CO, 2011.

[19] E. Carrera, M. Petrolo, On the effectiveness of higher-order terms in refined beam theories, Journal of Applied Mechanics 78 (2) (2010) 021013.1–021013.17.

[20] K. J. Bathe, Finite element procedure, 1996.

[21] E. Carrera, M. Petrolo, E. Zappino, Advanced elements for the static analysis of beams with compact and bridge-like sections, Submitted.

[22] V. Z. Vlasov, Thin walled elastic beams., Israel Program for Scienti

c Transaction.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Low Temperature Gas Generator Propellants

R. Puccio, F. Maggi, G. Colombo, and L. T. De Luca
Aerospace Engineering Dept., Politecnico di Milano
34, via La Masa, 20156 Milan, Italy

R. De Amicis
Consultant AVIO S.P.A.
00034 Colleferro (Rome), Italy

Keywords: *low temperature, gas generators, sodium azide based propellants, HTPB binder, fuse wires.*

Abstract

Low temperature propellants for gas generators are commonly used in automotive safety devices. The application of this technology to a spacecraft/launcher attitude control system is the aim of this research. Constraints required for gas generator propellants are in terms of burning rate and gas temperature. Sodium Azide-based formulations are studied with the addition of a metal-alkaline nitrate, ammonium perchlorate and HTPB.

A systematic experimental characterization of the tested propellants was conducted in a windowed strand burner with servo-controlled combustion pressure and a digital video camera: burn rates were measured at three different pressures by either an optical video technique or a standard fuse wires technique; temperature profiles were measured with Pt/Pt-Rh thermocouples (S-type). The obtained flame temperatures could further be decreased by adding cooling agents.

1 Introduction

Gas generators are mainly adopted in applications where high volumes of gas are required in a short time. Compacted powders or propellants grant high density and offer a convenient approach for potential gas storage in small volumes. A typical example of such technology is found in emergency devices such as air-bags, fire extinguishers, or kickstart charges for turbines. In some applications high temperature and toxicity of exhaust gases may represent a severe constraint. For this reason, low temperature and nontoxic gas generators were developed. A low-temperature gas generator is in general a propellant or a system of compressed pellets that release gases at relatively low temperature (less than 1200 K) with respect to standard propellants (say 2500-3000 K). For these formulations, heat feedback from the gas phase to the burning surface represents a critical point since the enthalpy for surface pyrolysis and condensed phase heating must be provided in order to get a self-sustained combustion [1].

There exist several patents available on this subject, most of them demonstrating the application of inorganic powders shaped in pellets. Reading through patents [2,3], some categories of powders are typically involved in the compositions if a certain level of toxicity is allowed.

- Oxidizers: oxygen-donors inorganic powders such as nitrates or perchlorates;

- Flame coolants: hydroxides, carbonates, hydrated hydroxides of some metals;
- Burning rate catalyst: a transition metal oxide or other oxide with known catalytic capabilities;
- Combustible powder: a metal in the group of boron, aluminum or zirconium.

Patents refer to these formulations having a flame temperature of about 1100 to 1500 K giving few indications about the composition. Most of gas generating compositions are based on alkali metal or alkaline earth metal azides with iron, nickel, cobalt or silica oxides. Cooling of flame may be necessary and can be pursued by adding a coolant in the mixture not exceeding the fraction of 35% by weight, or using cooling filters placed before the exhaust (Fig. 1). In some patent embodiments, the formation of a porous sinter during the combustion is used as a self-filtration of Condensed Combustion Products (CCPs).

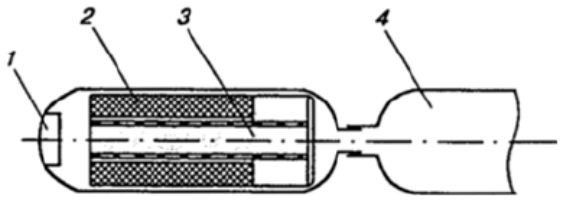


Fig. 1 Example of cold gas generator with cooling device: igniter (1); monolithic charge (2); chemical cooler (3); pressurized vessel (4) [1].

The application of such technology to a spacecraft/launcher attitude control system is the aim of this investigation. It requires burning rate around 4 mm/s at 20 bar, gas temperature less than 500 °C (773 K), and no CCPs. In particular for Launch Vehicles, based on Solid Rocket Motors (SRM) propulsion, possible roll can be produced during 1st, 2nd and 3rd stage SRM firings, higher than acceptable levels. To reduce these rolls, simple subsystems can be envisaged to be embarked, with few changes, on the relevant stages of the baseline Launch Vehicle and with the possibility to disembark if, after some flights, the detected roll levels result at acceptable values. To this end, simple solutions are under analysis based on the use of a Gas Generator with solid propellant, a three way flow distribution valve, able to pipe the generated gas to fixed nozzle thrusters in order to create three different control effects: no overall torque, clockwise torque and anti-clockwise torque.

2 Preliminary Formulations

Preliminary tests have been carried out with conventional rocket propellant ingredients. All fuel samples are based on HTPB as binder/fuel, in most cases coupled with Ammonium Perchlorate (AP) as oxidizer. Some compositions have been doped with Calcium Carbonate as coolant or with Iron Oxide as burning rate catalyst. The compositions under examinations are summed up in Table 1.

Table 1 Analyzed propellants compositions. Adiabatic flame temperatures calculated by thermochemical computations.

	Binder	Oxidizer	Coolant	Catalyst	Adiabatic Flame Temperature, [K]
Propellant 1	20% HTPB	30% AP 30% AN	20% $CaCO_3$	-	1211
Propellant 2	20% HTPB	80% $KMnO_4$	-	-	-
Propellant 3	20% HTPB	80% KNO_3	-	-	1592
Propellant 4	20% HTPB	80% $NaNO_3$	-	-	1836
Propellant 5	50% HTPB	50% AP	-	-	1190
Propellant 6	40% HTPB	60% AP	-	-	1242
Propellant 7	50% HTPB	49.5% AP	-	0.5% Fe_2O_3	1193
Propellant 8	40% HTPB	59.5% AP	-	0.5% Fe_2O_3	1245
Propellant 9	45% HTPB	55% AP	-	-	1215
Propellant 10	45% HTPB	54.5% AP	-	0.5% Fe_2O_3	1218
Propellant 11	40% HTPB	50% AP	10% $CaCO_3$	-	1183
Propellant 12	40% HTPB	49.5% AP	10% $CaCO_3$	0.5% Fe_2O_3	1186

Binder was composed by HTPB, DOA as plasticizer, and cured with IPDI. The catalytic element was Tin based. Propellants strands of approximately 10 grams have been produced.

2.1 Thermochemical Calculations

Through thermochemical computations the best possible compromise between quality of combustion and reduced flame temperature for

standard composite propellants has been analyzed (Fig. 2). If the presence of CCPs is considered a selection criterium, the optimal AP mass fraction is at about 70%. At this percentage the flame temperature is still too high. The reduction of the AP mass fraction is possible, subjected to the experimental check whether the flame is self-sustained or not.

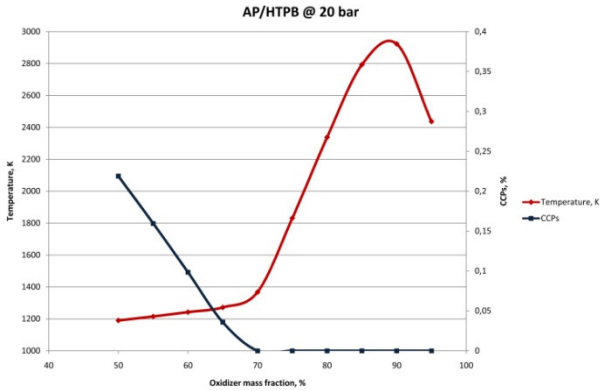


Fig. 2 Adiabatic flame temperature and CCPs mass fraction for AP/HTPB propellants at 20 bar.

The behavior of adiabatic flame temperature is evaluated adding other oxidizers with less oxidizing power than the AP ones. A common cooling agent like $CaCO_3$ is also tested (Fig. 3).

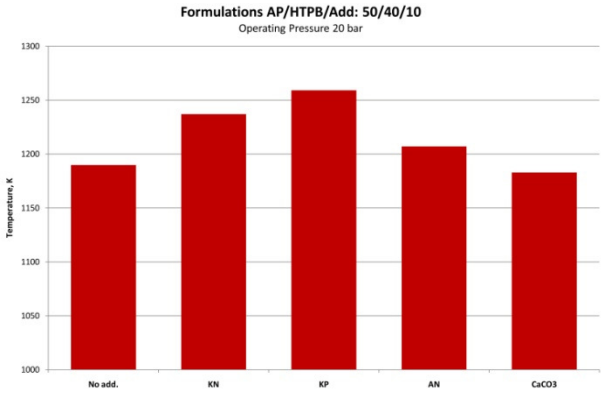


Fig. 3 AP/HTPB formulations with 10% of additive. Data refer to adiabatic flame temperature compared to AP/HTPB-50/50 formulation.

2.2 Experimental Results

Not all of the propellants prepared (Table 1) showed self-sustained combustion at 20 bar. Most of them did not support flame after detachment of the ignition wire from the burning surface, and after quenching left a great amount of residuals, in terms of fine black powders or sometimes sticky liquid. In a few cases, a massive residual skeleton was observed. An overall results summary is shown in Table 2.

Table 2 Summary of experimental results.

	Self-sustained combustion at 20 bar	Burn rate at 20 bar, [mm/s]	Ballistic exponent	Residuals
Propellant 1	No	-	-	No
Propellant 2	Yes	-	-	Yes
Propellant 3	No	-	-	No
Propellant 4	No	-	-	No
Propellant 5	No	-	-	Yes
Propellant 6	Yes	1.05	0.31	Yes
Propellant 7	No	-	-	No
Propellant 8	Yes	3.6	0.31	Yes
Propellant 9	No	-	-	No
Propellant 10	Yes	3.05	0.23	Yes
Propellant 11	No	-	-	No
Propellant 12	Yes	1.12	-	Yes

3 Sodium Azide-based Propellants

3.1 Preliminary SA-based Formulations

A wide literature on air-bag gas generators and cold propellants refer to Sodium Azide (SA) as the best compound for the scope [4,5,6]. So early experiments were addressed to evaluate the compatibility of SA with HTPB binder and various oxidizers. Based on these tests a preliminary campaign of SA and HTPB propellants was carried out. All samples prepared have a mass of 5 grams. Their composition comprises a quantity of HTPB in the range of 10-25%, and SA in the range of 60-75%. An addition of Potassium Nitrate (KN) or cooling agent Lithium Fluoride (LiF) in the range of 10-20% was also considered. The experimental characterization of samples took place in a windowed strand burner with servo-controlled combustion pressure and a digital video camera to acquire data. The combustion of these formulations under inert atmosphere was not satisfactory under both 20 and 5 bar of pressure. The propellant ignition, made by an

incandescent wire, was difficult, and the combustion was not self-sustained. Other attempts were made under atmospheric conditions, but results were not much different. Additional samples containing alkaline metal oxide, showed the same behavior of previous formulations under inert atmosphere. However, ignition under atmospheric conditions allowed to measure burning rate. The combustion is self-sustained, and the velocity measured is ~0.5 mm/s consistent with results in [4]. A massive residual skeleton remains after the combustion, about 45% of the total mass.

3.2 SA Thermochemical Computations

First of all the behavior of SA and HTPB compound at different percentages were investigated.

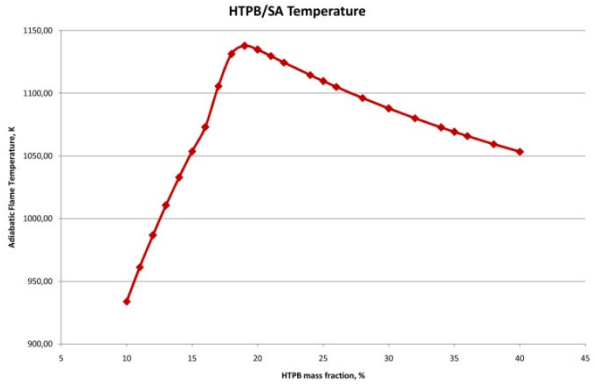


Fig. 4 Theoretical adiabatic flame temperature of SA and HTPB formulations vs. HTPB mass fraction at 20 bar.

What can be noticed from Fig. 4 is that SA, with low HTPB mass fraction (less than 20%), reacts with all carbon of the binder; such reaction increases the adiabatic flame temperature from 934 K to 1135 K. With more than 20% HTPB the temperature decreases up to 1053 K at 40% HTPB. In the range of 20-40% HTPB a lot of carbon remains in the combustion products as graphite instead reacting with SA (to form sodium cyanide). Thus, an amount not larger than 15% HTPB should be used in order to have an initial temperature less than 1050 K.

In order to achieve lower temperatures a coolant ingredient can be used. So different coolants for SA-based formulations were considered, such as $CaCO_3$, SiO_2 and LiF.

Calcium carbonate is a common coolant in aerospace applications.

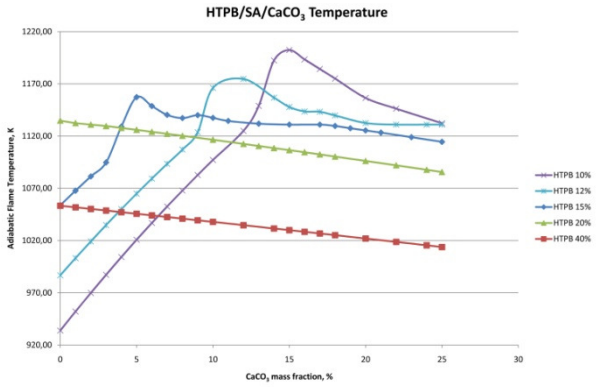


Fig. 5 Theoretical adiabatic flame temperature of SA-based propellant with $CaCO_3$ coolant, at different HTPB percentages, at 20 bar.

As can be seen in Fig. 5, with less than 20% HTPB the behavior of $CaCO_3$ is like an oxidizer, so it increases rapidly the flame temperature. To have a real coolant effect on the propellant with $CaCO_3$, the quantity of HTPB must be above 20% of total mass. However temperature decrease rate is very low, say about ~2 K for each 1% of $CaCO_3$. Therefore, the calcium carbonate does not seem to be the best choice for the target temperature.

Another possible coolant is the silicon dioxide, commonly used with SA and KN to form alkaline silicate. In fact silicon dioxide reacts producing a sintered, coherent combustion residue which is easy to filter.

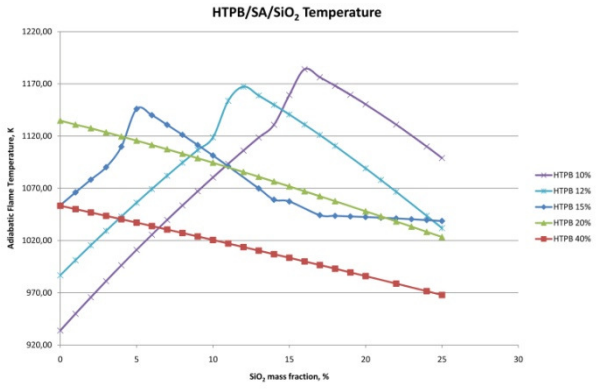


Fig. 6 Theoretical adiabatic flame temperature of SA-based propellant with SiO_2 coolant, at different HTPB percentages, at 20 bar.

Like calcium carbonate, for percentages lower than 20%, silicon dioxide (Fig. 6) acts as an oxidizers so that the flame temperature increases. Differently from calcium carbonate, the temperature decrease rate with silicon dioxide is appreciable. In fact with the 25% silicon dioxide the temperature can decrease up to 100 K from the initial value. The choice of silicon dioxide seems useful if the quantity of HTPB is larger than 40% so that adiabatic flame temperature could decrease below the 1000 K. This high percentage of HTPB could represent a problem for the ignition of the propellant. Despite that, the silicon dioxide is useful to sinter the solid combustion products, so that is simple to filter.

As seen in [1], lithium fluoride is a good coolant too. In particular, as can be seen in Fig. 7, even for low percentages of HTPB (less than 20%), the effect of lithium fluoride is to reduce flame temperature. With only 10%-12% of HTPB the flame temperature, at any lithium fluoride percentages, remains quite below the 1000 K limit. A severe problem with this type of coolant are the difficulties encountered in the combustion of the propellant. In early experimental tests it was been seen that a compound with more than 15% of LiF does not ignite under inert atmosphere (N_2) at any pressure value.

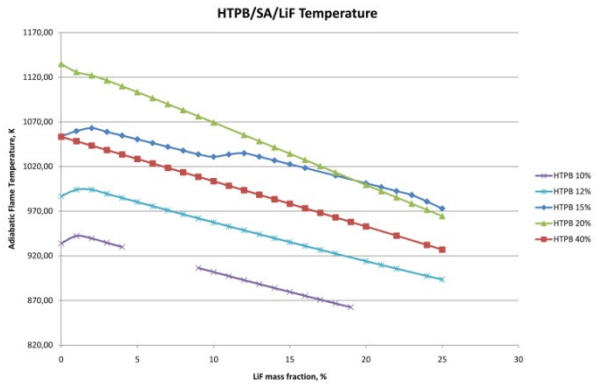


Fig. 7 Theoretical adiabatic flame temperature of SA-based propellant with LiF coolant, at different HTPB percentages, at 20 bar.

3.2.1 Ammonium Perchlorate Additive

Next step was to identify the minimum quantity of AP to ensure propellant ignition and self-sustained combustion at 20 bar in inert environment. The new formulations investigated must also decrease the adiabatic flame temperature at acceptable values. Through thermochemical computations a heavy decrement in adiabatic flame temperature is reached reducing the AP mass fraction to 0.5%. At this value the temperature is set to 1025 K. In Fig. 8 flame temperature and CCPs mass fraction vs. AP percentage are shown.

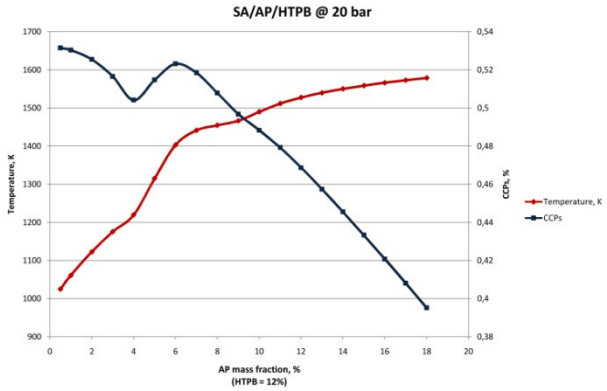


Fig. 8 Adiabatic flame temperature and CCPs mass fraction for SA/AP/HTPB formulations, for different AP mass fractions, at 20 bar.

Now the previous formulations with silicon dioxide and lithium fluoride coolant are considered with a small quantity of AP (0.5%). In Section 4 some compositions considered with AP have a minimum quantity of Fe_2O_3 below 1%. At this mass fraction the adiabatic flame temperature shown in Fig. 9-10 increase at most by ~29 K.

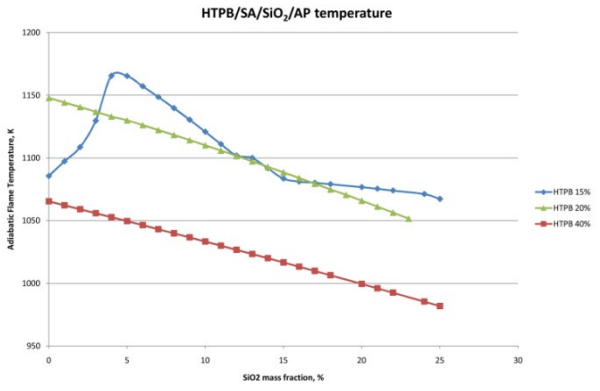


Fig. 9 Theoretical adiabatic flame temperature of SA-based propellants with SiO_2 mass fraction. Different HTPB percentages involved and 0.5% AP, at 20 bar.

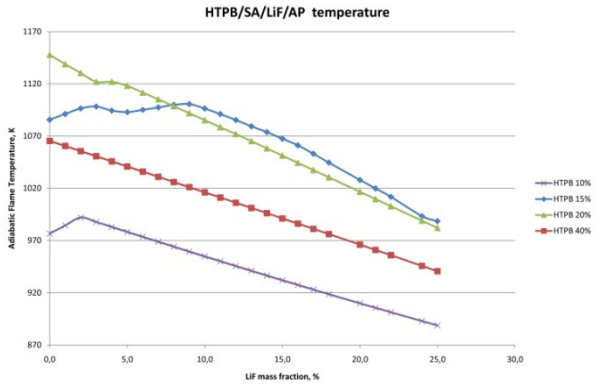


Fig. 10 Theoretical adiabatic flame temperature of SA-based propellants with LiF mass fraction. Different HTPB percentages involved and 0.5% AP, at 20 bar.

4 Experimental Results

SA-based formulations considered during tests are summarized in Table 3.

Table 3 Analyzed SA-based propellant formulations. Adiabatic flame temperature calculated by thermochemical computations.

	Binder	Fuel/Oxidizer	Coolant	Catalyst	Adiabatic Flame Temperature, [K]
PSA-1a	12% HTPB	70% SA 18% AP	-	-	1579
PSA-2b	10% HTPB	85% SA 2% AP	-	3% Fe_2O_3	1119
PSA-3c	10% HTPB	87% SA 2.5% AP	-	0.5% Fe_2O_3	1121
PSA-4a	10% HTPB	85% SA 0.75% AP	4% LiF	0.25% Fe_2O_3	1013

4.1 Burn Rate Measurements

For the first formulation PSA-1a (propellant comprising AP in the range of 20-15%) the burning rate measured at 20 bar is ~13 mm/s; for the second, with AP 5-10 µm, is ~12 mm/s (Fig. 11). These results demonstrate too high burn rate for the proposed target, and also a high ballistic exponent (Fig. 12). Anyway they are in accordance with AP propellant combustion properties. So it is necessary to reduce the quantity of AP at the minimum to ensure ignition. From CEA's calculations (see Section 3.2.1), the adiabatic flame temperature is set to ~1579 K. Obviously this temperature is too high, and the reason of that is mainly the quantity of AP.



Fig. 11 Some frames of SA-based propellants with AP (5-10 µm) combustion at 20 bar, at 50fps.

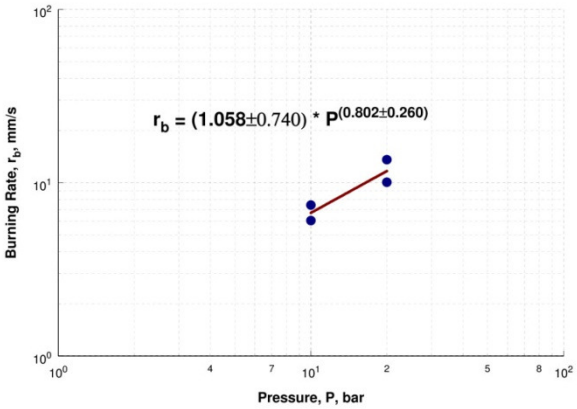


Fig. 12 Vielle's law rapresentation for PSA-1a propellant, with AP 5-10 µm.

Formulations with AP percentage less than 3% have also been investigated. As seen in Section 3.2.1 such reduction of AP quantity allows a decrement of temperature of at least 400 K. In order to maintain an adequate burning rate a small quantity of Iron Oxide is also added. Two different percentages are considered: the first propellant with percentage less than 5% (PSA-2b), and the second less than 1% (PSA-3c). The first propellant presents poor quality of combustion in air, mainly because the presence of too much Nanocat Iron Oxide and its fine particle size ~0.003 microns. The second (PSA-3c) has a better behavior, even in inert environment. The adiabatic flame temperature of the first propellant is 1118 K, for the second is 1121 K. For the latter propellant combustion products are theoretically 49% condensed type and 51% gaseous state. The combustion of PSA-3c propellant samples show an high quantity of smoke that does not allow to measure the burning rate in a conventional way. A further improvements required the addition of some coolant ingredients, in particular Lithium Fluoride that guarantee a noticeable decrement of temperature (see Section 3.2.1). In PSA-4a

propellant a percentage of 4% LiF is considered. Experimental burning rate results are obtained using the conventional fuse wire technique. Burn rate measured at 20 bar is ~2.2 mm/s and ~1.7 mm/s at 10 bar. In Figure 13 the Vielle's law is represented.

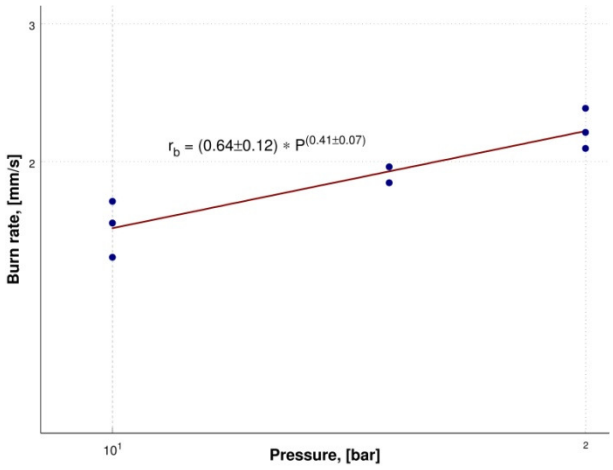


Fig. 13 Vielle's law rapresentation for PSA-4a propellant.

4.2 Temperature Measurements

In order to evaluate gas temperature, experimental tests with Pt-Pt/Rh thermocouples have been done. The thermocouple is placed ~5 mm in front of the sample burning surface. Pressure in strand burner is set to maximum 20 bar, due to chamber limitation. Propellant PSA-1a has an adiabatic flame temperature, computed with CEA software, of ~1578 K.

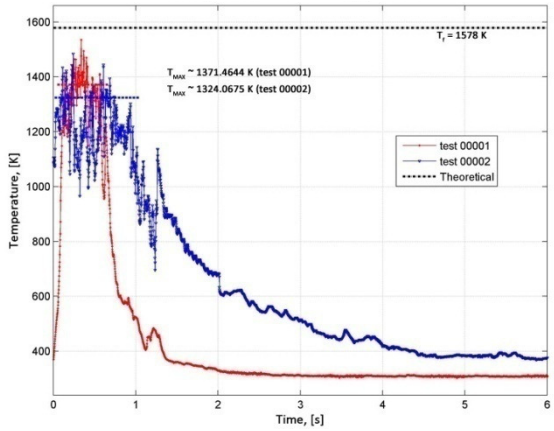


Fig. 14 Gas temperature measured with Pt-Pt/Rh thermocouple for PSA-1a, at 15 bar.

Figure 14 shows the experimental temperature profiles for the two tests done with PSA-1a propellant.

PSA-3c propellant has an adiabatic flame temperature of 1121 K. In Figure 15 the temperature profiles are represented for the two tests done in comparison with the third test with the thermocouple inserted in the propellant.

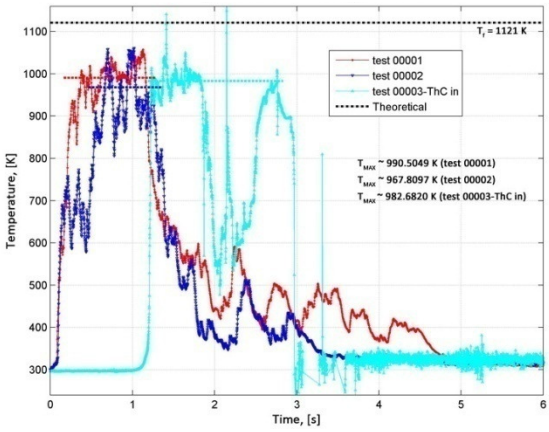


Fig. 15 Gas temperature measured with Pt-Pt/Rh thermocouple for PSA-3c formulation, at 15 bar.

Propellant PSA-4a, with 4% LiF, demonstrate a lower gas temperature of ~850 K without laser radiation, in spite of the theoretical 1013 K.

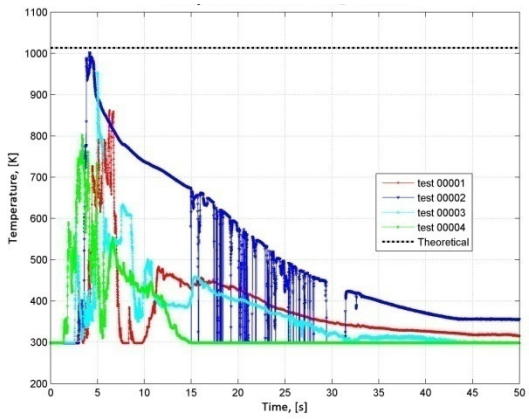


Fig. 16 Gas temperature measured with Pt-Pt/Rh thermocouple for PSA-4a, at 10 bar.

5 Conclusions and Future Developments

In this article conventional rocket propellants are shown to be able to achieve a minimum temperature of ~1200 K (adiabatic flame

temperature). SA and coolant ingredients allow reaching temperatures below 1000 K. Moreover, experimental results point out that the measured temperatures are lower than the theoretical ones by at least 100 K, typically. With propellant PSA-3c the temperature of ~990 K is reached without any coolant. With propellant PSA-4a a lower temperature of ~850 K, very closed to the target temperature of 800 K, has been acquired.

Another aspect to be taken into account is the formation of solid combustion residuals that could damage the pipelines and distribution valves of a potential attitude control system. For the formulations analyzed in this article a solid residuals is always obtained especially with low flame temperature compositions, but the residuals could be simply removed using a filtering system.

Future developments will focus on the analysis of a chemical filtering system that fulfil the double function of solid residual filter and cooling filter. Moreover SA, highly toxic, may be replaced with some other non toxic ingredient.

Acknowledgments

Financial support by Avio Space Propulsion Division, Italy is gratefully acknowledged. Nanocat was courteously provided by MACH I, USA. Special thanks are due to Dr. A. Bandera and Ms. E. Marchesi for setting the exploratory formulations and carrying out the initial ballistic tests.

References

[1] Shandakov V.A., Puzanov V.N., Komarov V.F., and Borochkin V.P., “Method of generating cold gases in solid–fuel gas generators”, *Combustion, Explosion, and Shock Waves*, Vol. 35, No. 4, July–August 1999, pp. 75–78.

[2] Garner E.F., “Low temperature gas generator propellant”, *U.S.Patent 3,912,562*, October 1975.

[3] Blomquist H.R., “Cool burning gas generating material for a vehicle occupant protection apparatus”, *U.S.Patent 6,875,295 B2*, April 2005.

[4] DiBiaso H.H., English B.A., and Allen M.G., “Solid–phase conductive fuels for chemical microactuators”, *Sensors and Actuators A: Phys.*, No. 111, 2004, pp. 260–266.

[5] Mario A. DiValentin, “Metal oxide/azide gas generating compositions”, *U.S.Patent 3,996,079*, December 1976.

[6] George W. Goetz, “Azide gas generating composition”, *U.S.Patent 4,062,708*, December 1977.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Atmospheric modeling by SCRAT experiment during the 2010 ESA BEXUS 10 Flight Campaign

G. Rodeghiero and L. Olivieri
CISAS University of Padova, Italy

A. Francesconi
CISAS University of Padova, Italy

Keywords: *stratosphere, oscillations, attitude*

Abstract

This paper discusses the atmospheric analyses that were carried out using pressure and temperature data collected during the flight of the SCRAT experiment onboard the BEXUS 10 balloon. Here are underlined the vast perspectives in terms of scientific data evaluation, analysis and modeling that can be retrieved by the use of simple commercial temperature and pressure sensors. An environmental model accounting for radiative and convective energy transfer based on the temperature and pressure data has been developed. The diurnal effects on the stratospheric temperature and other related physical quantities as the lapse rate and the water vapor content have been derived. As a by product, the onboard temperature sensor used as a Sun sensor can infer many information on the attitude of the balloon by the study of the temperature oscillations induced by the balloon's rotation during the floating phase.

1 Introduction

SCRAT (*Spherical Compact Rechargeable Air Thruster)* is a low-thrust cold gas actuator, ejecting gas through a nozzle at high speed in order to create force. The aim of SCRAT was the development and testing of the actuator on the stratospheric platform provided by BEXUS balloon. As a by-product SCRAT allowed to measure the pressure and temperature of the atmosphere and to study the related physical quantities as the density, the atmospheric lapse rate and the water vapor content. In addition, from the analysis of the temperature spectrum, a direct correlation between the temperature and the attitude dynamics of the balloon was figured out. The framework in which the experiment has been designed is the REXUS/BEXUS Programme [1], supported by the European Space Agency, the Swedish Space Corporation and DLR. It provides a unique opportunity to gain experience in space field and related researches by flying an experiment on a sounding rocket (REXUS) or a stratospheric balloon (BEXUS) along spatial and sub-spatial trajectories. The project is addressed to groups of students working together for a year to design, develop and test an experiment in a real space mission.

2 Background and motivations

Balloons can explore a wide range of altitudes (15m – 45 km), and thanks to their variable size,

between 3000 and 1.200.000 m^3, they can lift payloads up to 3000 kg for a time flight spanning from a few hours to a few weeks. The main studies concern astronomy, since the instrument can work above the disturbing atmosphere, planetary and atmospheric physics, microgravity by dropping off payloads from the gondola where there is a large free landing area and the test of new aerospace and re-entry vehicles. Since the first decades of the last century a lot of stratospheric balloons were launched from many different locations, but they remain still today a suitable platform to study the atmosphere, the climate and the pollution of the Earth as evidenced by the important results from missions like *Concordiasi* [2] and programs like the "*Programme Nationale de Chimie Atmospherique*" [3]. BEXUS 10 was launched from the ESRANGE base (67°N Lat, 20° E Lon), about 200 km above the Arctic polar cycle. Typically a polar flight can explore a deeper region of the stratosphere since the gravitational force experimented near the poles is greater and compresses the atmospheric structure into more flattened shell with respect to the lower latitudes. The study and the characterization of the atmospheric environment is a crucial topic also for the past and future Solar System missions. The great interest of the Space Agencies towards Mars, with missions as the *Mars Exploration Rover Mission* [4], and Titan, with the *Huygens* Probe [5], together with investigations by stratospheric balloons and sounding rockets have led to achieve key technologies and experience for the design of future missions as *Curiosity* [6], *ExoMars* [7] and many others.

3 BEXUS 10 Balloon

As given in Fig. 1 the BEXUS 10 balloon was essentially composed of four parts: the balloon, the re-entry parachute, the EBASS unit and the gondola hosting the scientific payload. The EBASS unit was responsible for: altitude control, flight termination, load cell controlled emergency termination, onboard GPS and housekeeping data. The stratospheric balloon fully inflated had a volume of 12.000 m^3 and lifted 263 kg (total weight of gondola, EBASS and parachute) at a mean altitude of 25 km; the flight train was equipped with a 120 m^2 area parachute to control and decelerate the gondola during the descent phase into the atmosphere, the total length of the flight train was about 138 m.

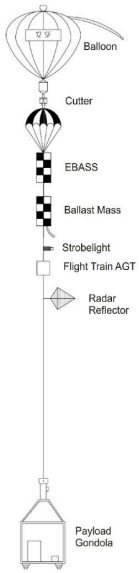


Fig. 1 BEXUS 10 flight train, photo courtesy SSC [8]

3.1 BEXUS 10 Flight Campaign

The balloon was launched at 1:07 UT the 9th of October, 2010 from the Esrange Space Center in Sweden.

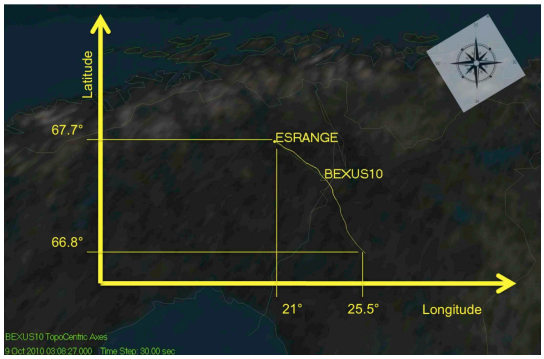


Fig. 2 BEXUS 10 trajectory above Sweden and Finland

After the launch, the balloon moved upward at a mean velocity of 5 m/s reaching the top of its flight (25 km) in 1 h 15 min; the telecommand for balloon cut-off and re-entry phase was sent 4h 30 min after the lift-off. The gondola landed

in a forest of Finland, 515 km in the South-East direction with respect to Esrange as displayed in Fig. 2.

4 Sensor layout

To achieve the goal of the atmospheric study, SCRAT was equipped with a pressure and temperature sensor, *Series 26PCDFA6D* and *PT100 Series DA-082* [9] respectively, measuring the static pressure and temperature with a sampling rate of 100S/s. The layout of the environmental sensors is reported in Fig. 3.

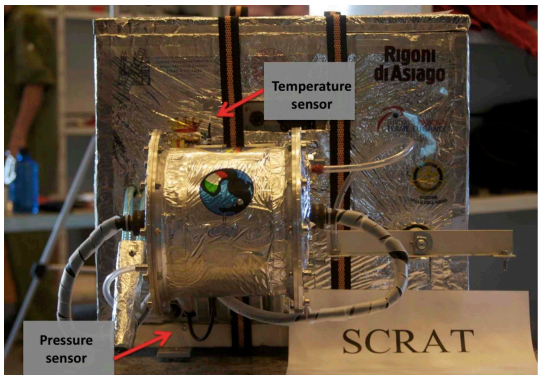


Fig. 3 Sensor layout and position in SCRAT experiment flight model

Unfortunately the pressure sensor experienced some problems during the flight due to the very low temperatures. Its operating temperature range was limited to -40°C and 85°C and it underwent operational problems due to the low temperatures (minimum -71°C ± 1°C) reached during the balloon flight, giving an unreliable pressure value for altitude greater than 7 km i.e. temperatures lower than -40°C as reported in Fig. 4.

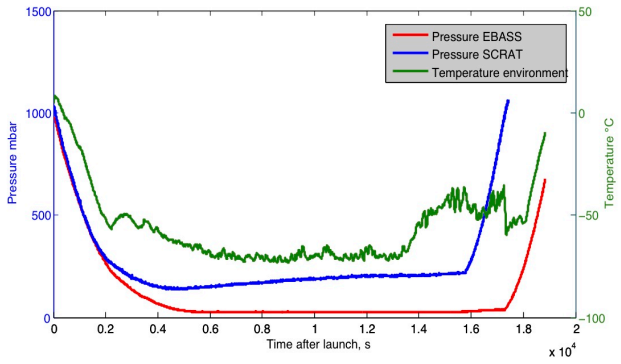


Fig. 4 Pressure data retrieved by SCRAT and EBASS

Given these considerations the pressure data above 7 km were retrieved from the EBASS unit sensor. Concerning the EBASS unit, it was equipped with a pressure environmental sensor model *DRUCK PMP 4015.*

5 Flight campaign results

In this part of the article are discussed in detail the atmospheric and physical results together with the data analysis approach adopted. The treated topics are briefly introduced below in Fig. 5:

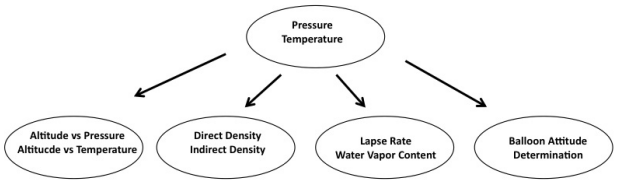


Fig. 5 Data analysis overview

5.1 Pressure and temperature profiles

The pressure at the ESRANGE level (350 msl) was *980.2 ± 4.3 mbar* and the lowest pressure at the top of the BEXUS 10 flight was *23.6 ± 4.3 mbar*. The pressure profiles in ascent and descent phases show the same exponential trend as can be observed in Fig. 6.

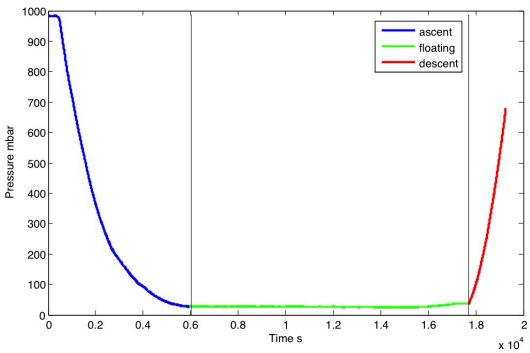


Fig.6 Environmental pressure during the flight ascent (blue), float (green) and descent (red)

The ascent and descent profiles are perfectly coincident as Fig. 7 attests: the two profiles are folded and a theoretical function is overlapped to them.

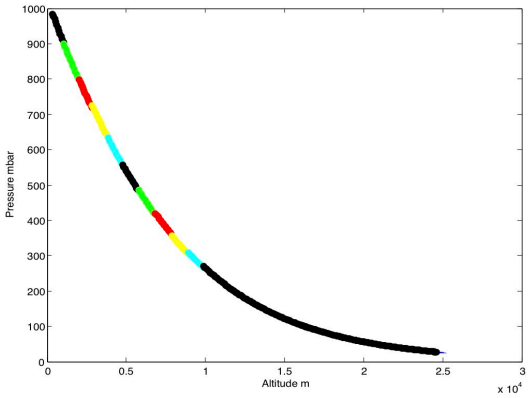


Fig. 7 Exponential decrease of the pressure reproduced by the hypsometric equation Eq. 1

The theoretical law is the hypsometric equation (Eq. 1), and it links the pressure with the altitude under the approximation of a stratified atmosphere whose layers are isothermal:

$$P_1 = P_0 \exp(\frac{z_1 - z_0}{H}) \qquad (1)$$

with P_0 and P_1 respectively the pressure at ESRANGE level and at an higher sample level, z_0 and z_1 their altitudes and the constant H, defined as *scale height*. The constant H contains many terms including temperature. Estimating the constant value for each layer and calculating the hypsometric equation we obtain the plot in Fig. 7. The correlation between the theoretical and empirical data is satisfactory and the pressure profiles are perfectly reproduced by the hypsometric equation (colourful line Fig. 7). It has to be noted that for altitudes greater than 10 km there is a unique scale height, which fits the observed pressure for about 15 km, this is due to the fact that the low and middle stratosphere is a quasi-isothermal layer.

The temperature experimented by SCRAT during its flight is reported in Fig. 8; the spectrum is wide and spans between *4.5°C* at the ground level down to *-71°C* at the top of the BEXUS 10 trajectory. A deeper investigation upon the retrieved temperature is reported in section 5.4.

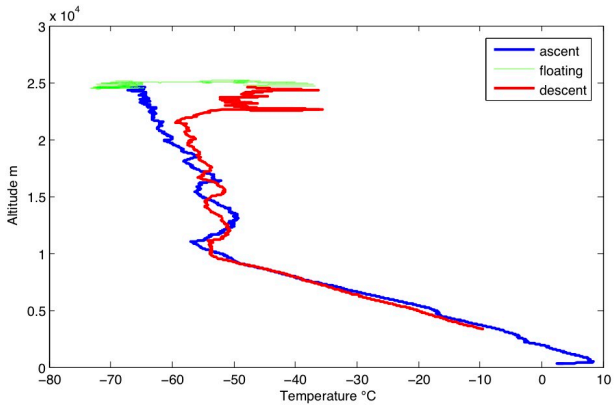


Fig.8 Environmental temperature during the flight

5.2 Density estimation

The atmospheric density has been determined by two different and uncorrelated methods: a direct estimation via the *thermodynamic* laws, and an indirect estimation through the *drag forces* experimented by the system *gondola-parachute* during the descent phase.

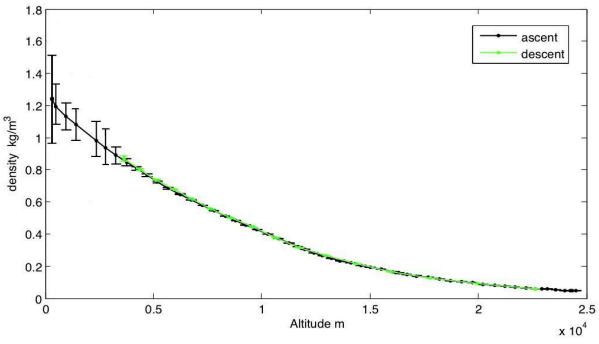


Fig.9 Direct density estimation for ascent and descent phase

The direct density estimation can be derived by the measurement of the temperature and the pressure using Eq. 2 whenever the mean molecular weight M of the gaseous mixture is known:

$$\rho = \frac{P \cdot M}{R \cdot T} \qquad (2)$$

In the case of the air $M = 28.966\ kg/kmol$, P is the pressure, T the temperature and $R = 8.314 \times 10^3 Nm/kmolK$ the gas constant. The ascent and

descent profiles are plotted in Fig. 9 attesting a good accordance. The density estimation by the drag forces is based on the empirical evidence that any falling object into the atmosphere experiences a net drag force opposite to its motion as reported in Eq. 3:

$$F_{drag} = \frac{1}{2} C_D A \rho v^2 \qquad (3)$$

Being C_D the drag coefficient (shape-dependent), A the area of the object, v its velocity and ρ the density of the air (the dragging element). For the present purpose the drag coefficient C_D and the area A refer to the parachute deployed after the cut-off phase, while the velocity is retrieved by the GPS receiver onboard the EBASS unit. Subsequently the cut-off phase the system *gondola-parachute,* after an initial acceleration, reached quickly a constant velocity denoted as *terminal velocity* and lower than *10 m/s* (see Fig. 10), that is established when the weight (F_g) of the system is balanced by the surrounding drag force ($F_g = F_{drag}$).

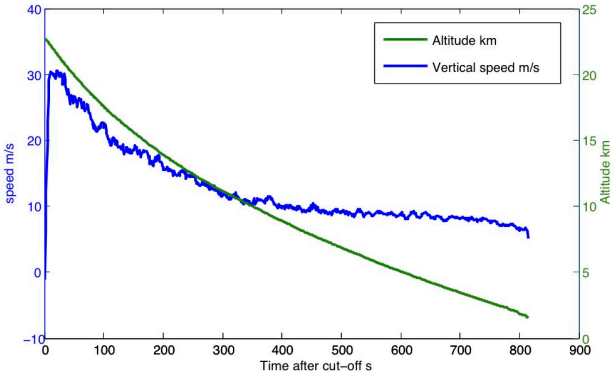


Fig.10 Altitude and velocity of the balloon during descent phase

Once the terminal velocity is reached, Eq. 3 can be reversed to derive the air density. This approach, even if it is affected by a larger uncertainty (see Fig. 11), is particularly effective whenever we have to deal with a planetary atmosphere whose composition is partially or totally unknown.

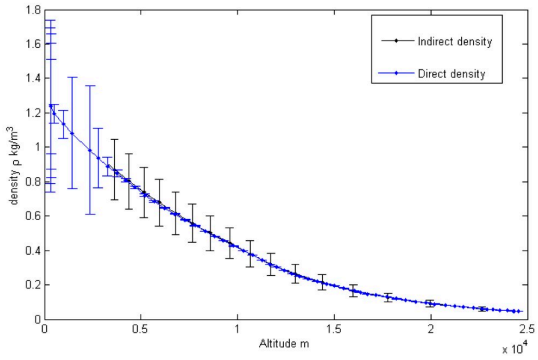


Fig.11 Comparison between direct and indirect density estimation

5.3 Atmospheric lapse rate

By the measurement of the temperature at different altitudes a new parameter can be measured and characterized: the atmospheric lapse rate, i.e. the rate of cooling of the atmosphere with the height. Usually the temperature decreases linearly with the altitude at a mean rate of about Γ = *-6.5°C/km* as we depart from the Earth ground because the infrared emission of the soil decreases. The lapse rate provides also interesting information about the water vapour content of the atmospheric mixture.

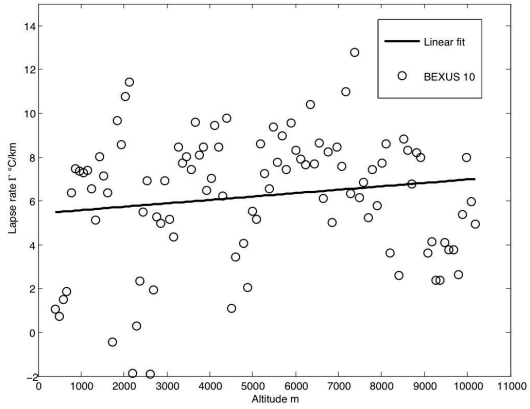


Fig.12 Lapse rate retrieved by SCRAT experiment

The estimation of the lapse rate during BEXUS 10 flight has carried to the result reported in Fig. 12 with a mean value Γ = *-6.2 ± 2°C/km.* The explanation of this atmospheric cooling deals with the mechanism of heat transportation inside the Earth atmosphere. Being evident the presence of large convective motions in our

atmosphere a *dry adiabatic lapse rate* can be introduced as described by Eq. 4 [10].

$$\Gamma_{dry} = \frac{dT}{dz} = -\frac{g}{C_P} = -9.8°C/km \qquad (4)$$

With g gravity acceleration and C_P specific heat at constant pressure. The result of Eq. 4 is however in contrast with the SCRAT measurement (Γ = *-6.2 ± 2°C/km*). There must be some other factor that controls the atmospheric lapse rate limiting the decrease of the temperature with the height. The solution has to be searched in a new formulation of Eq. 4 by the introduction of the so called *saturation adiabatic lapse rate* given in Eq. 5 [10].

$$\Gamma_{wet} = \frac{dT}{dz} = -\frac{g/C_P}{1+(L/C_P)(dw_S/dT)} \qquad (5)$$

where w_s is the mass of saturated water vapor per mass of air and L is the latent heat of condensation in *kJ/kg* that is a polynomial expression of the environmental temperature. The saturation adiabatic lapse rate Γ_{wet} takes into account the condensation heat of water vapor in the energy balance of the air parcel under investigation. The presence of a certain quantity of water vapor, that has a high heat of condensation, makes rising Γ_{wet} at values between *-4* and *-6.5 °C/km.*

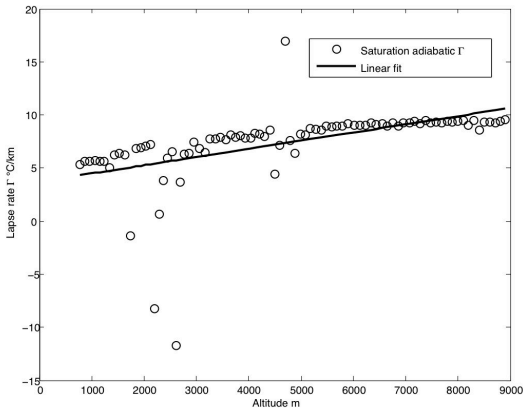


Fig.13 Simulated saturation adiabatic lapse rate over a water surface

In fact a wet air parcel that is rising the atmosphere by convection undergoes a cooling process that makes the water vapor condensing and releasing a certain quantity of energy called *condensation heat L*. The value retrieved by SCRAT is well described by the saturation adiabatic lapse rate and perfectly enclosed in the interval values of Eq. 5. In Fig. 13 and 14 have been simulated the saturation adiabatic lapse rates for two different kinds of ground surfaces: water and ice. Comparing these results with those retrieved by SCRAT it can be noted that the flight data represent an intermediary scenario.

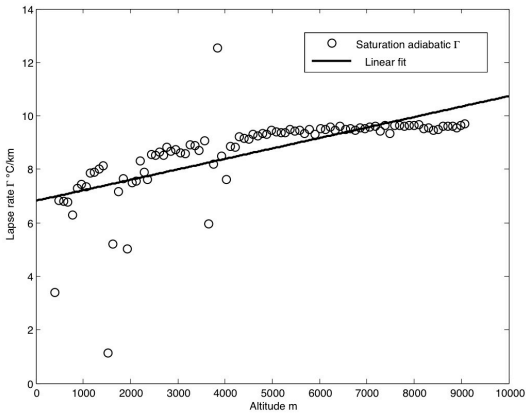


Fig.14 Simulated saturation adiabatic lapse rate over an icy surface

The term w_s in Eq. 5 can be written as follows [10]:

$$w_s = 0.622\frac{e_s}{P-e_s} \qquad (6)$$

with P environmental pressure and e_s saturation water vapor pressure. The values of e_s obey to the *Goff-Gratch* model and are tabulated in literature. By the assumption of this model and the measurement of the pressure the mixing ratio of the atmospheric water vapor can be estimated as given in Fig. 15.

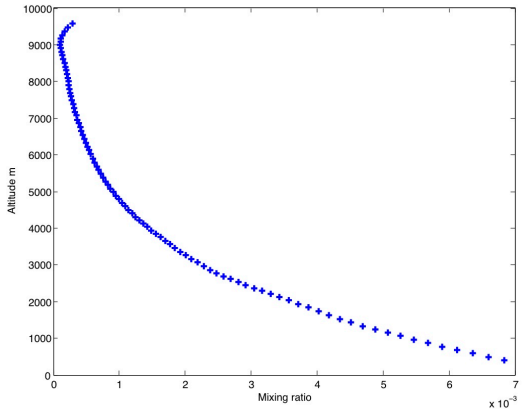


Fig.15 Mixing ratio of water vapor in Troposphere

5.4 Balloon attitude determination

This last section deals with the analysis of the diurnal effects on the stratospheric temperature during the BEXUS 10 flight. As above mentioned, the balloon was launched at 1:07 UT during the night, and it flew for five hours until 6:07 UT early in the morning.

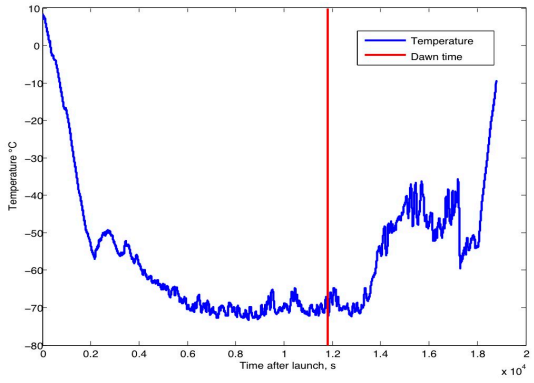


Fig.16 Temperature spectrum during the flight. Dawn time (red line)

The Sun raised at 4:17 UT and the last segment of the floating phase was during the daytime leading to an interesting correlation of the balloon's dynamics and the retrieved environmental temperature. As can be observed in Fig. 16 there are many oscillations in the temperature spectrum after the dawn (red line). These oscillations are thought to be originated by the rotation of the balloon with respect to the Sun. Evidences of the rotating behaviour of the stratospheric balloons have been figured out in many other stratospheric flights as reported by NASA [11] and H H Wang et al. [12] and it is thought that the rotation of stratospheric balloons is due to the convective motions of the Helium masses enclosed in the balloon heated up by the Sun. The intent of this section is to partially infer the attitude of the balloon by the use of the temperature sensor as a *Sun sensor*. To achieve useful information about the attitude of the balloon during the diurnal segment of the flight we need the data from the GPS receiver, model *Rockwell Jupiter*, the temperature oscillations recorded by the temperature sensor and the AGI's software *Satellite Tool Kit*, *stk* [13], to create a virtual reference mission as close as possible to the real BEXUS10 flight. First of all, to verify the interpretation of the temperature oscillations a thermal model accounting for the radiation heat transfer between the main bodies has been implemented.

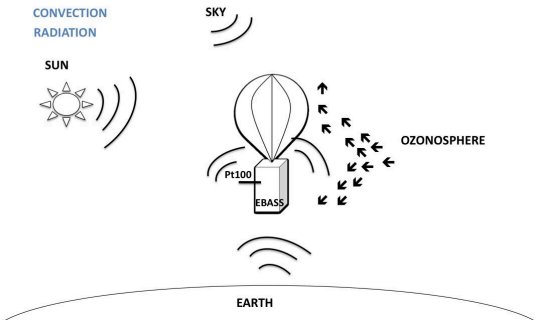


Fig.17 Picture of the bodies involved in the thermal model transferring heat by radiation and convection

At a first instance only radiation and a weak contribution of ozonospheric convection have been considered; conduction has been neglected since the bodies aren't in direct contact preventing any conduction heat transfer. Each body exchanges simultaneously heat with all the others by radiation as reported in Fig, 17. So for example the term $\dot{Q}_{pt}(t)$ in Eq. 7 represents the total amount of energy exchanged by the temperature sensor (*pt*) with all the surrounding bodies:

$$\dot{Q}_{pt}(t) = \sum_{t_{dawn}}^{t_{cut-off}} \dot{Q}_i(t) = \dot{Q}_{Bl}(t) + \dot{Q}_{\oplus}(t) + \dot{Q}_{\Theta}(t) + \dot{Q}_{O_3}(t) + \dot{Q}_{sky}(t) \qquad (7)$$

balloon (*Bl*), Sun ($\odot$), ozonosphere (O_3), Earth ($\oplus$) and sky (*sky*). The temporal window considered for the simulation goes from the dawn until the cut-off phase beyond that the motion of the gondola is too irregular to recognize regular rotations. Of all the heat transfer terms in Eq. 7 the most peculiar is the $Q_{\odot}(t)$ term, (Eq. 8) that represents the heat coming from the Sun. It is modulated by a sine function with variable period to reproduce the oscillations observed in the temperature spectrum.

$$\dot{Q}_{\Theta}(t) = \frac{A_{pt}}{3} \cdot \alpha_s \cdot C_{\Theta} \sin\left(\frac{2\pi t}{T}\right) \qquad (8)$$

with A_{pt} representing the portion of the temperature sensor area receiving a solar flux equal to $C_{\odot}\approx 1366 W/m^2$ and weighted by the periodic function and the absorption coefficient α_S of the incident solar radiation.

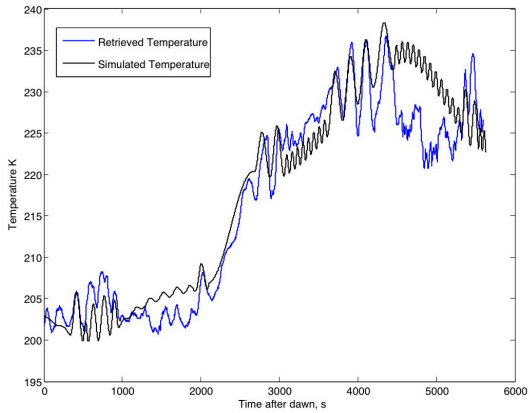


Fig.18 Retrieved and simulated temperature spectrum from the dawn onwards

The result of this thermal model is quite surprising as Fig. 18 shows. The fundamental harmonic of the temperature spectrum is well reproduced by the model (black line) and about 15 oscillations induced by many balloon rotations are fairly simulated.

All the simulated temperature maxima are interpreted as a direct heat radiation transfer between the temperature sensor and the Sun, one facing the other, while the minima as an eclipse of the temperature sensor that enters the cone of shadow of the balloon. To have a cross correlated interpretation of these data a simulated mission scenario has been created using *stk* software. The trajectory of BEXUS 10 has been reconstructed based on the GPS coordinates retrieved during the flight. The access to the Sun of BEXUS 10 (the visibility between the two objects) has been computed based on the precise stk Sun ephemeris and the GPS data. A first important result is the exact time of the dawn at the stratosphere level (4:17 UT) as reported above. From the dawn onwards all the temperature oscillations with an amplitude equal or greater than $\Delta T=10$ °C have been imputed to the balloon's rotation. Then the temperature sensor has been reproduced by a geometrical conic sensor attached to the balloon's body. The sensor is defined by the cone amplitude (see Fig. 19) that establishes the *field of view, fov,* of the object; many simulations have been done tuning the fov to the final value of *180°*. The basic idea behind this simulation is trying to reproduce the temperature oscillations as the consequence of sunlight-to-eclipse cycle of the temperature sensor driven by the balloon rotational motion.

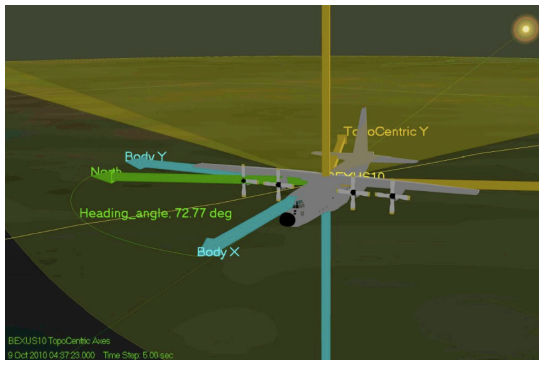


Fig.19 Conic sensor (yellow) with fov = 180° attached to the balloon (grey aircraft)

In this perspective the fov of the sensor should contain the Sun for about half the period of rotation making rise the temperature at a maximum; then as the Sun exits the fov because the sensor enters the cone of shadow of the balloon the sensor radiates to the other bodies and cools causing a minimum in the temperature spectrum.

The last step of the stk scenario consists on the balloon attitude setting and determination based on the rotations discovered in the study of the environmental temperature. For the present purpose a *NordEastDown* reference frame*, NED r.f.,* has been assumed for the balloon flight. The NED r.f. has the X–axis is directed to North, the Y–axis to East and the Z-axis points toward the Earth's surface, aligned with the instantaneous normal to the Earth's surface along the gravity field direction (see Fig. 20). In this manner the balloon's attitude configuration during the ascent and floating phase is that of an object with the Z–axis constantly aligned with normal to the surface. The balloon's attitude has to be set with respect to an external r.f. by defining the yaw–pitch–roll angles. The effects of the rotations on the NED r.f. have been studied using a *Earth Centered Fixed, ECF*, r.f.. This terrestrial r.f. rotating with the Earth has its origin at the centre of the Earth itself, the Z–axis passing through the North pole, the X–axis passing through the equator at the prime

meridian and the Y–axis, determined by the right-hand rule, crossing the equator at 90◦ longitude.

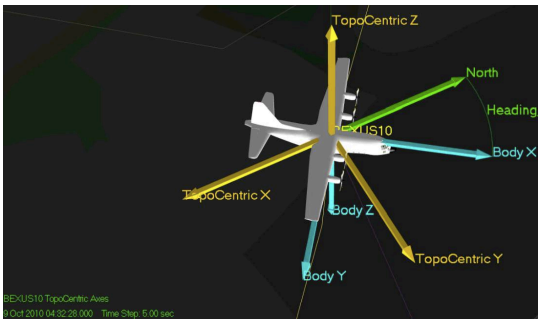


Fig.20 NED r.f. (azure arrows) in comparison to Topocentric r.f. (yellow arrows)

For the night flight segment the balloon's attitude has been considered completely frozen into the ECF r.f. since there are no information in absence of the Sun. From the dawn onwards, a rotational degree of freedom has been given to the system and in proximity of each temperature oscillation a complete rotation of the NED r.f. has been inserted in the attitude determination file. The rotation has been supposed around the Z-axis of the NED r.f. and simulated by a complete rotation of the *yaw* angle. In this scenario the access duration is assumed to be half the rotation period since the temperature sensor has a fov that covers 2π *str*, half of the entire solid angle. The interval of access between the Sun and the sensor have been retrieved and doubled by stk and compared with the periods obtained by the thermal model discussed above.

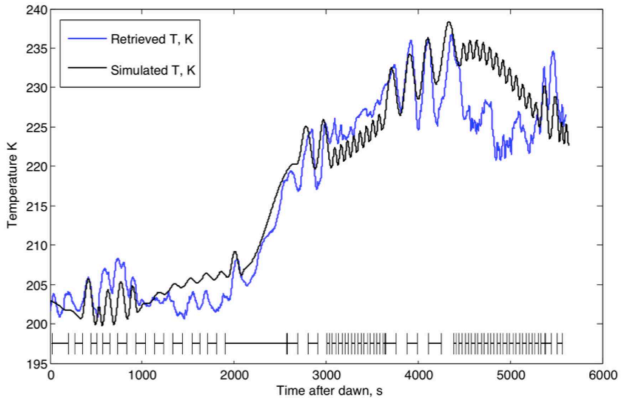


Fig.21 Temperature spectra and Sun accesses retrieved by stk software for a 180° fov sensor

In Fig. 21, overlapped to the retrieved and simulated spectra, there are (bottom Fig. 21) the intervals of access obtained with stk.
As a by product it is possible to estimate also the rotational speed of the balloon on the basis of the thermal model, of stk and also from the *Fourier analysis* of the temperature spectrum. The results, which attest slow rates of rotation, are collected in Table 1 and in good agreement with those experimented by NASA [10].

Thermal model	$\nu = 0.34 \pm 0.06$ rpm*
stk	$\nu = 0.32 \pm 0.07$ rpm
Fourier analysis	$\nu = 0.41 \pm 0.08$ rpm

Table 1 Balloon rotational speed (* rates per minute)

6 Conclusion

The modelling of the temperature and pressure data retrieved during BEXUS 10 flight by SCRAT experiment has led to a wide range of applications and atmospheric analyses. The atmospheric density together with the lapse rate and water vapour content has been determined. Moreover the study of the stratospheric temperature oscillations has led to achieve many interesting information about the dynamics and the attitude of the balloon during the daily segment of the floating phase.

Acknowledgments

Authors are grateful to the REXUS/BEXUS Team for the great support during the whole Programme and to the ESRANGE staff for the precious collaboration during the entire flight campaign. Authors want also to thank GMS SPAZIO srl for having provided a free license for the stk software.

References

Proceedings references

[1] http://www.rexusbexus.net/

[2] http://www.cnrm.meteo.fr/concordiasi/

[3]http://www.insu.cnrs.fr/co/actions-sur-projets/pnca-programme-national-de-chimie-atmospherique

[4] http://marsrover.nasa.gov/home/

[5] http://www.esa.int/esaMI/Cassini-Huygens/

[6] http://www.nasa.gov/mission_pages/msl/

[7] http://exploration.esa.int/

[8] *BX–BEXUS–user–manual–v6–0–20Dec09*

[9] http://it.rs-online.com

[10] '*Theory of Planetary Atmospheres'*, J. W, Chamerlain and D. M. Hunten, Academic Press Inc, 1987

[11] *NATIONAL SCIENTIFIC BALLOON FACILITY RECOMMENDATIONS FOR GONDOLA DESIGN APRIL 1, 1986 http://www.nsbf.nasa.gov*

[12] "*Analysis of the Motion Control Methods for Stratospheric Balloon-Borne Gondola Platform" H H Wang et al., Journal of Physics: Conference Series 48 (2006) 1295–1300*

[13] http://www.agi.com

3rd CEAS Air&Space Conference
21st AIDAA Congress

Anti-contamination by biochemically active structures

U. Reidt, C. Heller, H. Maier-Boetzel, L. Chancé, K. Bauer
EADS Innovation Works, Sensors, Electronics & Systems Integration, 81663 Munich, Germany

Keywords: *dirt prevention/rejection, ice prevention/rejection, micro/nano-coating, enzyme*

Abstract

With respect to the ability to reduce and remove organic contamination, including insects, on a laminar airfoil, biocatalytically active agents were here investigated, in solution or immobilized on airfoil materials surfaces. The feasibility and working principle of those structures were demonstrated within this study, and results of experiments reported. Several combinations of enzymes and linkers were tested for the degradation of proteins including those from real insect extracts, and the influence of several environmental factors was measured. Active surfaces with a higher performance as commercial ones were also obtained.

1 Introduction

Insect debris and icing on the leading edge of an airfoil causes the flow to become turbulent at a very early stage [11], inducing more drag and an increase of fuel consumption. It is thus desirable to prevent or remove debris as much as possible, and preferably before the cruise flight is reached. An approach is to develop a surface treatment or coating that can prevent insects from sticking to the aircraft's surface. The present study deals with the functionalization of surfaces by binding active molecules (enzymes) in order to detach the organic debris from the surface. To our knowledge, very little work has been undertaken so far to try to detach insect by biochemical tools [7].

Contamination caused by insects can be expected primarily within the first 600m of altitude during climb flight. Due to environmental conditions, only little water is available in the atmosphere, and the ambient temperature goes down to -60°C. In a regular flight, all cleaning process requiring water have a time window of a few minutes for these reasons. A self-cleaning surface would avoid the need of constantly applying a cleaning liquid onto the airfoil.

Proteases and other digestive enzymes are very common in nature and can be easily produced. Many of them have been optimized for specific applications such as washing procedures. In a first choice, trypsin and chymotrypsin were selected as biocatalytic proteins. They are available in large quantities at a reasonable price, and have a high number of cleavage sites in almost all proteins [6]. Their activity and ability to digest various types of proteins was first investigated using several analytical methods. Their resistance to

environmental factors such as temperature and medium quality was then tested. Finally, enzymes were immobilized on different types of surfaces and the resulting samples were tested and compared to commercially available activated surfaces.

2 Measurements of enzymes activity and protein degradation ability in solution

To determine the activity of the enzymes and to validate the catalytic reaction, we have used three different analytical methods. We have focused on enzyme kinetics. For complete characterization of enzymes, the exact behavior of the kinetics and possible inhibitory factors is needed.

2.1 Spectrophotometry

Spectrophotometric assays are performed using TAME (*p-toluene-sulfonyl-L-arginine methyl ester*) as a substrate for trypsin and BTEE (*benzoyl-L-tyrosine ethyl ester*) as a substrate for chymotrypsin, respectively. In both cases, the clear substrate is converted to a chromogenic molecule upon enzymatic reaction. In the case of TAME, the reaction product has an absorption maximum at 247 nm wavelength, whereas the BTEE product absorbs at 256 nm. The absorption – monitored with a spectrophotometer – is proportional to the enzymatic activity. We used a GE Healthcare Ultraspec 2100 pro spectrophotometer. Absorbance readings were taken every 10 seconds for a total assay time of 600 seconds.

The trypsin activity was first determined using a well-established assay [5]. After mixing substrate and assay buffer, the reaction was started by adding the trypsin to the solution at different concentrations. As shown in Fig. 1, the activity increases with the enzyme concentration.

In many cases, enzymatic activity can also be dependent on the substrate concentration in a non-linear manner. To figure out the optimum substrate concentration, measurements with increasing TAME concentrations were performed. Keeping constant the trypsin concentration at 100 µg/mL, substrate concentrations between 2 mM and 100 mM were tested. The optimum substrate concentration for the assay was between 0.01 and 0.025 M TAME.

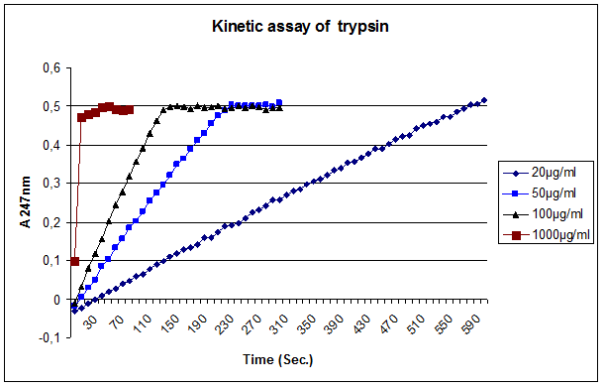


Fig. 1 Trypsin assay with different enzyme concentrations

The catalytic activity of trypsin can be inhibited by the presence of trypsin inhibitor from chicken egg white. Using a concentration of 10 mg/mL trypsin inhibitor to a trypsin concentration of 50 µg/mL, the proteolytic reaction was nearly completely inhibited. Enzyme inhibition can serve as a mean to determine the reaction mechanism as well as a reference for other inhibitory molecules that might be present in the environment.

We extended our series of tests to a second protease, chymotrypsin. A similar procedure was adopted, and the increase in absorption was recorded for two different concentrations of the enzyme (Fig. 2).

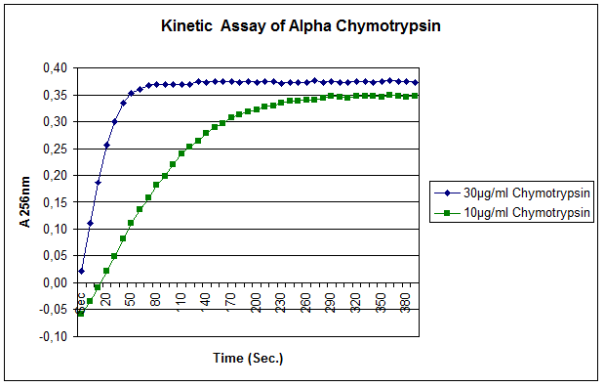


Fig. 2 Kinetic assay of alpha chymotrypsin

2.2 Electrophoresis method

A second approach is to follow the degradation of a protein or a protein mixture. Upon enzymatic digestion, a protein will be cut down to smaller units (peptides). Using polyacrylamide gel electrophoresis (SDS-PAGE), the proteins and degradation products are separated according to size. Thus, degradation can be monitored.

The enzymatic activity of trypsin and chymotrypsin was previously quickly tested and quantified using artificial substrates. However, the natural substrates of these enzymes are proteins. We have chosen BSA (*Bovine Serum Albumin*) as a model substance to monitor proteolytic activity of both enzymes.

BSA was digested with different concentrations of trypsin. Reactions were started by adding an increasing volume (0.5 µL – 32 µL) of a trypsin stock solution to trypsin buffer containing 100 mg/mL BSA. All reactions were incubated for 1 minute at room temperature and stopped by adding SDS PAGE. The apparent molecular weight of trypsin is 25.6 kDa whereas uncleaved BSA has a size of 65 kDa. An increase in protease concentrations results in smaller fragments (more degradation).

The same experiment was performed by using chymotrypsin. The SDS PAGE indicates a very similar result as with trypsin (Fig. 3).

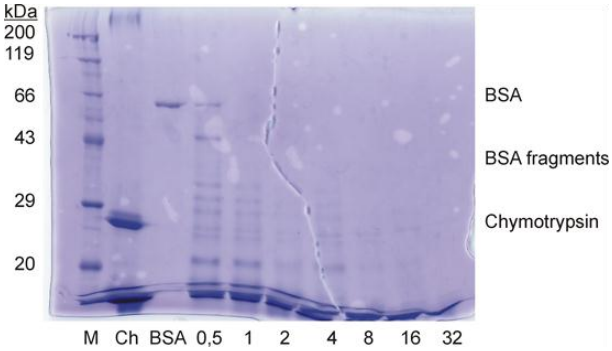


Fig. 3 Degradation of BSA with alpha chymotrypsin. M=molecular weight marker, Ch=Chymotrypsin, and 0.5-32= increasing amount (in mL) of chymotrypsin added to reaction

The drosophila is a fruit fly that was chosen as a model insect in the course of the project. To study the degradation of the insect by the enzymes, 20 flies were homogenized in presence of 1 mL phosphate buffered saline (PBS). The extract was ready to use after spinning down the debris. 100 µL each of drosophila extract has been digested with various volumes of a 10 mg/mL trypsin solution for 1 min at room temperature. As a negative control, pure Drosophila extract without trypsin has been used. The digestion reaction was stopped by adding 20µl of the reaction to 20µl of SDS (*sodium dodecyl sulfate*) sample buffer. After denaturation at 110°C for 15 min, samples were loaded on a SDS gel. After electrophoretic separation the protein bands were visualized by staining the gel with Coomassie brilliant blue.

2.3 Fluorescence measurements

This last method uses a fluorescently labeled protein, FTC-casein, which delivers small peptide fragments carrying fluorescein when digested. We used a fluorescence plate reader (Packard Fluorocount) to detect the fluorescence signal from 0 to 240 minutes after the start of the reaction.

The experiment shows the kinetic digestion of FTC-casein using Trypsin for the reaction. As negative controls, FTC-casein and FTC-casein with Trypsin inhibitor solution have been used. The results show that the FTC-casein is a highly sensitive method for monitoring the trypsin reaction, more than with electrophoresis (Fig. 4). The activity of as few as 10 ng trypsin could be easily monitored and the reaction could be completely suppressed by adding an inhibitor.

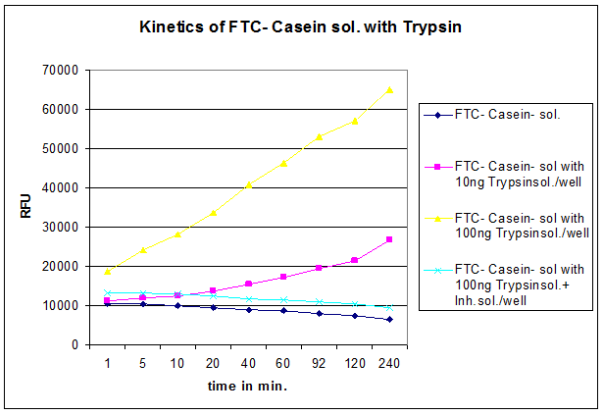


Fig. 4 Trypsin digestion of FTC-casein in solution

3 Influence of environmental factors

In the application foreseen in this project, the enzymes will have to work not in a defined laboratory environment but rather under harsh conditions. Following tests show the ability of the enzyme to resist to several factors.

3.1 Temperature

To simulate cold environmental conditions, solutions of trypsin of different concentrations were frozen for 1 hour at -20 °C and then thawed and kept at room temperature for 90 min. Standard photometric trypsin assay showed that the enzymatic activity was not influenced by this freezing-thawing cycle.

Exposure to heat was then tested. Enzyme solutions were incubated at different temperatures for 15 min each, then cooled down to room temperature for another 15 min and immediately used in standard photometric trypsin assay (Fig. 5).

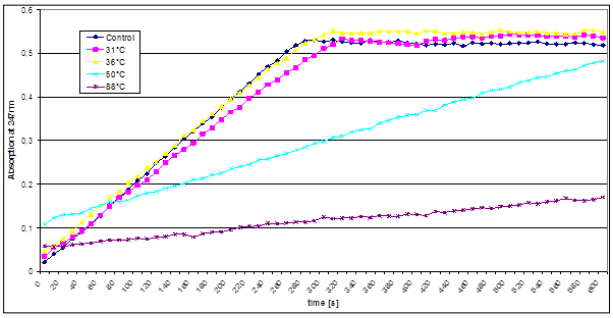


Fig. 5 Activity of trypsin after heating to different temperatures

The enzymatic activity decreases with incubating temperature, due to the fact that the optimum temperature of the trypsin is the body temperature (37 °C). Above that, protein denaturing occurs. However, if enzyme stability at higher temperatures would be necessary or favorable, it might be possible to replace these enzymes with heat stable ones from thermophilic organisms [4].

3.2 Influence of glycerol

To minimize the formation of ice on aircraft wings, anti-freezing and de-icing solutions are currently used before takeoff. To mimic such a situation we have tested glycerol as an anti-freezing agent that can prevent solutions from freezing at temperatures of -20°C and below (depending on the concentration). Here, we have studied the influence of glycerol at a concentration of 10% as an example. The addition of 10% glycerol has only little influence on trypsin activity.

3.3 Activity in different media

All the enzyme assays described above have been performed under laboratory conditions, i.e. in well-defined solutions made with ultra-pure water. However, this does not reflect the real situation which occurs outside the laboratory. We have therefore tested the enzyme activity in media other than the reaction buffer, i.e. in tap water and in water collected from a puddle (rain water).

From the results presented in Fig. 6, it is clear that media other than the optimum buffer can indeed lower the enzymatic activity. Again, this is not surprising, as natural occurring enzymes are always working best in their normal environment. Again, it might be possible to replace these enzymes with natural or genetically engineered enzymes that show optimum activity under different conditions.

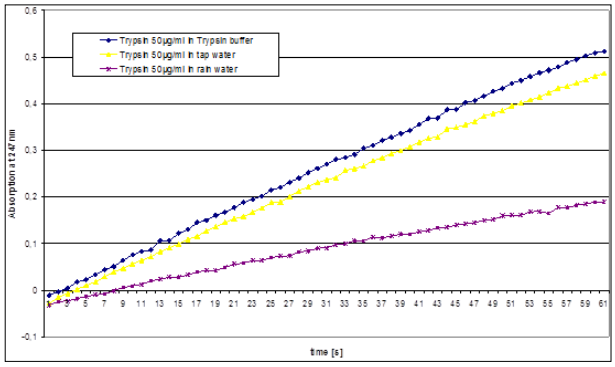


Fig. 6 Influence of different reaction media on the trypsin activity

4 Immobilization of enzymes

After establishing and validating the enzymatic reactions in solution, the immobilization of enzymes was tested. In general, a biocatalytically active surface is designed as in Fig. 7. The surface has to be activated first to make it accessible to chemical

reactions. The linker molecule can covalently link the enzyme to the activated surface. In most cases, it also contains a so-called "spacer" between its active groups, which consists of a flexible chain of atoms. This method is commonly used for coupling proteins or cells to surfaces in applications such as medical implants, protein biochips (microarrays). Biosensors and bioreactors [1],[2],[3],[8],[9],[10]].

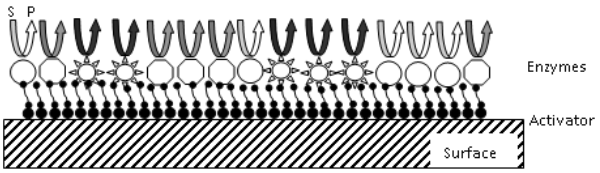


Fig. 7 General design of a biocatalytically active surface. Each type of enzyme catalyzes a specific reaction (symbolized by the bent arrows)

Several surface materials were chosen, as well as activator molecules and linkers to build biocatalytically activated surfaces. The trypsin activity was measured for different combinations and digestion of proteins by the immobilized trypsin was then recorded.

4.1 Immobilization procedure

Square pieces of 1 x 1 cm^2 in size of different materials were thoroughly cleaned by incubating them in a NaOH solution for 30 min and then treating them with ultrasound for 15 min. The pads were then washed several times with distilled water and neutralized with HCl solution, and incubated in a silane solution for activation. After a baking step, the cross linker solution and the trypsin solution (100 mg/mL) were applied to the surface. Activated samples were then stored in the cold.

Samples were used in photometric assays. A buffer solution containing TAME was used to monitor the activity of the enzyme, as described in the following paragraph. It should be high lightened that the reactions were monitored in a standard photometric cuvette in a volume of 3 mL with the surface sample on the bottom. TAME molecules have to travel rather long distances by diffusion to get to the enzyme layer. Moreover, the amount of enzyme molecules present is much lower than with liquid enzyme. This leads to lower reaction rates than in previous activity measurements.

4.2 Test of various components

4.2.1 Activator molecules

Two different activator molecules were used to obtain amino functionalized surfaces: APTES (*Aminopropyltriethoxysilane*) and AEAPTMS (*Aminoethylaminopropyltrimethoxysilane*) solutions. A glass surface was used in this test, and BS^3 (*Bis (Sulfosuccinimidyl) suberate*) was chosen as a linker for the trypsin. The enzymatic reaction in presence of the TAME substrate was then measured for 60 min with a 1 min interval using the spectrophotometer. No significant difference in enzymatic activity was found between the two experiments (slope of the curve).

A negative control, made of glass plates that have not been activated by silanization, shows no enzymatic activity. This proves that the immobilization of trypsin is performed through covalent linkage (chemical bonds) and not through physical adsorption.

4.2.2 Cross linkers

Three different linkers were tested: BS^3, but also EDC (*Ethylendiisocyanate*) and GA (*Glutaraldehyde*). They were applied onto the pads after their activation with APTES. Trypsin was then attached to the surface. For comparison, we also used commercially available glass surfaces, already activated with amine groups (*Genetix*). A negative control test was also used with a non-activated glass pad. The activity of the enzyme measured with the spectrophotometer is presented in Fig. 8. It shows that all 3 cross-linking methods perform very well without significant difference between each of them. The glass plates activated with our own methods show a much higher performance than the commercial surfaces that were tested.

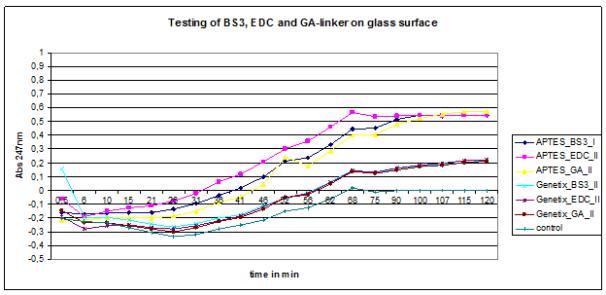


Fig. 8 Activity of immobilized trypsin bound to glass using different cross linkers

4.2.3 Surface materials

Finally, other surface materials than glass were tested such as aluminium, carbon fibre composite and titanium, all of them being used as aircraft constituents. All materials were activated with APTES and trypsin was immobilized to the surfaces using the different cross-linkers previously presented.

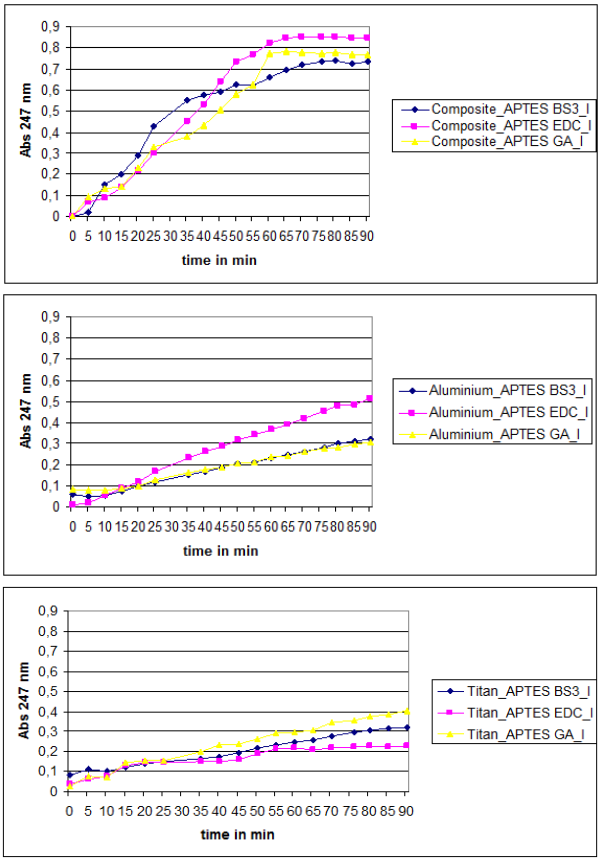


Fig. 9 Activity of immobilized trypsin bound to carbon fiber composite, aluminium and titanium surfaces

The results show that the enzyme can successfully be coupled to all the surfaces, with a small influence of the type of cross linker used once again. The composite material seems to bind more trypsin than the other surfaces.

4.3 Digestion of proteins by the activated surface

The last tests realized were dealing with the ability of the activated surfaces to digest proteins. Two aluminium pads were activated with APTES and trypsin was immobilized via BS^3 cross linker. The FTC-casein solution was prepared by adding 10 µL of the 5 mg/mL stock solution to 5 mL trypsin buffer (dilution 1:500). The positive control consisted in a solution of liquid trypsin (1 µL of a 2.5 µg/µL stock solution) mixed with 1 mL of FTC-casein. The negative control was made of an aluminium pad activated but not treated by BS^3 cross linker. After various amount of time, the samples were taken out of the reaction and a trypsin inhibitor solution was added. They were then diluted with a solution of trypsin buffer and a measure of the fluorescence was performed (Fig. 10).

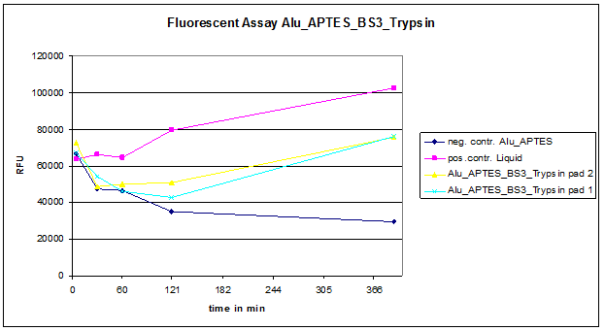


Fig. 10 Digestion of FTC-Casein with immobilized trypsin

Another measurement was realized to measure the digestion of the FTC-casein in presence of Drosophila extract. Like before, the fluorescence level was regularly measured. Results show an increasing level with time for both the pure FTC-casein as substrate and for the mixture of FTC-casein and Drosophila extract. However, the latter one shows a smaller increase in fluorescence, indicating that there is a competition between the two substrate, and thus that the insect proteins are being digested by the enzyme (Fig. 11).

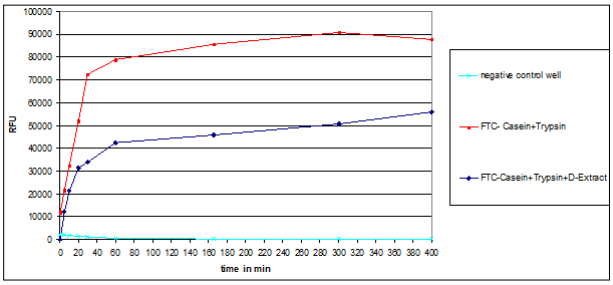


Fig. 11 Digestion of FTC-casein + Drosophila extracts mixture with immobilized trypsin

5 Conclusion

The studied enzymes showed their ability to quickly degrade various types of proteins including fruit fly debris. Several measurement tools were used. They also showed an interesting tolerance to several environmental factors such as temperature variations, occurring during a flight.

The immobilization of enzymes on surfaces was successfully realized and their digestive potential was measured. A higher performance than commercially available surfaces was measured. Simple cross linkers such as glutaraldehyde showed equally good results as more sophisticated ones with spacer molecules. This result could lead to a cost reduction of the active surface manufacturing as the sophisticated cross-linkers are much more expensive than conventional ones.

More research has to be done to determine exactly the requirements of the active surface for an efficient cleaning, and in which environmental conditions. The activity of the immobilized enzymes has also to be improved, and solutions should be found to obtain a reusable biocatalytically modified surface without losing too much efficiency.

References

[1] Angenendt, P., "Microarray Technology as a Universal Tool for High-Throughput Analysis of Biological Systems", *Drug Discovery Today*, Vol. 10, No. 7, 2005, pp. 503-511.

[2] Cao, L., *Carrier-bound Immobilized Enzymes: Principles, Application and Design*. Wiley-VCH, Weinheim, 2005.

[3] Guo, A., Zhu, X.Y., "The Critical Role of Surface Chemistry in Protein Microarrays", *Functional Protein Microarrays in Drug Discovery*, 1st ed., CRC Press, 2006, Chap. 4.

[4] Hickey, A.M., Marle, L., McCreedy, T., Watts, P., Greenway, G.M., Littlechild, J.A., "Immobilization of thermophilic enzymes in miniaturized flow reactors", *Biochemical Society Transactions*, Vol. 35, part. 6, 2007, pp. 1621-1623.

[5] Hummel, B.C.W., "A modified spectrophotometric determination of chymotrypsin, trypsin, and thrombin". *Canadian Journal of Biochemistry and Physiology*. Vol. 37, 1959, pp. 1393-1399.

[6] Keil, B., "Trypsin", *The Enzymes*, 3rd ed., Academic Press, New York, 1971, pp. 249-275.

[7] O'Donoghue, D., Young, T.M., Pembroke, J.T., O'Dwyer, T.F., "An investigation of surfactant and enzyme formulations for the alleviation of insect contamination on Hybrid Laminar Flow Control (HLFC) surfaces", *Aerospace Science and Technology*, Vol. 6, 2002, pp. 19-29.

[8] Sharmin, F., Rakshit, S., H. Jayasuriya, H., "Enzyme Immobilization on Glass surfaces for the development of Phosphate detection Biosensors", *Agricultural Engineering International: the CIGR Ejournal*, Manuscript FP 06 019, Vol. 9, 2007.

[9] Sobek, J., Bartscherer, K., Jacob, A., Hoheisel, J., Angenendt, P., "Microarray Technology as a Universal Tool for High-Throughput Analysis of Biological Systems", *Combinatorial Chemistry & High Throughput Screening*, Vol. 9, 2006, pp. 365-380.

[10] Tebbe, D., Thull, R., Gbureck, U., "Influence of spacer length on heparin coupling efficiency and fibrinogen adsorption of modified titanium surfaces", *BioMedical Engineering OnLine*, 6:31, 2007.

[11] Young, T.M., Humphreys, B., Fielding, J.P., "Investigation of hybrid laminar flow control (HLFC) surfaces", *Aircraft Design*, Vol. 4, 2001, pp. 127-146.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Fatigue crack propagation through adhesively bonded stiffened panels: a model to describe the skin crack behavior underneath the stiffeners

N. Gioia, G. Molinari, I. Meneghin, G. Ivetic and E. Troiani
University of Bologna, Second Faculty of Engineering, DIEM - Aerospace Engineering

Keywords: Damage Tolerance, Adhesively bonded stiffened panels, Fatigue crack propagation

Abstract

Adhesively bonded stiffened panels can provide the best damage tolerance proprieties if compared to the conventional integral and riveted panels. Bonded panels can, indeed, take big advantage by the long fatigue crack propagation periods through the skin underneath the bonded stiffeners. To exploit the aforementioned benefit of bonded panels for safe and weight effective structures, reliable prediction models taking into account the fatigue crack propagation under the bonded stiffeners have to be provided.
A well known analytical literature model, developed for the damage tolerance characterization of patch repairs, was implemented and subsequently validated by fatigue crack propagation test results concerning wide stiffened panels. The implemented model confirmed to be a reliable and effective method for the damage tolerance design optimization of thin-walled bonded structures.

1 Introduction

The Slow Crack Growth [SCG] is cornerstone of the Damage Tolerance [DT] design criteria widely applied to aeronautical structures. In accordance with SCG, the structure must be designed in order to guarantee slow Fatigue Crack Propagation [FCP] under the regulatory loads; as a consequence, any damage introduced in the structure can be detected before fatigue propagating to the critical dimension that determines a catastrophic failure.
Airframes are thin-walled structures constituted by stiffened panels, i.e. thin sheets (skin) reinforced by riveted, integral or adhesively bonded stiffeners (e.g. stringers, frames and doublers) (Fig. 1). In presence of a through the thickness skin crack, adhesively bonded stiffeners are more effective in hindering the fatigue skin crack propagation than riveted and integral stiffening elements (Fig. 2). Bonded stiffeners take indeed big advantage by the long FCP periods the skin crack propagates underneath them; in Fig. 2 the typical FCP behavior of an integral, riveted and bonded panel are compared (the FCP period under the bonded stiffener [ΔNunderB] is longer than the period taken under the riveted stiffener [ΔNunderR]). As a consequence, bonded stiffened panels can provide the best FCP performances in comparison with the riveted and integral ones [1, 2], allowing reliable and weight-effective solutions in accordance with the SCG design criterion.
In order to exploit the beneficial DT performances of the adhesively bonded stiffened panels, reliable FCP predictions models that can describe the long fatigue skin crack propagation period underneath the bonded stiffeners are fundamental. This beneficial period is

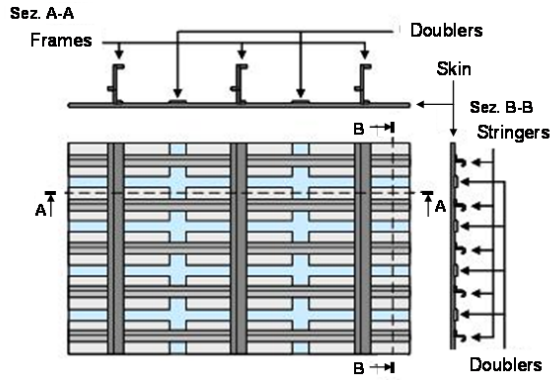


Figure 1: Adhesively bonded stiffened panel with orthogonally arranged stiffeners (stringers, frames and doublers) bonded on the skin.

currently neglected in the design of bonded stiffened panels due to the inadequacy of the employed design methods [2].

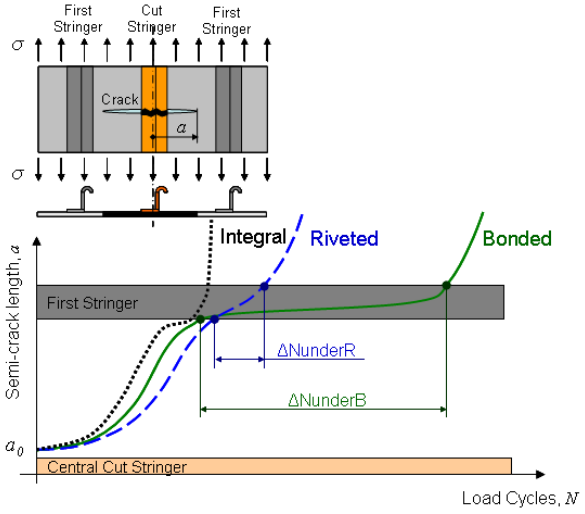


Figure 2: General fatigue crack propagation behavior of a riveted, integral and bonded stiffened panel.

An analytical model, based on the inclusion analogy of bonded reinforcements for the Stress Intensity Factor (SIF) calculation, was proposed by L. R. F. Rose for the DT characterization of patch repairs [3]. It was implemented in the present work in order to describe the crack retardation effect of bonded stiffeners on skin cracks fatigue growing underneath them. The supplementary effect of adhesive delaminations that can develop around the advancing crack tip at the adhesive interface between the bonded stiffener and the cracked skin was considered in the model as well.
Rose's model was successfully validated by FCP tests results concerning wide adhesively bonded stiffened panels representative of typical wide-body fuselage applications; the experimental data were provided in frame of a cooperation with a leading European aeronautical company. The model was implemented in order to improve the capability of an analytical tool developed by the authors on the basis of the displacement compatibility method [4, 5], named LEAF (Linear Elastic Analysis of Fracture) [6], in predicting the DT performances of adhesively bonded stiffened panels. The integration of LEAF with the Rose's model provided an effective a reliable tool for the preliminary design phase of new structural concepts.

2 Analytical model

A fatigue crack propagating through the skin underneath an adhesively bonded stiffener is retarded by a crack hindering mechanism known as crack-bridging [3, 7]. As a consequence, the stiffener restrains the skin Crack-Opening Displacement (COD) in the stiffener covered area, as sketched in Fig. 3(a), reducing the driving force that promotes the crack propagation.
An analytical model to predict the COD of a skin crack propagating underneath an adhesively stiffener bonded on a loaded panel was proposed by L. R. F. Rose [3]. The crack-bridging mechanism was modeled by a continuous distribution of linear springs acting between the crack faces (Fig. 3(b)); the crack-opening induced by the remote stress applied to the panel $[\sigma]$ is constrained by the crack-closure stress, bridging-stress $[\sigma_b]$, established at the crack faces by the linear springs.
The Rose's model is depicted in Eq. 1, where E is the elastic modulus of the skin and k a parameter representative of the spring stiffness (spring stiffness constant per unit force).
The σ_b formula in Eq. 1 describes the linear elastic behavior of the modeled springs, being the bridging-stress proportional to the COD, i.e. the

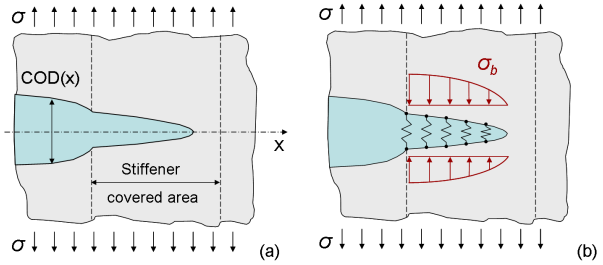


Figure 3: Skin COD restrained by the crack-bridging mechanism (a) modeled by a continuous distribution of springs acting between the crack faces in accordance with the Rose's model (b).

spring extension, and the spring stiffness $[kE]$.

$$COD(x) = f(\sigma - \sigma_b)$$
$$COD(x) = f\left(\sigma - kE\,\frac{COD(x)}{2}\right) \quad (1)$$

Equation 1, based on the dislocation theory developed by Bilby and Eshelby [8], is a differential-integral equation that must be numerically solved in the unknown COD.
The spring stiffness parameter $[k]$ is an index that quantifies the load transfer from the cracked skin to the bonded stiffener. It can be calculated in accordance with the single-strap joint elastic theory (Fig. 4 and Eq. 2) [3]. The in-plane stiffness of the strap joint depends on the displacement $[\delta_A]$ due to the adhesive shear-strain induced by the remote stress $[\sigma]$. In the likely presence of adhesive delamination $[del]$ extended around the advancing skin crack tip, the additional displacement $[\epsilon_R del]$ due to the reinforcement strain over the delaminated area must be considered.

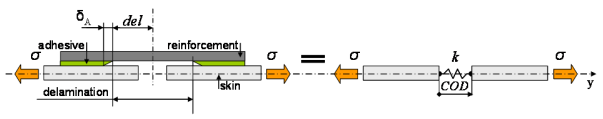


Figure 4: Sketch of a single strap-joint.

$$k = \frac{\sigma}{E_p\,COD} = \frac{\sigma}{E_p\,(\delta_a + \varepsilon_r\,del)} \quad (2)$$

Wider is the delamination and lower is the spring stiffness, thus reducing the effectiveness of the bonded stiffener in retarding the skin crack propagation.
In accordance with the Linear Elastic Fracture Mechanics (LEFM), the COD is directly related to the SIF by Eq. 3, being G and ν respectively the skin shear-modulus and Poisson's ratio and $[a]$ the semi-crack length.

$$SIF(a) = COD(x)G\left(\frac{1-\nu}{2}\right)\sqrt{\frac{2\pi}{a-x}} \quad (3)$$

The FCP rate $[da/dN]$ can be evaluated by means of the well-known Paris' law as a function of the calculated SIF range $[\Delta K]$(difference of the $[K]$ calculated at the maximum and minimum applied stress).

$$\frac{da}{dN} = C_p \Delta K^{m_p} \quad (4)$$

where $[C_p]$ and $[m_p]$ are skin material constants.

3 Experimental model validation

FCP tests results of wide adhesively bonded stiffened panels representative of typical wide-body fuselage applications were employed to validate the implemented analytical model [9].
The tested panels were constituted of a flat skin with seven equally-spaced adhesively bonded stringers, as sketched in Fig. 5. A through-thickness notch was machined orthogonally to the stringers in the middle of the panel; both the skin and the central stringer were cut. During the FCP tests, a fatigue load was applied to the panel orthogonally to the initial notch, promoting its straight propagation.
The test campaign was performed by a leading European aeronautical company. Different combinations of skin and stiffener materials and dimensions were investigated. In order to show the reliability of the implemented crack-bridging model, the three most significant test-analytical correlations are reported in the present work. 2524-T3, 1.6mm thick aluminum skin with seven identical 7349-T76511 Z-shaped bonded stringers panels are herein considered. The investigated panels differ only

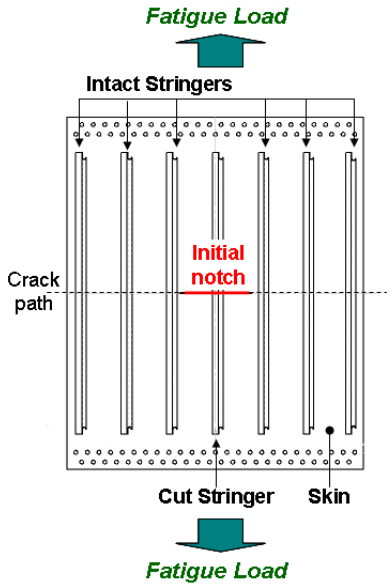


Figure 5: Sketch of the fatigue tested adhesively bonded stiffened panel.

for the stringer cross-sectional area (89mm^2, 151mm^2 and 207mm^2).
In Figures 6, 7, 8 the recorded FCP rate underneath the first stringer adjacent to the central cut one (see sketch in Fig. 2) are reported together with the calculated predictions. Different delamination extensions (*del*) around the advancing crack tip are taken into account, as well as, a delamination that fatigue grows together with the skin crack underneath the bonded stiffener.

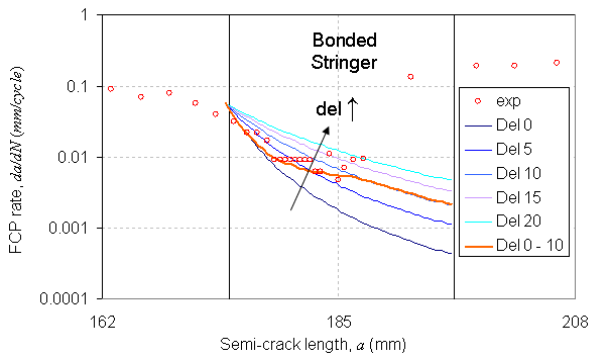


Figure 6: Analytical - test correlation for the tested panel with 89mm^2 cross-sectional area bonded stringers.

The analytical predictions show a good agree-

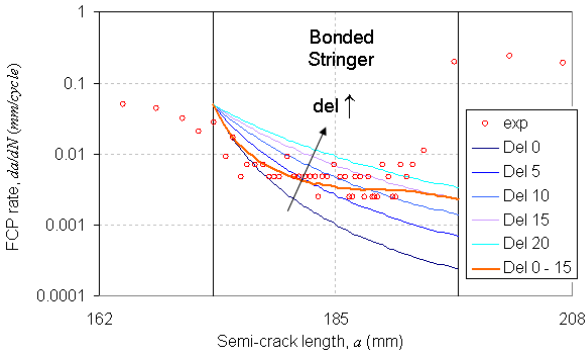


Figure 7: Analytical - test correlation for the tested panel with 151mm^2 cross-sectional area bonded stringers.

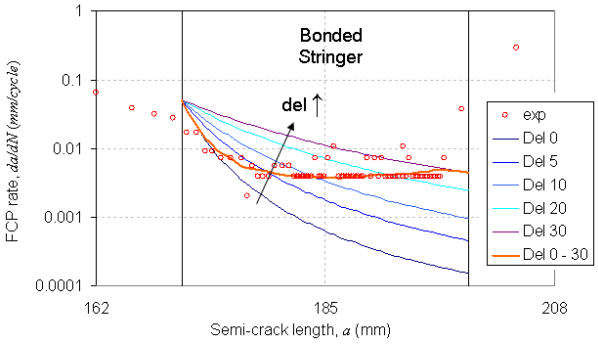


Figure 8: Analytical - test correlation for the tested panel with 207mm^2 cross-sectional area bonded stringers.

ment with the experimental results. The assumption of an adhesive delamination growth coupled with the fatigue skin crack propagation leads to the best agreement with the experimental data.

4 Conclusions

On the basis of a well-know literature, an analytical model was implemented to describe the fatigue skin crack propagation through a stiffened panel underneath adhesively bonded stiffeners. The test-analytical correlation showed the reliability of the implemented model. The integration of the implemented model with a tool developed by the authors for the damage tolerance assessment of stiffened panels based on the displacement compatibility method provides an effective tool for the preliminary design of bonded structures.

Acknowledgments

Grateful thanks are given to Nikolaus Ohrloff and Marco Pacchione (Airbus) for the experimental data and technical advices.

References

[1] Schijve J., "Crack stoppers and ARALL laminates", *Tech. Report LR-589, Delft University of Technology* Delft, Holland, 1989.

[2] Schmidt H.J., "Damage Tolerance Technology for current and future aircraft structure", *Plantema Lecture 23rd ICAF Symposium*, Dalle Donne C. Ed., Hamburg, Germany, 2005, pp. 1-41.

[3] Rose L.R.F., "Theoretical analysis of crack patching", *Bonded repair of aircraft structures*, Baker A.A. and Jones R. Ed., Martinus Nijhoff Publ., New York, USA, 1988, Chap. 5, pp. 77-105.

[4] Poe C.C., "Stress intensity factor for a cracked sheet with riveted and uniformly spaced stringers", *NASA Technical Report, NASA TR R-358*, Langley Research Center, Hampton, Va., USA, 1971.

[5] Poe C.C., "The effect of broken stringers on the stress intensity factor for a uniformly stiffened sheet containing a crack", *NASA Technical Memorandum, NASA-TM-X-71947*, Langley Research Center, Hampton, Va., USA, 1973.

[6] Molinari G., Meneghin I., Melega M., Troiani E., "Parametric damage tolerant design of advanced aeronautical stiffened panels by means of LEAF analytical tool", *Proceedings 2nd Aircraft Structural Design Conference RAeS2010 (Royal Aeronautical Society)*, London, UK, 2010.

[7] Bao G., Suo Z., "Remarks on crack-bridging concepts", *Applied Mechanics Review, ASME Book, American Society of Mechanical Engineers*, 1992, vol. 45, pp. 355-366.

[8] Bilby B.A., Eshelby J.D., "Dislocations and the theory of fracture", *Fracture: An advanced treatise (ed.) H Liebowitz (New York: Academic Press)*, 1968, vol. 1, pp. 100.

[9] Meneghin I., Pacchione M., Vermeer P., "Experimental Investigation on Bonded Structures for increased Damage Tolerant Aircraft Structures", *Proceedings ICAF2009 conference (International Committee on Aeronautical Fatigue)*, Rotterdam, Holland, 2009.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Large Attitude Maneuvers of a Spacecraft with Flexible Solar Panels

E. Paolini, F. Curti and M. Parisse
School of Aerospace Engineering, University of Rome La Sapienza, Italy

M. Battilana
CGS S.p.A., an OHB Technology Company, Milan, Italy

Keywords: *attitude, control, flexibility, thrusters, reaction wheels.*

Abstract

This paper faces with the problem of the attitude maneuver control of a spacecraft with large solar panels, actuated by thrusters and reaction wheels. A functional separation has been performed: thrusters are used for attitude maneuvers with a feed forward approach, while reaction wheels are used in closed loop to refine the maneuver. Thrusters allocation is optimized to increase the resulting torque in order to reduce the fuel consumption for the maneuvers. Since the attitude dynamics is influenced by the vibrations due to the flexibility of the solar panels, this effect is taken into account: the attitude dynamics equations are derived using a Lagrangian approach to include the flexibility of the arrays. An optimal control has been implemented in order to reduce fuel consumption while maneuvering the satellite. Numerical simulations have been run to evaluate the AOCS performance, during the large attitude maneuvers, and the pointing accuracy. Results are presented.

1 Introduction

Nowadays many satellites require the ability to perform large angle maneuvers in a fixed time, or in some cases even in minimum time. The problem is often approached trying to simplify the spacecraft dynamics. Therefore deep studies on the rigid body control have been developed for large angle maneuvers [1,2]. Minor efforts have been focused on the control of rigid bodies with flexible appendages. However the problem became of large interest in the mid 90s. Moreover, three axes maneuvers are usually performed by means of a sequence of single-axis rotations. Such maneuvers may take a relatively long time, and their duration can be reduced considerably by performing simultaneous three axes slews. This work deals with the three-axes slew of a spacecraft with flexible appendages. This problem is not only infinite dimensional for the arrays flexibility, but also highly non-linear. Works facing with these problems are for example [3,4,5]. However, the problem can be simplified, considering a discrete parameters model.

The objective of the spacecraft under analysis is the observation of deep space for astronomy purposes from a LEO orbit.
A layout of the spacecraft with its body reference frame, centred in the satellite centre

of mass (CoM), is depicted in Fig. 1. The line of sight of the instruments in the payload module is along X_{sat}; the panels deploy along Y_{sat}. The sun direction shall be kept always in the $-X_{sat}$, Z_{sat} plane. The panels can rotate around Y_{sat} to improve the angle of incidence of the sun rays. Each panel is 1.2 m wide and 5 m long. Table 1 shows the spacecraft dimensions and the estimated moments of inertia: the body frame depicted in Fig. 1 is supposed to coincide with the principal axes reference frame, thus inertia matrix is diagonal. The position of the centre of mass is supposed to lie along X_{sat}; its position varies from 1266.6 mm from the -X_{sat} face at BOL to 1353.5 mm at EOL, ensuring the CoM to lie always within the platform module.

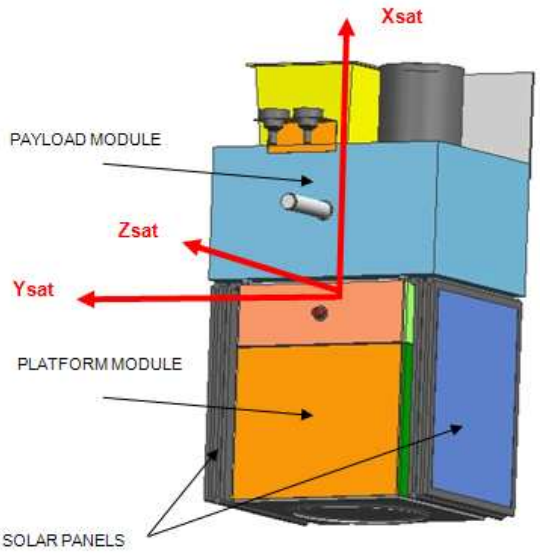


Fig. 1 Spacecraft layout and body frame definition.

Table 1 Spacecraft dimensions and moments of inertia.

PLATFORM X LENGTH [mm]	**1445**
PLATFORM Y LENGTH [mm]	**1240**
PLATFORM Z LENGTH [mm]	**1240**
PAYLOAD X LENGTH [mm]	**1505**
PAYLOAD Y LENGTH [mm]	**1650**
PAYLOAD Z LENGTH [mm]	**1274**
XX INERTIA [kgm^2]	**400**
YY INERTIA [kgm^2]	**1000**
ZZ INERTIA [kgm^2]	**1200**
SOLAR PANEL MASS [kg]	**30**

The axis of the line of sight of all instruments in the payload module (X_{sat}) shall be rotated for astronomical observations. Considering the maximum field of view of the instruments, the amplitude of the maneuver is $\pm 45°$ around Y_{sat} and Z_{sat}. In addition, the attitude control shall keep the sun direction always in the $-X_{sat}$, Z_{sat} plane; thus a third rotation, around X_{sat}, is required for re-pointing the sun vector in the desired direction. The amplitude of this rotation is calculated applying the rotation matrix to the initial sun vector, and imposing the final sun vector to have a null second component, a positive first one and a negative third component. The result is that the maximum span for the maneuver around X_{sat} is $\pm 90°$.

The maximum time allowed for the maneuver is 360 s. The accuracy required for the manuever is 0.01° along Y_{sat} and Z_{sat}, 0.05° along X_{sat}. For this accuracy it is necessary to consider the flexibility of the arrays. Generally, in fact, the necessity to simulate spacecraft flexibility depends on the level of accuracy required to ACS, because *every spacecraft is flexible to some extent* [6].

2 Dynamics model

A discrete parameters approach has been selected for its better interaction with the Simulink environment. Fig. 2 shows the model of the flexible spacecraft. Each panel is considered as a cantilever beam of length *l*, with a spring of stiffness *K* at the root and a mass *m* at the tip. In this way the first mode of vibration is considered: α represents the angle of deflection of the solar arrays, assumed in opposite verse for each panel.

The analytical mechanics approach has been utilized, finding the equations of motion by means of the Lagrange equations. The Lagrangian problem is formulated in terms of quasi-coordinates [7], thus it takes the form in Eq. (1):

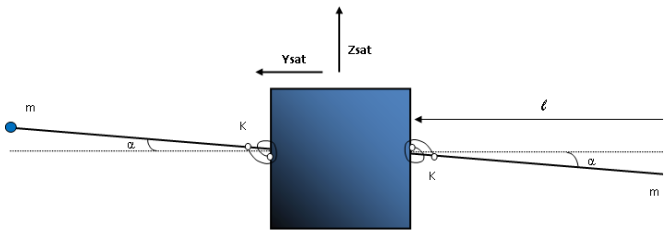


Fig. 2 Solar arrays discrete parameters schematization.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \omega}\right) + [\tilde{\omega}]\left(\frac{\partial L}{\partial \omega}\right) = Q \qquad (1)$$

L is the Lagrangian function, Q the torque vector, ω the angular rate vector in body coordinates and $\tilde{\omega}$ the skew-symmetric matrix representing the cross product. The equation of the angle α can be expressed by means of the usual Lagrange equation, adding the presence of a dissipation function D, that acts as a natural damping for the system, represented by the Rayleigh damping function [7,8] as in Eq. (2):

$$D = 2\left(\frac{l^2\dot{\alpha}^2}{2}\right)K_d \qquad (2)$$

K_d is the natural damping factor of the system (in kg/s). Thus the equation for the panels deflection can be written as in Eq. (3):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\alpha}}\right) - \left(\frac{\partial L}{\partial \alpha}\right) + \left(\frac{\partial D}{\partial \dot{\alpha}}\right) = 0 \qquad (3)$$

The Lagrangian function L is the difference between the kinetic energy and the potential energy, that are expressed in this case as in Eq. (4) [9]:

$$T = \frac{1}{2}\vec{\omega}^T J\vec{\omega} + \sum_{i=1}^{4}\frac{1}{2}I_s{\omega_{ri}}^2 + \sum_{i=1}^{4} I_s\omega_{ri}\vec{\omega}\cdot\hat{a}_i + \frac{1}{2}m\left(\vec{V}_1\cdot\vec{V}_1 + \vec{V}_2\cdot\vec{V}_2\right) + I_p\dot{\gamma}^2 \qquad (4)$$

$$U = \frac{1}{2}K(l^2\alpha^2 + l^2\alpha^2) = Kl^2\alpha^2$$

$\vec{\omega}_{ri}$ represents the i-th wheel angular rate. $\vec{V}_1$ and $\vec{V}_2$ are the linear velocities of the two masses, $J = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix}$ is the matrix of inertia of the rigid body in body coordinates, not including the presence of the appendages; the body frame is assumed to be the principal frame of inertia, and this implies J to be diagonal. I_p is the beamlike-panel moment of inertia and I_s the wheel polar moment of inertia. $\hat{a}_i$ is the unit vector that expresses the orientation of the i-th wheel axis in the body frame.

In this way it is possible to obtain the equations governing the system: three for the angular rates and one for the deflection. To avoid undesired algebraic loops the system is put in a form as in Eq. (5):

$$T\begin{bmatrix}\dot{p}\\ \dot{q}\\ \dot{r}\end{bmatrix} = -\vec{\omega}\times T\vec{\omega} + U + \tau \qquad (5)$$

where T is the inertia matrix of the entire system, U the control vector and τ the term containing all the coupling terms due to the flexibility. Quaternions kinematics' equations are integrated to compute the spacecraft attitude. The vibration frequency of the panel is evaluated to be 2.2 Hz.

3 Thrusters configuration

The necessity to make maneuvers in relatively short times led to the idea of using reaction thrusters. The selected hardware is an on-off hydrazine-based thruster able to produce 1N thrusts with a rise time of 60 ms, a decay time of 150 ms and a MIB of 2 ms. Thrusters can be accommodated only in the platform module, and in such a way that there is no firing against the payload. The more natural position for thrusters is clearly in the corners of the platform module, so that the arms are maximized, allowing to save fuel. Moreover, if two thrusters are placed in the opposite

direction with respect to the CoM and fired together, the arising torque is independent from the CoM movements due to the reduction of the fuel n the tanks during the spacecraft operative life. An 8 thrusters configuration with a redundancy around *X* has the advantage that it can ensure a higher torque for one and two axes maneuvers, but on the other hand it cannot give a perfect maneuverability for three axes maneuvers. In fact if a thruster gives torques along more than one axis, it is not possible to obtain the desired components of the control torque in body coordinates. Therefore a 12 thrusters configuration has been adopted, as in Fig. 3. This accommodation allows to control the torque along each body axis independently, thus to obtain any of the possible torque components in space.

Once the configuration is chosen, another step can be done in order to minimize the fuel consumption: the thrust directions trade-off. The simple base concept is that the more the thrust direction is tilted with respect to the normal to the satellite, the longer the arm is, with a consequent gain in fuel consumption, until an optimal condition is reached. From the optimal point, continuing inclining the thrust direction would result in a shorter arm, thus in a worst configuration. Obviously the tilt angle shall be the same for each couple of opposite thrusters, to have a null translation when they are actuated. To evaluate the optimal tilt angle, it has been evaluated the torque as function of this angle, as shown in Fig. 4 for *Mx*, the torque around *X*. The curve presents a maximum, corresponding to the optimal tilt angle. The graph is parameterized with respect to the margin to be taken from the satellite borders for thrusters mounting. The analysis demonstrates that considering a margin for the mounting of 20 cm, the optimal tilt angle of the thrusters that provide torques around *X* is 55.9°, while that for thrusters that give torques around *Y* is 49.8°. The thrusters that provide torques around *Z* have been not tilted to avoid plume impingement on the solar arrays.

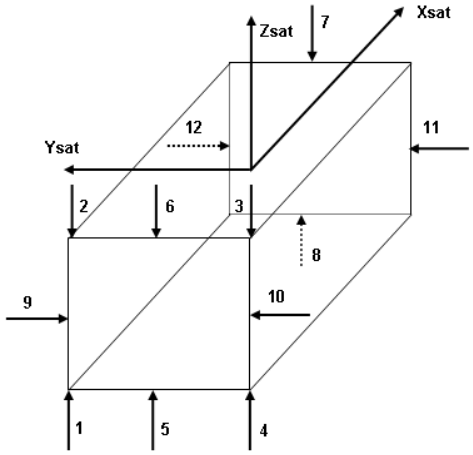


Fig. 3 Thrusters accommodation on the platform module.

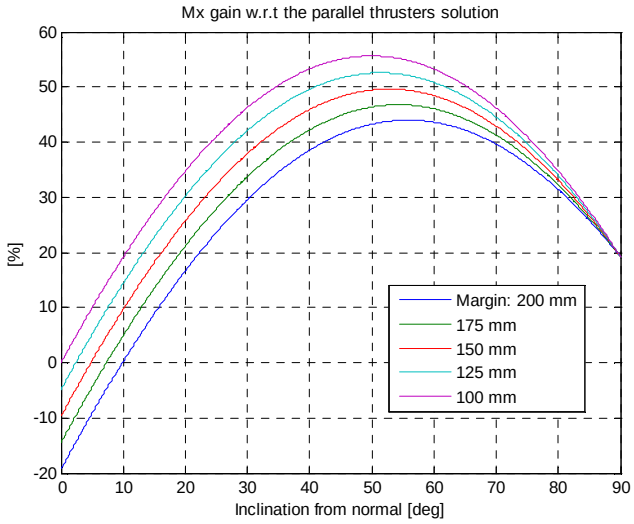


Fig. 4 Mx gain in torque w.r.t the not-inclined thrust directions solution.

4 Control

The complete maneuver has been split into two phases: the *maneuver* and the *fine pointing*. The *maneuver* phase aims to perform the initial part of the maneuver control: its function is to reduce the error between the current attitude and the target attitude beneath a certain threshold. The *fine pointing* phase aims to bring the attitude from the threshold error to the final target, in compliance with the requirements on the accuracy. The schemes for the two phases are shown in Fig. 5. The *maneuver* phase is the part in which the major

effort is required to the ACS hardware, because more than the 95% of the starting attitude error, from the starting condition to the switching condition, is compensated in this phase. In the *fine pointing* phase the attitude error is small: this is a refinement of the maneuver made in the previous phase, and therefore thrusters cannot be useful for intrinsic problems related to their low accuracy. In general, the purely open-loop maneuvers do not require any measurement for feedback, thus there is no possibility of closed-loop instability; nevertheless, in practice, the open-loop schemes are sensitive to the parameters uncertainties and the unexpected disturbances. Therefore a combination of feedforward and feedback control is desirable. For this reason, in the *maneuver* phase thrusters are used in open-loop to provide the change in angular rates needed to perform the maneuver, while reaction wheels are used in closed-loop to compensate disturbances. A slewing guidance gives intermediate targets to the reaction wheels at the controller working frequency. Instead, in the *fine pointing* state a fixed guidance towards the final target is used. An important aspect to consider in the framework of this separation of the controller into two different phases is the criterion to automatically switch from one phase to the other. Different solutions can be taken into account: the chosen one is such that when the *fine pointing* phase is reached, reaction wheels are able to bring attitude within the accuracy in less than 60 s. It is reasonable to admit that the controller can also switch backward, from the *fine pointing* phase to the *maneuver* phase, for example if external disturbances increase the absolute value of the speed and/or the attitude error beyond the limits. Therefore the implementation in the simulator shall grant a two-ways switch.

The chosen control law for the reaction wheels is a PD control known as quaternion feedback [9], and has the form in Eq. (6):

$$\begin{aligned} T_{cx} &= 2K_{p1}q_1q_4 + K_{d1}\omega_1 \\ T_{cy} &= 2K_{p2}q_2q_4 + K_{d2}\omega_2 \\ T_{cz} &= 2K_{p3}q_3q_4 + K_{d3}\omega_3 \end{aligned} \qquad (6)$$

where T_{ci} is the control torque along the i-th body-axis, q_i the components of the error quaternion and ω_i the angular rate error components. K_p and K_d are the proportional and derivative gain matrices. The scalar part of the quaternion, q_4, multiplies the first terms to make the control law robust to its changes in sign, as suggested in [8].

The open-loop control for thrusters is an optimal control that aims to minimize the cost function in Eq. (7).

$$\min J = \int_{t_0}^{t_f} \sum_{i=1}^{3} u_i C u_i^T \, dt + \widetilde{X_f} S \widetilde{X_f}^T \qquad (7)$$

$$\begin{aligned} &subject\ to: \quad \dot{x} = f(x,t) + Bu\,, \quad |u_i| \le 1 \\ &with\ x(t_0) = x_0\ and\ x(t_f) = x_f \end{aligned}$$

where $\widetilde{X_f}$ is the error in the final state, t_f is the final time and *C* and *S* the weight matrices. The first term represents the minimization of the propellant consumption, while the second term the minimization of the error on the final state with respect to the target.

However, this solution can be found analytically only for simple problems, for example for the minimum fuel maneuvers of rigid bodies. The dynamics considered in this analysis is strongly non linear for the presence of the solar arrays, and therefore the analytical approach cannot be used. As a consequence, a solution via a numerical approach has been considered. The Matlab function *fmincon* is able to solve the minimization of cost functions in which the states follow a non-linear dynamics, and in presence of constraints.

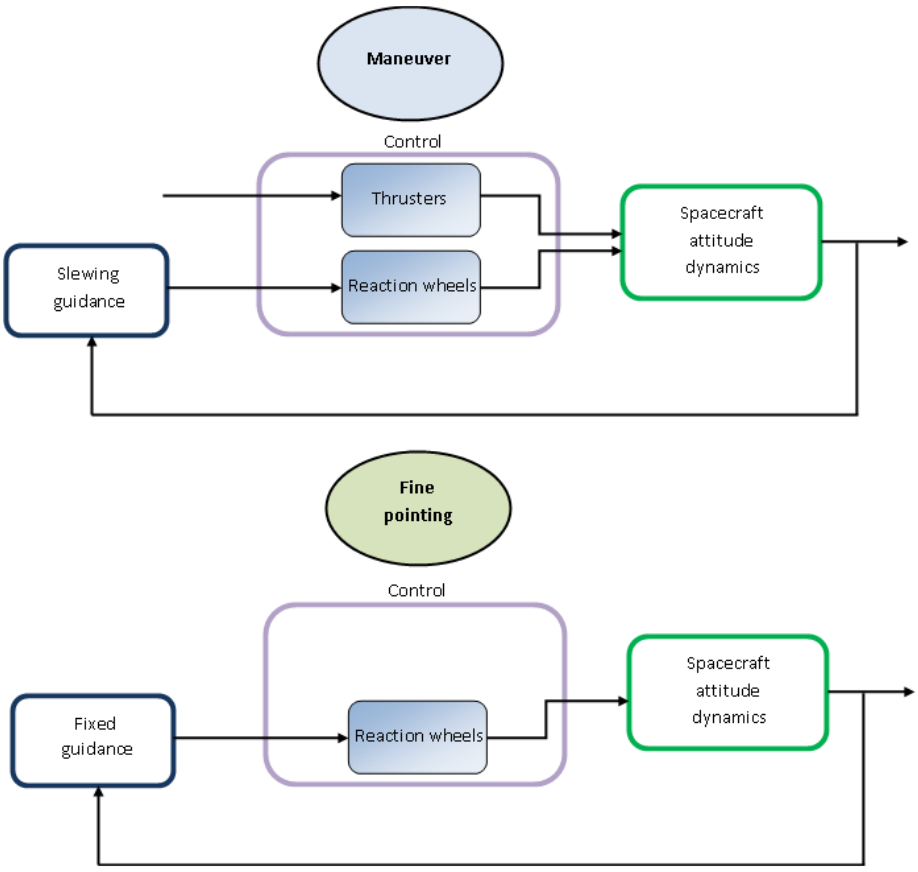


Fig. 5 Control architecture for the two states.

To help the software in solving this not easy problem, two important assumptions have been made:

- the control profile is considered to be piecewise constant: this reduces the number of variables.
- The control profile is considered to be an odd function, that is, anti-symmetric with respect to the half time of the thrusters' firing.

The second assumption is more realistic than it can seem at a first glance: in fact it ensures to have a null angular momentum at the end of the thrusters activation, condition that should be fulfilled by a well-designed maneuver control, while holding the generality of the solution. However these two assumptions surely reduce the search for a minimum to a local optimality problem, that is the found solution will be the best among the possible solutions defined by the two above conditions.
The optimizer software has to work off-line, because the knowledge of the final state is necessary to the computation of the optimal solution. It simulates the dynamics of the spacecraft for the entire time in which thrusters give their contribution. At each step of the optimization process the spacecraft dynamics is simulated, taking into account only the thrusters contribution. Therefore *reaction wheels are not considered in the optimization process*. This ensures that the result is an optimum thrusters contribution, without the help of the wheels. The only simplification made to the simulated dynamics is the *neglecting of the gyroscopic torque*. In a minimum fuel logic, it is reasonable that thrusters have to provide the minimum possible effort, while the reaction wheels have to take care of everything they are able to. In this framework, reaction wheels can therefore compensate for the gyroscopic torque, which is in the order of $1 \cdot 10^{-2}$ Nm, absolutely in the range of the wheels torques. Thus, the slewing guidance of the wheels is based on a dynamics that neglects gyroscopic contribution, and when the real dynamics is simulated, wheels interpret the gyroscopic torque as a disturbance, and act to compensate for it. The compensation is well performed; the error on the final state between a simulation neglecting the gyroscopic torque and a simulation considering it, with the reaction wheels in the control loop, is in the order of $1 \cdot 10^{-2°}$ in eigen-angle.

A Pulse Width Modulation (PWM) technique is used to modulate the thrust of the on-off thrusters. Also this effect is neglected in the optimization process, where adjustable torques are considered as provided from the thrusters. As for the gyroscopic torque, the PWM effect is compensated by reaction wheels in the real dynamics simulation. This avoids numerical problems created by the presence of the PWM in the optimization process.

In the minimum fuel consumption problem, the spacecraft dynamics is simulated at each step of the optimization process. Thus the state profile during the entire firing time interval using the optimal control is known at the end of the process. This profile represents the state trajectory if only thrusters are used, and therefore it can be used as a slewing guidance for the reaction wheels in the simulation of the "real" dynamics.

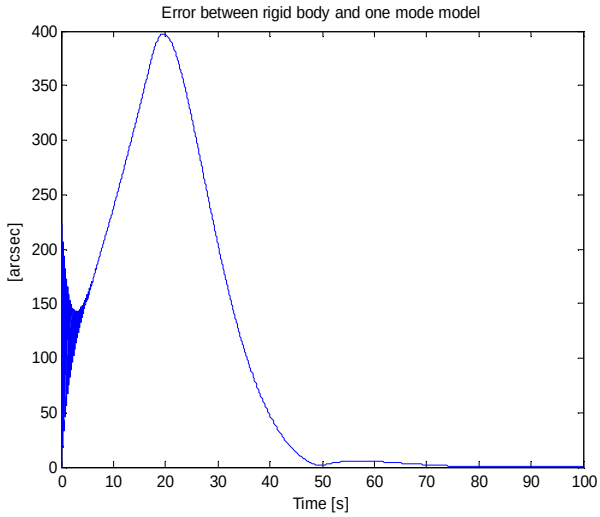


Fig. 6 Difference in eigen-angle between the rigid body and the flexible body models, in arcsec.

5 Simulations and results

The model of the flexible attitude dynamics described in section 2 has been implemented in Matlab Simulink environment. The model has been tested to verify its correct functioning. Several characteristic cases have been performed: rigid body, harmonic oscillator, spinning body, body at rest with oscillating panels, etc… The results of the simulations are identical to the expected behavior in all the cases. Therefore the model is considered validated. Details of the test campaign are in [10].

A comparison of the model with a simple one representing the attitude dynamics of a rigid body demonstrates that the difference is not negligible during the fine pointing phase, as it is shown in Fig. 6. The simulation starts from initial conditions for attitude and speed of switch to the *fine pointing* phase, and initial deflection of the arrays of 0.05°. The system is forced to reach zero by the reaction wheels. It is to notice that the difference between the attitude of the two models, expressed in eigen-angle, is more than 10 times the required accuracy of 36 arcsec during fine pointing.

One axis maneuvers are examined first. The worst-case maneuver of this kind is a 90° rotation around *X*. The control profile obtained from the optimization process is shown in Fig. 7. Thrusters fire for the first 200 s. Defining the fuel consumption as the integral over time of the sum of the modules of the control variables (being *u* dimensionless the result is in seconds), the consumption for this control profile is 30.1 s, which corresponds to a hydrazine consumption of 40 grams considering a specific impulse of 150 s. The control profile obtained from the optimization process has been used then for the simulation of the entire maneuver, resulting in the states profiles shown in Fig. 8 and Fig. 9. 8° from the final target are reached in the first 200 s: as expected, reaction wheels have perfectly compensated for gyroscopic torque and PWM effects. After the thrusters contribution, the *maneuver* phase is completed by the reaction wheels, until the switch conditions are reached and the *fine pointing* phase is performed. The maneuver reachesa the final target within the accuracy nd widely within the required time, as shown in Fig. 10.

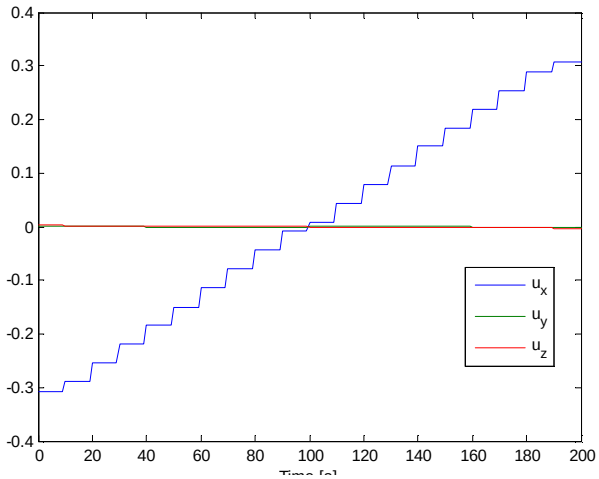


Fig. 7 Control profiles for a (90°,0,0) maneuver.

The array deflection is reported in Fig. 11. It is clear that the frequent burns due to the Pulse Width Modulation excite continuously the panels, resulting in wide oscillations for the first 200 s. While thrusters are firing in the first 200 s, wheels have to compensate for gyroscopic torque and PWM effects. The maximum speed peak reached by the wheels is of 2500 rpm, which is in the acceptable range for the selected hardware.

Results for the other one axis maneuvers are analogous, with a saving in time in the order of 50-60 s with respect to the maximum allowed time for reaching the desired accuracies, and with a consumption of no more than 35 s.

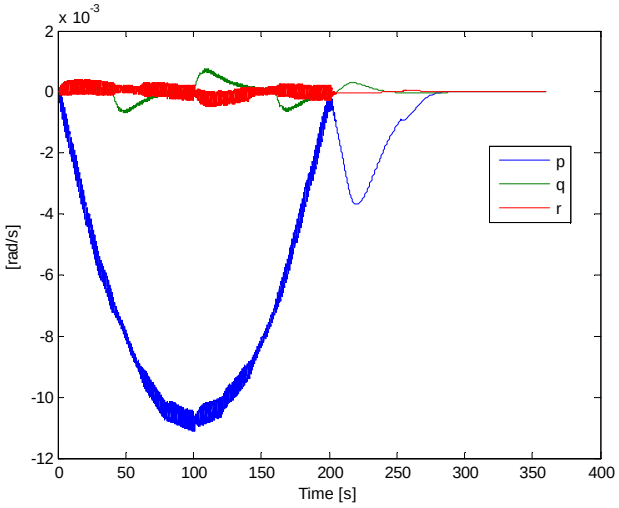


Fig. 8 Angular rates profiles for a (90°,0,0) maneuver.

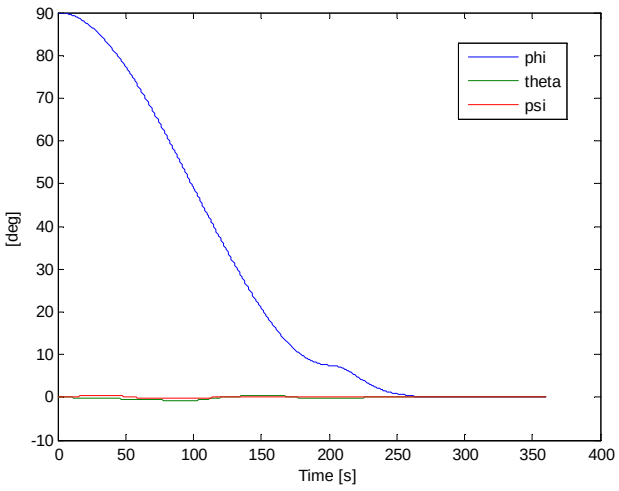


Fig. 9 Attitude profiles for a (90°,0,0) maneuver.

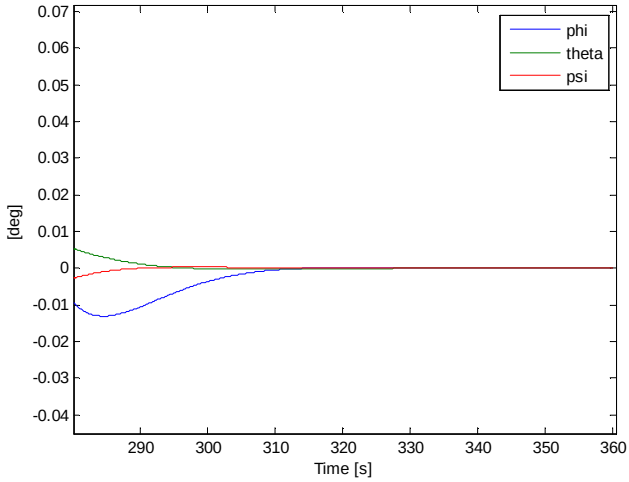


Fig. 10 Accuracy is reached in less than 290 s.

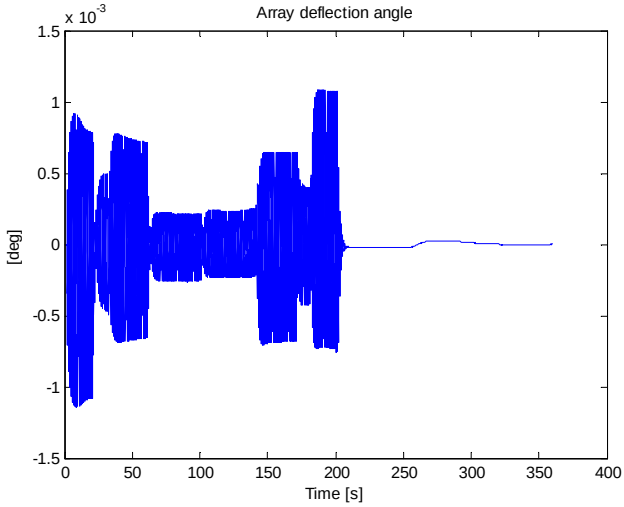


Fig. 11 Array deflection angle during maneuver.

The worst-case for three axes maneuvers is a (-90°,45°,45°) slew. In this case the control profile obtained from the optimization process is that in Fig. 12. This control ensures a consumption of 51.4 s, 69 g of hydrazine. The attitude profile obtained for the entire maneuver is depicted in Fig. 13.
The maneuver is correctly performed: the attitude reaches the target in the fixed time within the required accuracy. The accuracy range is reached 40 s before the fixed final time.

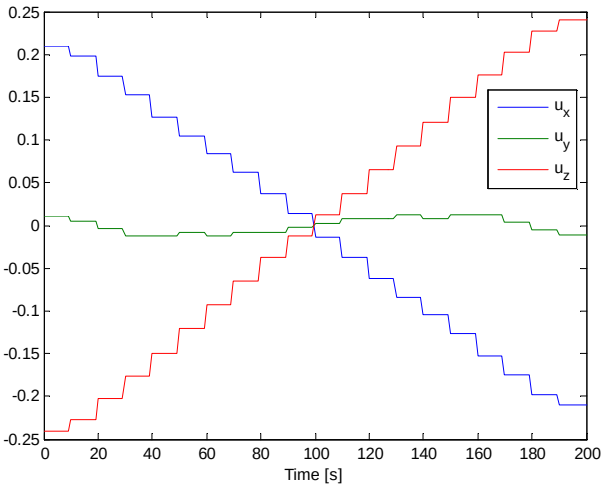


Fig. 12 Control profiles for a (-90°,45°,45°) maneuver.

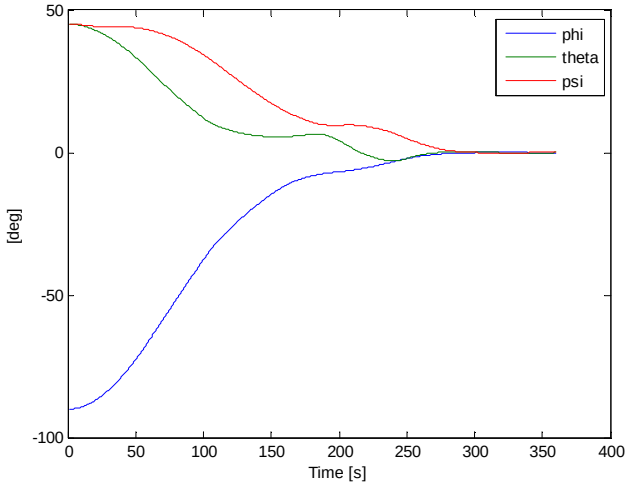


Fig. 13 Attitude profiles for a (-90°,45°,45°) maneuver.

Table 2 summarizes the obtained results, giving evidence of the behavior of the developed control strategy with the maneuver requirements.

Table 2 Results obtained with the developed control strategy.

ONE AXIS MANEUVERS	
Error after 360 s	$1 \cdot 10^{-5}$°
Time to reach the accuracy range	290-328 s
Fuel consumption	30.1-35.2 s
THREE AXES MANEUVERS	
Error after 360 s	$1 \cdot 10^{-4}$°
Time to reach the accuracy range	320 s
Fuel consumption	51.4 s

6 Conclusion

The analytical model of the flexible dynamics has been developed, implemented and fully validated. Thrusters have been accommodated on board the platform module in order to reduce fuel consumption, ensuring at the same time sufficient thrusts for the maneuvers. Thrust directions have been tilted with respect to the satellite faces in order to increase the available torques, and optimal tilt angles have been found numerically.
The control strategy has been developed to obtain a minimum energy-minimum error strategy, that is a control that brings the attitude close to the target, saving the maximum possible amount of fuel. Since the problem was not solvable analytically, a numerical approach has been used. An optimizer algorithm has been developed using the Matlab function *fmincon*. At the end of the optimization process the optimum control profile is obtained and utilized for a simulation of the entire maneuver. Tests demonstrate that the considered control strategy allows performing the worst-case maneuvers, ensuring to reach the final target within the required accuracy and time. After 360 s, the maximum allowed time for the maneuver, the attitude error with respect to the target is always below $1 \cdot 10^{-4}$°, that is two

orders of magnitudes less with respect to the required accuracy. The range of admissible error – the accuracy range – is always reached in 320-330 s, saving 40 s w.r.t the maximum allowed time in the worst-case three axes maneuver, and even 70 s for the worst-case one axis maneuver around X_{sat}. Propellant consumption is far inferior to that of a not-optimal maneuver.

References

[1] Vadali S.R., Carter M.T., Singh T., Abhyankar N.S., "Near-minimum-time maneuvers of large structures: theory and experiments", *Journal of Guidance, Control and Dynamics*, Vol. 18, No. 6, 1995.

[2] Hall C.D., Tsiotras P., Shen H., "Tracking rigid body motion using thrusters and reaction wheels", AIAA 98-4471.

[3] Singh G., Kabamba P.T., McClamroch N.H., "Planar, time-opimal, rest-to-rest slewing maneuvers of flexible spacecraft", *Journal of Guidance*, Vol. 12, No. 1, 1989.

[4] Ebrahimi A., Moosavian S.A.A., Mirshams M., "Control of space platforms with flexible links using command shaping method", *Iranian Journal of Science & Technology*, Vol. 32, No. B1, 2008.

[5] Anthony T.C., Wie B., Carroll S., "Pulse-modulated control synthesis for a flexible spacercaft", *Journal of Guidance*, Vol. 13, No. 6, 1990.

[6] Wertz J.R., *Spacecraft Attitude Determination and Control*, Kluwer Academic Publishers, 1978.

[7] Meirovitch L., *Methods of analytical dynamics*, McGraw-Hill.

[8] Sidi M.J., *Spacecraft dynamics and control. A practical engineering approach*, Cambridge University Press, 1997.

[9] Wie B., Barba P.M., "Quaternion feedback for spacecraft large angle maneuvers", *Journal of Guidance*, Vol. 8, No. 3, 1985.

[10] Paolini E., "Large attitude maneuvers of a spacecraft with flexible appendages", Master Degree Dissertation in Astronautical Engineering, School of Aerospace Engineering, University of Rome La Sapienza, Italy, 2011.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Fast and Accurate Numerical Integration of Relative Motion in Spacecraft Formation Flying

G. Baù and A. Valmorbida
Phd Candidates, University of Padova, CISAS "G. Colombo", Padua, Italy

E. Lorenzini
Professor, University of Padova, CISAS "G. Colombo", Padua, Italy

A. Francesconi
Research Fellow, University of Padova, CISAS "G. Colombo", Padua, Italy

Keywords: orbit propagation, formation flying, regularized special perturbation methods

Abstract

The special perturbation method DROMO developed by Peláez in 2006 for the perturbed two-body problem is employed to propagate the relative motion in spacecraft formation flying, and the performance of the new method, named DROMO-FF, is analyzed. DROMO is a very fast and accurate regularized method which involves a set of seven integrals of the pure Keplerian motion whose physical meaning is described in the paper. We propose to propagate the absolute motion of n spacecraft simultaneously by using DROMO with the introduction of new dependent variables, necessary for the synchronization, and to determine the relative dynamics by differentiating the absolute states. After investigating the influence on the performance due to the numerical integration of the new variables, we show that DROMO-FF is significantly more accurate than Cowell's method for the same computing time, or equivalently, faster for the same accuracy. A second approach to propagate relative motion wherein linearization is performed with respect to the formation baricenter is presented and compared to DROMO-FF. It is shown that for closed formations round-off does not affect the accuracy of DROMO-FF.

1 Introduction

The study of relative motion between satellites is of great importance in formation flying technology, which consists of a cluster of spacecraft flying in formation. Usually the relative motion of the generic spacecraft, called "chaser", is expressed in an orbital reference frame moving with a reference spacecraft, called "target". The first model of relative motion was developed by Hill [8] in his work on the lunar theory. Clohessy and Wilthshire [5] (CW) adapted the Hill's equations to the problem of satellite rendezvous in the 1960s. The Hill-CW equations of relative motion were derived assuming that the distance between the target and the chaser is small compared with the radius of the target circular orbit. These equations can be analitically integrated if the motion is supposed to be unperturbed, and their solutions are used to obtain a rough prediction of the relative dynamics for short time spans. Tschauner and Hempel [12] first found an analytical solution for the case of elliptical reference orbits. Many other formulations of the state transition matrix for this case have been produced, like those proposed by Melton [9] and Yamanaka Ankersen [13] for an elliptical orbit of arbitrary eccentricity, or that proposed by Broucke [4].

Analytical solutions to the problem of relative motion provide a tool for the preliminary design of a formation flight. Besides, thanks to the high speed of their computation, they are preferred to complete

models, which require numerical integration, in an on-board computer to predict the relative position and velocity with respect to an other spacecraft. However, the big drawback of analytical formulas is that they lose accuracy with time because of the limiting hypothesis they are based on.

Equations of motion derived by the variation of parameters (VOP) method are effective in dealing with the long-term study of the motion of asteroids, comets, as well as natural and artificial satellites subject to relatively small perturbation forces, mainly because, unlike the methods formulated in rectangular coordinates, describe the evolution of elements (or integrals of the motion), which exhibit no error propagation with respect to the unperturbed two-body problem. The VOP can be further improved introducing a two-body regularization which cancels out the dependence on the eccentricity of the propagation error of the physical time in the pure Keplerian motion.

A different strategy towards the improvement of a VOP method has been followed by Peláez et al. In a recent article [10] Peláez has proposed a new formulation for the two-body problem, borrowing elements of rigid-body dynamics. The idea is to track the evolution of an orbital reference frame moving with the particle and link a new set of generalized orbital elements to this frame. The result is a propagator, called DROMO, that shows an improvement in accuracy (or, equivalently, a reduction of computing time) with respect to other efficient regularized methods and a more compact and simple formulation of the equations of motion.

In a recent work [1] Baù applied DROMO to spacecraft formation flying. Two approaches were proposed. In the first, model, named here DROMO-FF, the method is applied to each spacecraft simultaneously by mean of new dependent variables introduced to achieve the synchronization. From the absolute dynamic states, the relative dynamics is determined. In the second approach, named here DROMO-G, after performing a Taylor's expansion of the gravitational terms, the method is used to propagate the two-body motion of the baricenter of the formation, which is chosen as the reference target point. The equations of the relative motion with respect to the baricenter are expressed in an orbital reference frame with inertial attitude.

In this article we analyse in detail the performance, in terms of accuracy and speed of computation, achieved by adopting DROMO-FF instead of the usual Cowell's method for the problem of prapagating the relative motion of a spacecraft formation flying. The main goal is to give evidence that an orbit propagator of this kind is suited to be employed for an on-board relative orbit prediction, embedded in a Guidance Navigation and Control (GNC) system. Besides, for ground simulations, it is important that the method is highly accurate in long-term integrations, in order to test via software the robustness and reliability of a system of sensors and of the GNC system. Actually, we have a software simulator for spacecraft formation flying applications, which is organized in the following blocks: a) dynamics propagator; b) guidance block that computes the desired trajectory to be followed; c) navigation block that implements a state observer based on Kalman filtering to obtain an optimal estimation of the current system state using a model of the system and measurements from sensors; d) control block that computes optimal control action as input for the actuators in order to track the desired trajectory; e) sensors and actuators blocks that model sensors and actuators, respectively, taking into account noise, resolution and biases.

2 DROMO special perturbation method

The starting idea of DROMO special perturbation method (see [10]) is the decomposition of the position vector of a point mass $\mathbf{x}$ into the product of its magnitude $R = \|\mathbf{x}\|$ and its direction $\mathbf{i} = \mathbf{x}/R$. The decomposition in the projective coordinates $(R,\, \mathbf{i})$, so-called after Ferrándiz, is a preliminary operation in obtaining a set of linearized equations of motion (Deprit [6]).

Non-dimesionalization is performed and the differential equations with respect to non-dimensional time τ are derived for:

1. $z = 1/r$, where r is the non-dimensional orbital radius;

2. $u = v_r$, where v_r is the non-dimensional radial velocity;

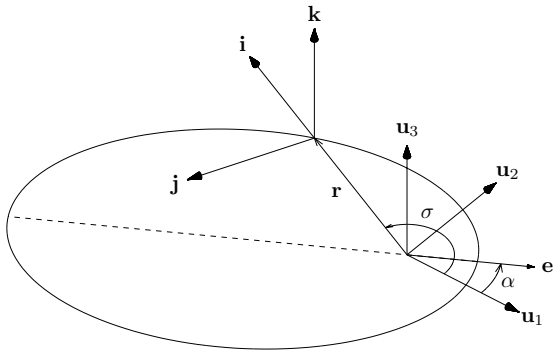


Figure 1 – Orbital reference frames $\mathcal{R}$ and $\mathcal{U}_0$ for the generic osculating orbit. The dash line represents the apse line of the osculating ellipse.

3. $\psi = h$, where h is the non-dimensional specific angular momentum;

4. the components of the unit quaternion $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ related to the orbital refrence frame $\mathcal{R} = \langle \mathbf{i}, \mathbf{j}, \mathbf{k} \rangle$, being[1]:

$$\mathbf{i} = \frac{\mathbf{x}}{R}\,, \quad \mathbf{k} = \frac{\mathbf{x} \times \mathbf{v}}{\|\mathbf{x} \times \mathbf{v}\|}\,, \quad \mathbf{j} = \mathbf{k} \times \mathbf{i}\,, \quad (1)$$

where $\mathbf{v}$ is the velocity vector.

The next step done in [10] is to change the independent variable from τ to fictitious time σ, according to the Sundman transformation:

$$\frac{d\tau}{d\sigma} = \frac{r^2}{h}\,.$$

Finally, the variation of parameters technique is applied and the differential equations of the elements attached to the quantities z, u, ψ and $\boldsymbol{\varepsilon}$ are determined. In the next section we explain the meaning of the elements used in DROMO.

2.1 Integrals of the motion

The easier element to explain is q_3, which is equal to the inverse of the specific angular momentum:

$$q_3 = \frac{1}{\psi}\,. \quad (2)$$

[1] Actually Peláez chose **j** opposed to the direction of the angular momentum vector, and **i** along the radial direction (as above).

The other elements have not a such direct interpretation as q_3. Let $\mathbf{a}_p = (a_{px},\, a_{py},\, a_{pz})^{\mathrm{T}}$ be the perturbing acceleration vector (non-dimensional) expressed in the orbital frame $\mathcal{R}$, which was defined in (1). The angular velocity of $\mathcal{R}$ is given by (see [10]):

$$\boldsymbol{\omega} = \frac{r}{h} a_{pz}\, \mathbf{i} + \frac{h}{r^2}\, \mathbf{k}\,.$$

Let us consider now a reference frame $\mathcal{U}$ which is rotating with respect to $\mathcal{R}$ at the angular velocity:

$$\boldsymbol{\Omega}_{rel} = -\frac{h}{r^2}\, \mathbf{k}\,,$$

then, the angular velocity of $\mathcal{U}$ is:

$$\boldsymbol{\Omega} = \frac{r}{h} a_{pz}\, \mathbf{i}\,.$$

We see that when the component a_{pz} is equal to zero, $\mathcal{U}$ remains fixed with respect to inertial space. In general, there are ∞^1 reference frames $\mathcal{U}$ that rotate with the angular velocity $\boldsymbol{\Omega}$. They all have one axis coinciding with $\mathbf{k}$ and the other two axis rotated clockwise around $\mathbf{k}$ of $\sigma + C$ with respect to the directions of $\mathbf{i}$ and $\mathbf{j}$, where C is an arbitrary constant. For $C = 0$ the corresponding orbital frame $\mathcal{U}_0 = \langle \mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3 \rangle$ is defined by the rotation:

$$[\mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3] = [\mathbf{i},\, \mathbf{j},\, \mathbf{k}]\, Q_u\,, \quad (3)$$

where:

$$Q_u = \begin{pmatrix} \cos\sigma & \sin\sigma & 0 \\ -\sin\sigma & \cos\sigma & 0 \\ 0 & 0 & 1 \end{pmatrix}\,.$$

The reference frames $\mathcal{R}$ and $\mathcal{U}_0$ are shown in Fig. (1).

The elements attached to the inverse of the orbital radius z and the radial velocity u, named A and B in [10], are multiplied by ψ and the new elements:

$$q_1 = \psi A \qquad\qquad q_2 = \psi B$$

are introduced. The eccentricity vector, defined by (see for instance [2]):

$$\mathbf{e} = -\frac{\mathbf{x}}{r} - \frac{\mathbf{h} \times \mathbf{v}}{\mu}\,, \quad (4)$$

is projected on the reference frame $\mathcal{R}$ as follows:

$$\mathbf{e} = -\mathbf{i} - \frac{\mathbf{k}}{q_3} \times (u\, \mathbf{i} + s\, \mathbf{j})\,,$$

where s is the non-dimensional transverse velocity. By exploiting the relations:

$$s = q_3 + q_1 \cos\sigma + q_2 \sin\sigma$$
$$u = q_1 \sin\sigma - q_2 \cos\sigma \,,$$

and the identities $\mathbf{k} \times \mathbf{i} = \mathbf{j}$ and $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$, Eq. (4) yields:

$$\mathbf{e} = \left(\frac{q_1}{q_3}\cos\sigma + \frac{q_2}{q_3}\sin\sigma\right)\mathbf{i} - \left(\frac{q_1}{q_3}\sin\sigma - \frac{q_2}{q_3}\cos\sigma\right)\mathbf{j} \,.$$

Eq. (3) is used to project $\mathbf{e}$ on $\mathcal{U}_0$, it results:

$$\mathbf{e} = \frac{q_1}{q_3}\,\mathbf{u}_1 + \frac{q_2}{q_3}\,\mathbf{u}_2 \,, \tag{5}$$

and if the new variables:

$$\zeta_1 = \frac{q_1}{q_3} \,, \qquad \zeta_2 = \frac{q_2}{q_3}$$

are plugged into Eq. (5), we finally get:

$$\mathbf{e} = \zeta_1\,\mathbf{u}_1 + \zeta_2\,\mathbf{u}_2 \,.$$

The previous equation states that ζ_1 and ζ_2 are the projections of the eccentricity vector along the unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$:

$$\zeta_1 = e\cos\alpha \,, \qquad \zeta_2 = e\sin\alpha \,,$$

where α, shown in Fig. (1), is the angle between $\mathbf{u}_1$ and $\mathbf{e}$. At the initial time t_0, σ takes an arbitrary value σ_0. By setting:

$$\sigma_0 = \vartheta\,(t_0) \,, \tag{6}$$

it results $\alpha\,(t_0) = 0$, which means that $\mathbf{u}_1$ is parallel to the eccentricity vector at the initial time.

Now we deal with the elements attached to the unit quaternion $\boldsymbol{\varepsilon}$. It is shown in [10] that these elements are the components ε_{10}, ε_{20}, ε_{30} and ε_{40} of the unit quaternion related to the orbital frame $\mathcal{R}_0 = \langle \mathbf{i}_0,\, \mathbf{j}_0,\, \mathbf{k}_0 \rangle$ defined by the rotation:

$$[\mathbf{i}_0,\, \mathbf{j}_0,\, \mathbf{k}_0] = [\mathbf{i},\, \mathbf{j},\, \mathbf{k}]\, Q_0 \,,$$

where:

$$Q_0 = \begin{pmatrix} \cos\triangle\sigma & \sin\triangle\sigma & 0 \\ -\sin\triangle\sigma & \cos\triangle\sigma & 0 \\ 0 & 0 & 1 \end{pmatrix} \,,$$

with $\triangle\sigma = \sigma - \sigma_0$. We recognize that $\mathcal{R}_0$ belongs to the family of reference frames $\mathcal{U}$ introduced before, and, in particular, it corresponds to the choice $C = -\sigma_0$. As a consequence, this frame is invariant when the motion is unperturbed, and also when the disturbing acceleration is locked within the orbital plane. According to our assumption on σ_0 (see Eq. 6), the frame $\mathcal{R}_0$ is permanently rotated of $\vartheta\,(t_0)$ with respect to $\mathcal{U}_0$.

In general, for each choice of the value of C, it exists a SO(3) rotation with respect to an inertial reference frame based on the following Euler angles: the right ascension of the ascending node Ω, the inclination i and the angle $\overline{\omega} = \omega - \alpha - C$, where ω is the argument of periapsis. For example, $\overline{\omega} = \omega - \alpha + \sigma_0$ for $\mathcal{R}_0$, and $\overline{\omega} = \omega - \alpha$ for $\mathcal{U}_0$. The angles Ω, i and $\overline{\omega}$ are integrals of the motion, even when the acceleration vector is not zero and lies on the orbital plane.

Once the orbital reference frames $\mathcal{U}_0$ and $\mathcal{R}_0$ are defined, the elements used in DROMO:

$$\boldsymbol{\alpha} = \begin{pmatrix} q_1 & q_2 & q_3 & \varepsilon_{10} & \varepsilon_{20} & \varepsilon_{30} & \varepsilon_{40} \end{pmatrix}^{\mathrm{T}} \,, \tag{7}$$

are fully explained: q_3 is the inverse of the non-dimensional specific angular momentum (Eq. 2), q_1 and q_2, divided by q_3, give the projections of the eccentricity vector on $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively (Eq. 5), and ε_{10}, ε_{20}, ε_{30} and ε_{40} are the components of the unit quaternion which defines the orientation of $\mathcal{R}_0$ with respect to inertial space. The differential equations of these elements, derived in [10] through the variation of parameter technique, are:

$$\begin{aligned}
\frac{dq_1}{d\sigma} &= \frac{1}{q_3 s^2}\left[\sin\sigma\, a_{px} + \left(\frac{q_3}{s}+1\right)\cos\sigma\, a_{py}\right] \\
\frac{dq_2}{d\sigma} &= \frac{1}{q_3 s^2}\left[-\cos\sigma\, a_{px} + \left(\frac{q_3}{s}+1\right)\sin\sigma\, a_{py}\right] \\
\frac{dq_3}{d\sigma} &= -\frac{a_{py}}{s^3} \\
\frac{d\varepsilon_{10}}{d\sigma} &= \frac{a_{pz}}{2q_3 s^2}\left(\cos\triangle\sigma\,\varepsilon_{40} - \sin\triangle\sigma\,\varepsilon_{30}\right) \\
\frac{d\varepsilon_{20}}{d\sigma} &= \frac{a_{pz}}{2q_3 s^2}\left(\cos\triangle\sigma\,\varepsilon_{30} + \sin\triangle\sigma\,\varepsilon_{40}\right) \\
\frac{d\varepsilon_{30}}{d\sigma} &= -\frac{a_{pz}}{2q_3 s^2}\left(\cos\triangle\sigma\,\varepsilon_{20} - \sin\triangle\sigma\,\varepsilon_{10}\right) \\
\frac{d\varepsilon_{40}}{d\sigma} &= -\frac{a_{pz}}{2q_3 s^2}\left(\cos\triangle\sigma\,\varepsilon_{10} + \sin\triangle\sigma\,\varepsilon_{20}\right) \,,
\end{aligned} \tag{8}$$

where $\triangle\sigma = \sigma - \sigma_0$. The time equation takes the form:

$$\frac{d\tau}{d\sigma} = \frac{1}{q_3 s^2} \,. \tag{9}$$

2.2 DROMO accuracy

To test the efficiency of DROMO we select the example 2b at page 122 of the book by Stiefel and Scheifele [11]. The problem, which was also considered by Bond [3] and Peláez [10] for comparing different regularized methods, is to determine the position of a satellite after 288.12768941 mean solar days (msd) around the Earth under the perturbing action of the Earth oblateness and the Lunar gravitational attraction. The initial osculating orbit has an inclination of 30° with respect to the equatorial plane and an eccentricity of 0.95. The satellite is initially at the perigee at the distance $R_0 = 6800\,\mathrm{km}$. Details on the implementation of the two perturbations are available in the book [11].

We compare in terms of accuracy the following methods: the element formulation derived from the Kustaanheimo-Stiefel regularization scheme by Stiefel and Scheifele [11], the L-C extension developed by Fukushima [7], DROMO [10] and Cowell's method. Apart from the last, the other propagators adopt an independent variable which is not the physical time, but a fictitious time. In order to determine the value of the fictitious time corresponding to 288.12768941 msd, we created an iterative procedure based on the Newton-Raphson algorithm. At each iteration the differential equations are integrated by an explicit Runge-Kutta 5(4) pair of Dormand and Prince (DOPRI5) which implements local extrapolation. Once we know the final value of the independent variable, the comparison between the special perturbation methods is performed by means of an embedded Runge-Kutta numerical method 4(5) of fourth order with Cash-Karp parameters.

The components of the final position vector assumed as correct are:

$$\begin{pmatrix} \tilde{x}_f \\ \tilde{y}_f \\ \tilde{z}_f \end{pmatrix} = \begin{pmatrix} -24219.0501 \\ 227962.1064 \\ 129753.4424 \end{pmatrix} \text{ km} \,,$$

and were obtained by integrating DROMO, Sperling-Bürdet regularization [3] and Stiefel-Scheifele's variables with DOPRI5 at the maximum accuracy, and taking the common figures. Two errors are computed:

$$\mathrm{RSS} = \|\mathbf{r}_f - \widetilde{\mathbf{r}}_f\| \qquad \triangle R = |r_f - \widetilde{r}_f| \,,$$

where $\widetilde{\mathbf{r}}_f$ and $\mathbf{r}_f$ are the correct and the approximated final position vectors of magnitude $\widetilde{r}_f$ and r_f respectively.

Table 1 – Comparison of special perturbation methods for the problem of oblate Earth plus the Moon.

	Steps	RSS [km]	$\triangle R$ [km]
Sti-Sche	3100	0.0308	0.0208
L-C ext	3100	1.5534	1.5376
DROMO	3100	0.0973	0.0973
Cowell	10000	120.82	65.98

The results are shown in Table (1). DROMO is competitive with the other two regularized methods. Recent work is showing that the performance of DROMO can be improved. These promising results give evidence that the concept at the basis of DROMO is the right way for achieving highly accurate orbit propagation. Finally, note that Cowell produces very large errors, even if we are aware that the numeric integrator employed is not suited to this method.

3 Formation Flight

Let us consider n spacecraft in formation flight around a primary, and let $\mathcal{I} = \{O;\, \mathbf{i}_X,\, \mathbf{i}_Y,\, \mathbf{i}_Z\}$ be an inertial reference frame with the origin placed at the primary center of mass. At time t_0 the dynamic state of each spacecraft is given by the inertial position and velocity vectors $\mathbf{r}_0$ and $\mathbf{v}_0$. The problem of determining the time evolution of the formation is usually stated in these terms: *propagate the relative dynamic states of the spacecraft with respect to a reference orbit, and propagate this orbit*. The reference spacecraft (real or virtual) should be chosen in such a way that it represents the formation position in space, and the spacecraft relative positions can be easily understood. This formulation of the problem is suitable for planning station keeping and formation keeping control strategies.

We propose two approaches to determine the relative motion of the spacecraft with respect to a refer-

ence mass. In the first we apply DROMO special perturbation method [10] to each spacecraft, in the second only to the formation center of mass.

3.1 DROMO-FF

DROMO special perturbation method [10] is applied to each spacecraft in order to obtain a synchronous propagation of the relative dynamic states with respect to a selected reference mass. The new propagator created in this way is named DROMO-FF.

If Eqs. (8) were integrated for each mass in sequence, we would get, in general, the position and the velocity vectors of the spacecraft at different intermediate times from the initial to the final time of integration. In order to propagate the trajectories simultaneously, we have to assume a unique time imposed by a generic mass of the formation, for instance mass 1, so that:

$$\tau_1 = \tau_i \,, \qquad i = 2, \ldots, n \,.$$

As a consequence σ_i, with $i = 1, \ldots, n$, will take, in general, different values for the spacecraft at the same (non-dimensional) time τ_1, as can be inferred from the equation:

$$\frac{d\sigma_i}{d\tau_1} = q_{3,i}\, s_i^2 \,, \qquad i = 1, \ldots, n \,, \tag{10}$$

which is derived from Eq. (9). Therefore σ_1 is the independent variable and the other quantities σ_i, with $i = 2, \ldots, n$, are added to the set of the dependent variables. To summarize, we have one differential equation for time:

$$\frac{d\tau_1}{d\sigma_1} = \frac{1}{q_{3,1}\, s_1^2} \,, \tag{11}$$

the $n-1$ differential equations:

$$\frac{d\sigma_i}{d\sigma_1} = \frac{q_{3,i}\, s_i^2}{q_{3,1}\, s_1^2} \,, \qquad i = 2, \ldots, n \,, \tag{12}$$

where Eqs. (10) and (11) have been exploited, and the $7 \times n$ variational equations:

$$\frac{d\boldsymbol{\alpha}_i}{d\sigma_1} = \frac{q_{3,i}\, s_i^2}{q_{3,1}\, s_1^2}\, \frac{d\boldsymbol{\alpha}_i}{d\sigma_i} \,, \qquad i = 1, \ldots, n \,, \tag{13}$$

where $\boldsymbol{\alpha}$ is a column vector containing the elements (see Eq. 7), and the term $d\boldsymbol{\alpha}_i/d\sigma_i$ on the right hand represents the set of Eqs. (8) for the i-th mass. By integrating the $8 \times n$ Eqs. (11) - (13) we propagate the motion of the spacecraft at the same times. Finally, the relative dynamic states of the spacecraft with respect to the reference mass, for instance mass 1, are derived by differentiation:

$$\delta\mathbf{s}_i = \mathbf{s}_i - \mathbf{s}_1 \,, \qquad i = 2, \ldots, n \,.$$

3.2 DROMO-G

In this second model, named DROMO-G, the linearized relative motion with respect to the formation center of mass is propagated, together with its trajectory. The absolute motion of each spacecraft is the composition of the reference mass translational motion with the motion relative to this mass. Let us choose the formation center of mass (G) as the reference point of mass M, equal to the total mass of the formation. Then, the position and the velocity vectors of each spacecraft can be expressed by the sum:

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_G + \delta\mathbf{r} \\ \mathbf{v} &= \mathbf{v}_G + \delta\mathbf{v} \,, \end{aligned}$$

where $\delta\mathbf{r}$ and $\delta\mathbf{v}$ are the relative position and velocity vectors with respect to the center of mass. The time derivatives of $\mathbf{r}$ and $\mathbf{v}$ are given by:

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= \mathbf{v} \\ \frac{d\mathbf{v}}{dt} &= -\frac{\mu}{r^3}\mathbf{r} + \mathbf{g}_{\mathrm{p}}\left(\mathbf{r}\right) + \mathbf{a}_{\mathrm{p}}\left(\mathbf{r},\, \mathbf{v},\, t\right) \,, \end{aligned} \tag{14}$$

where the vector $\mathbf{g}_{\mathrm{p}}$ includes the perturbative gravitational effects under consideration, and $\mathbf{a}_{\mathrm{p}}$ is the resultant vector of the perturbative non-gravitational accelerations, which, in general, can be external or internal to the formation, so that $\mathbf{a}_{\mathrm{p}} = \mathbf{a}_{\mathrm{p}}^{\mathrm{ext}} + \mathbf{a}_{\mathrm{p}}^{\mathrm{int}}$.

The dynamic state of the center of mass varies according to the equations:

$$\begin{aligned} \frac{d\mathbf{r}_G}{dt} &= \mathbf{v}_G \\ \frac{d\mathbf{v}_G}{dt} &= \frac{1}{M}\sum_{i=1}^{n} m_i \left\{ -\frac{\mu}{r_i^3}\mathbf{r}_i + \mathbf{g}_{\mathrm{p},i}\left(\mathbf{r}_i\right) + \right. \\ &\left. + \mathbf{a}_{\mathrm{p},i}^{\mathrm{ext}}\left(\mathbf{r}_i,\, \mathbf{v}_i,\, t\right) \right\} \,. \end{aligned} \tag{15}$$

Let us expand the gravitational terms in Eq. (14) about the reference path with Taylor's theorem. Stopping at the first order, we generate the expansion:

$$-\frac{\mu}{r^3}\mathbf{r}+\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}\right)=-\frac{\mu}{r_G^3}\mathbf{r}_G+\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}_G\right)+ \\ +\left[\mathbf{G}\left(\mathbf{r}_G\right)+\mathbf{G}_{\mathrm{p}}\left(\mathbf{r}_G\right)\right]\boldsymbol{\delta}+\mathcal{O}\left(\boldsymbol{\delta}^2\right)\ , \quad (16)$$

where $\mathbf{G}\left(\mathbf{r}_G\right)$ and $\mathbf{G}_{\mathrm{p}}\left(\mathbf{r}_G\right)$ are the gravity gradient matrix and the gradient matrix of the perturbative gravitational accelerations calculated at the center of mass position vector. Eq. (16) is plugged into Eq. (15) and after applying the definition of $\mathbf{r}_G$ we have:

$$\frac{d\mathbf{v}_G}{dt}=-\frac{\mu}{r_G^3}\mathbf{r}_G+\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}_G\right)+ \\ +\frac{1}{M}\sum_{i=1}^{n}m_i\left[\mathcal{O}\left(\boldsymbol{\delta}_i^2\right)+\mathbf{a}_{\mathrm{p},i}^{\mathrm{ext}}\left(\mathbf{r}_i,\,\mathbf{v}_i,\,t\right)\right]\ . \quad (17)$$

Comparing Eq. (17) with Eq. (14), we see that they appear in the same form: the center of mass behaves like a particle of mass M moving around the primary under the influence of the perturbative acceleration vector:

$$\mathbf{a}_{G,\mathrm{p}}=\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}_G\right)+\frac{1}{M}\sum_{i=1}^{n}m_i\left[\mathcal{O}\left(\boldsymbol{\delta}_i^2\right)+\mathbf{a}_{\mathrm{p},i}^{\mathrm{ext}}\right]\ .$$

Thus, we decide to employ DROMO to propagate the motion of the center of mass.

Let us derive the equations that govern the evolution of the relative dynamics. By plugging Eq. (16) into Eq. (14) yields:

$$\frac{d\mathbf{v}}{dt}=-\frac{\mu}{r_G^3}\mathbf{r}_G+\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}_G\right)+\left[\mathbf{G}\left(\mathbf{r}_G\right)+ \\ +\mathbf{G}_{\mathrm{p}}\left(\mathbf{r}_G\right)\right]\boldsymbol{\delta}+\mathcal{O}\left(\boldsymbol{\delta}^2\right)+\mathbf{a}_{\mathrm{p}}\left(\mathbf{r},\,\mathbf{v},\,t\right)\ . \quad (18)$$

The difference between the right-hand sides of Eq. (18) and Eq. (17) gives the following expression for the relative acceleration vector of the generic spacecraft:

$$\frac{d\delta\mathbf{v}}{dt}=\mathbf{g}_{\mathrm{p}}\left(\mathbf{r}_G\right)+\left[\mathbf{G}\left(\mathbf{r}_G\right)+\mathbf{G}_{\mathrm{p}}\left(\mathbf{r}_G\right)\right]\boldsymbol{\delta}+ \\ +\mathcal{O}\left(\boldsymbol{\delta}^2\right)+\mathbf{a}_{\mathrm{p}}\left(\mathbf{r},\,\mathbf{v},\,t\right)-\mathbf{a}_{G,\mathrm{p}}\ . \quad (19)$$

Equations (19) and:

$$\frac{d\delta\mathbf{r}}{dt}=\delta\mathbf{v}\ , \quad (20)$$

describe the time evolution of the relative dynamic state of each spacecraft. Finally, the independent variable is switched from time to σ through Eq. (9).

4 Results

The performance of DROMO-FF in terms of accuracy and computing speed is compared to Cowell's method, which is commonly chosen for the orbit propagation of spacecraft. The simulations have been run: 1) in the same computer - Intel Core 2 Duo 2 GHz, 1 Gb RAM; 2) with the same compiler - Matlab 7.8.0.347; 3) with the same integrating algorithm - an embedded Runge-Kutta 4(5) of fourth order with Cash-Karp parameters.

4.1 Absolute Keplerian motion

The motion of one mass is characterized by 8 dependent variables in DROMO, one of these variables is physical time. When the relative motion of n masses is determined using DROMO-FF the number of dependent variables becomes $8\times n$, including: the physical time, $7\times n$ elements or integrals of the unperturbed two-body problem (seven for each point mass), and $n-1$ variables which were added in order to synchronize the propagation of the spacecraft trajectories.

In the case of pure Keplerian motion, the elements are by definition constants, while time and the variables σ_i, with $i=2,\,\ldots,\,n$, vary with the independent variable, which is denoted by σ_1, or more simply σ. While time is not involved in the computation of the dynamic state of the generic *i*-th mass, the quantity σ_i explicitly appears in the expression of the orbital radius and fixes its orientation on the orbital plane. Thus, the numerical errors introduced in σ_i, mainly due to truncation and round-off, also affect the position and velocity of the i-th mass.

In this section we analyse the influence of the numerical integration of the $n-1$ differential equations of the variables σ_i, with $i=2,\,\ldots,\,n$, in the accuracy and the computing speed with respect to the case in which DROMO is applied to each mass in sequence.

Table 2 – Orbital elements.

	Semi-major axis (km)	Eccentricity	Inclination (rad)	Initial true anomaly (rad)
Mass 1	70000	0.9	0	0
Mass 2	70000	0.9	10^{-3}	π

We consider two unperturbed ellipses around the Earth, followed by mass 1 and mass 2 respectively (Table 2). Mass 1 is chosen as the reference mass, and its position and velocity are propagated with no errors, because they depend on DROMO integrals and the independent variable σ_1. The dynamic state of mass 2 is determined by DROMO integrals and by σ_2, which is a dependent variable, affected by numerical errors. We compute the error in the final position vector of mass 2 through the root sum square, defined as follows:

$$\mathrm{RSS} = \sqrt{(x - x_{\mathrm{ref}})^2 + (y - y_{\mathrm{ref}})^2 + (z - z_{\mathrm{ref}})^2}\,, \tag{21}$$

where $(x,\, y,\, z)$ are the inertial coordinates of mass 2 as computed by our model, and $(x_{\mathrm{ref}},\, y_{\mathrm{ref}},\, z_{\mathrm{ref}})$ are the same components calculated analytically (or, which is the same, by employing DROMO for mass 2). The derivative of σ_2 with respect to σ_1 (Eq. 12, for $i = 2$) in terms of the classical orbital elements semi-major axis (a) and eccentricity (e) appears in the form:

$$\frac{d\sigma_2}{d\sigma_1} = \left[\frac{a_1\left(1-e_1^2\right)}{a_2\left(1-e_2^2\right)}\right]^{3/2}\left(\frac{1+e_2\cos\sigma_2}{1+e_1\cos\sigma_1}\right)^2\,. \tag{22}$$

Figure (2) shows the RSS up to 100 revolutions. We selected a very unfavorable example with two highly eccentric orbits phased by 180 degrees, so that picks are generated in the derivative of σ_2. These picks make the numerical integration less accurate and amplify the error when mass 2 is at perigee. The RSS increases as mass 2 accumulates revolutions around the Earth up to the maximum value of 88 meters reached in the last perigee passage. By decreasing the eccentricity down to 0.3, for example, and keeping invariate the semi-major axis for both orbits, the maximum RSS would be less than 3 meters.

Spacecraft flying in formation move along orbits of increasingly similar shape the closer they are, thus we expect that the error in the position made by DROMO-FF will be small.

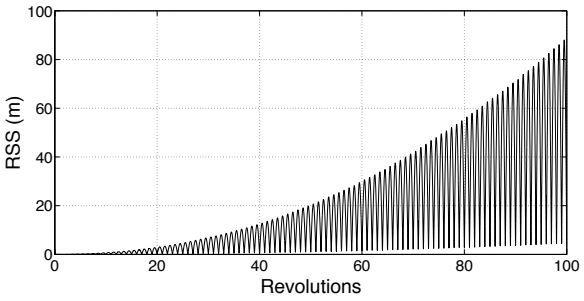


Figure 2 – Root sum square of the error in the position vector of mass 2, calculated by Eq. (21)**.**

Table 3 – Mean computing runtime and integration steps for 100 revolutions.

	Mean runtime (s)	Steps
DROMO-FF	132.43	135489
DROMO	81.02	111190

As regard the computation effort, we compare the integration steps and the computing time required to propagate the motion of masses 1 and 2 after 100 revolutions with DROMO-FF and with DROMO applied to each mass in sequence. We do not account for the computing time spent after the integration to calculate position and velocity starting from the dependent variables. Besides, to minimize the effect of uncontrolled factors we repeated the simulations 30 times and averaged the runtimes. Table (3) collects the results. DROMO-FF needs 63% more runtime due to the fact that the number of steps is 22% bigger and that each step is nearly one third of time slower. As we deal with less eccentric orbits this gap decreases until the simultaneous orbit propagation operated by DROMO-FF becomes faster than propagating mass 1 and mass 2 in sequence using DROMO.

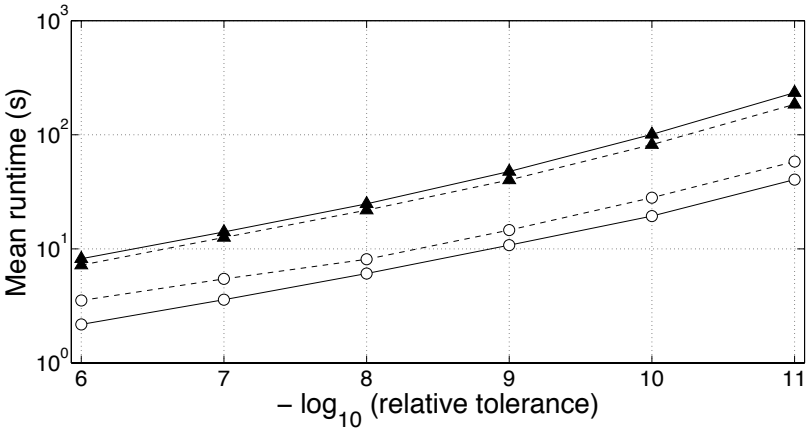


Figure 3 – The mean computing runtime of DROMO-FF (triangle marker) and Cowell's method (circle marker) for the two eccentricities 0.1 (markers connected by solid lines) and 0.7 (markers connected by dash lines) is plotted versus the relative tolerance of the RK 4(5) algorithm.

4.2 Relative Motion

We select the example reported in [13] to compare DROMO-FF with Cowell's method in terms of computing time and accuracy. Table (4) reports the classical orbital elements of the target along with the chaser position and velocity relative to the target at the initial time. The relative motion of the spacecraft under the influence of the perturbations due to the Earth oblateness and the Moon's third body gravitational attraction is propagated for a time interval of 50 periods of the target initial orbit. The constants implemented for the Earth, the Moon and the zonal harmonic J_2 are the same of the example 2b at page 122 of [11]. Figure (3) shows the mean computing runtimes of DROMO-FF and Cowell's method, each one being the average over 30 simulations, for different values of the relative tolerance adopted in the RK 4(5) algorithm. Figure (4) compares the accuracy in terms of the error:

$$E = \|\mathbf{r}_{\mathrm{T}} - \mathbf{r}_{\mathrm{T,\,ref}}\| + \|\mathbf{r}_{\mathrm{C}} - \mathbf{r}_{\mathrm{C,\,ref}}\| \ , \qquad (23)$$

where $\mathbf{r}_{\mathrm{T}}$ and $\mathbf{r}_{\mathrm{C}}$ are respectively the target and the chaser computed position vectors, while $\mathbf{r}_{\mathrm{T,\,ref}}$ and $\mathbf{r}_{\mathrm{C,\,ref}}$ are the corresponding "correct" position vectors. These were obtained by employing DROMO, Sperling-Bürdet regularization and Stiefel-Scheifele's method integrated by DOPRI5 with the maximum accuracy and keeping the common figures in the components.

Table 4 – Target and chaser initial conditions.

Target	
Eccentricity	0.1 and 0.7
Perigee height	500 km
Inclination	30 deg
Longitude ascending node	0 deg
Argument of perigee	0 deg
Initial true anomaly	45 deg
Chaser	
Position (Hill frame)	$[100,\ 10,\ 10]$ m
Velocity (Hill frame)	$[0.1,\ 0.1,\ 0.1]$ m/s

For the case of eccentricity equal to 0.1, the least accurate propagation by DROMO-FF produces an error of 0.37 m against the 605.38 m of the most accurate propagation by Cowell's method. Besides, the runtime is nearly 5 times smaller. The increase in the eccentricity up to 0.7 deteriorates the accuracy, expecially of DROMO-FF, which achieves a maximum error of 34.43 km. Nevertheless, even with the relative tolerance set equal to 10^{-7}, DROMO-FF is more accurate and faster, of respectively a factor 5.43 and 4.62, than Cowell's method integrated with the maximum accuracy.

In the previous section we dealt with the effect that the integration of the $n-1$ new variables σ_i, introduced to synchronize the orbit propagation of

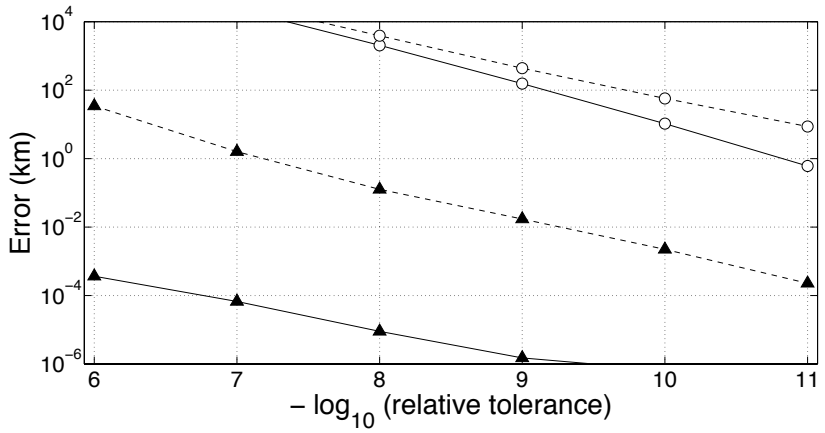


Figure 4 – The error, calculated by Eq. (23), of DROMO-FF (triangle marker) and Cowell's method (circle marker) for the two eccentricities 0.1 (markers connected by solid lines) and 0.7 (markers connected by dash lines) is plotted versus the relative tolerance of the RK 4(5) algorithm.

n masses, has on the computing time and accuracy of DROMO-FF. A different approach is to propagate each mass in sequence with DROMO. In the pure Keplerian motion this is equivalent in terms of accuracy to the analytical calculation, while as regard the computing time, it is wasted mainly in the numerical integration of the physical time, which is a dependent variable. Therefore, we are interested in checking also the performance of DROMO applied to the target and to the chaser in sequence. For the more eccentric case ($e = 0.7$) and the same values of the relative tolerance used to generate Figs. (3) and (4), the mean runtimes and the root sum squares of the errors in the final position vectors of each spacecraft are computed and then summed together. The orbit propagation of the target and the chaser in cascade improves the accuracy with a slight increase of the mean runtime: an error smaller then that of the most accurate simulation by DROMO-FF is made with a mean runtime more than 8 times shorter. However, by decreasing the relative tolerance the gap between the two appraoches reduces both in terms of accuracy and mean runtime.

4.3 Round-off error

There is a last issue we desire to address, the round-off error. DROMO-FF integrates the complete dynamics, whithout making any approximation. It can be less accurate for the same runtime than the integration of each mass in sequence by DROMO, but it presents the advantage of synchonizing the motion of the spacecraft so that we can appreciate the relative trajectories from the beginning to the end of simulations.

We have to pay attention on a possible source of error given by the differentiation operated in DROMO-FF to determine the relative dynamics from the absolute positions and velocities. If the spacecraft are so close that have α digits in common, and if the computer's word length is β digits, ($\beta \geq \alpha$), then the relative state vector obtained from the difference will have one or more components known to only $\beta - \alpha$ digits. The loose of accuracy just described is known as round-off error. The closer the spacecraft are the smaller is the number $\beta - \alpha$ of known digits and the bigger is the relative error on the relative dynamic state computation.

One way to overcome this error is to directly propagate the relative motion, as does DROMO-G, which was developed with this aim. Because linearization is performed, the method is particularly suited to close fomations, which, instead, are source of round-off error for DROMO-FF. Thus, we are interested in comparing the accuracy of DROMO-FF and DROMO-G when spacecraft are close one to the other. Let us select for the target at the initial time the same orbit reported in Table (4), with eccentricity equal to 0.7, apart from the true anomaly which is set equal to 0. For the chaser two displacements along the radial di-

Table 5 – Relative errors (%) of DROMO-FF and DROMO-G.

	10 cm	1 m
DROMO-FF	0.05	0.01
DROMO-G	2.01	2.17

rection of the target are considered: 0.1 m and 1 m. The motion is perturbed by J_2 and Moon and is propagated for 20 periods of the initial osculating orbit of the target. The accuracy on the final distance between the masses is calculated by imposing 40 steps per revolution in both the methods. The "correct" final distance is obtained with the same procedure of the previous examples and is respectively of 0.943 m and 9.435 m for the two cases. The relative errors are reported in Table (5). We see that DROMO-FF is very accurate and the relative error increases of 5 times after reducing of one order of magnitude the distance. Even if the distance between the spacecraft is so smaller than the orbital radius of the target, round-off does not affect the results because the main role in the accuracy is played by the global integration error.

5 Conclusions

We propose a method, named DROMO-FF, to simultaneously propagate the relative motion of n spacecraft in formation, based on the regularized formulation of the perturbed two-body problem developed by Peláez in 2006. DROMO-FF is compared with Cowell's method in terms of accuracy as well as computing time on the problem of propagating the motion of two spacecraft flying along perturbed orbits of small and high eccentricity. In the first case DROMO-FF reduces the error of at least a factor 1653 with moreover a saving in runtime for this case of a factor 5. In the latter DROMO-FF is 4 orders of magnitude more accurate and only from 2 to 3 times slower than Cowell's method. On the basis of these results we decided to employ DROMO-FF in our software simulator of spacecraft formation flying.

References

[1] Baù, G., Bombardelli, C., Peláez, J., Lorenzini, E., and Francesconi, A. A variational method for the propagation of spacecraft relative motion. In *20th AAS/AIAA Space Flight Mechanics Meeting*, volume 134 of *Advances in the Astronautical Sciences*, pages 2459–2474, San Diego, CA, February 14-17 2010. Paper AAS 10-266.

[2] Battin, R. H. *An Introduction to the Mathematics and Methods of Astrodynamics*. AIAA Education Series, AIAA, Reston, VA, revised edition, 1999.

[3] Bond, V. R., and Allman, M. C. *Modern Astrodynamics: Fundamentals and Perturbation Methods*. Princeton University Press, Princeton, NJ, 1996. P. 151.

[4] Broucke, R. A. Solution of the elliptic rendezvous problem with the time as independent variable. *Journal of Guidance, Control, and Dynamics*, 26(1):615–621, July-Aug. 2006.

[5] Clohessy, W. H., and Wiltshire, R. S. Terminal guidance system for satellite rendezvous. *Journal of the Aerospace Sciences*, 27:653–658, 674, Sept. 1960.

[6] Deprit, A., Elipe, A., and Ferrer, S. Laplace vs. stiefel. *Celestial Mechanics and Dynamical Astronomy*, 58:151–201, 1994.

[7] Fukushima, T. New two-body regularization. *The Astronomical Journal*, 133(1):1–10, January 2007.

[8] Hill, G. W. Researches in the lunar theory. *American Journal of Mathematics*, 1(1):5–26, 1878.

[9] Melton, R. G. Time-explicit representation of relative motion between elliptical orbits. *Journal of Guidance, Control, and Dynamics*, 23(4):604–611, 2000.

[10] Peláez, J., Hedo, J. M., and de Andréz, P. R. A special perturbation method in orbital dynamics. *Celestial Mechanics and Dynamical Astronomy*, 97(2):131–150, 2007.

[11] Stiefel, L., and Scheifele, G. *Linear and Regular Celestial Mechanics*. Springer-Verlag, New York Heidelberg Berlin, 1971.

[12] Tschauner, J., and Hempel, P. Rendezvous zu eine in elliptischer bahn umlaufenden ziel. *Astronautica Acta*, 11(2):104–109, 1965.

[13] Yamanaka, K., and Ankersen, F. New state transition matrix for relative motion on an arbitrary elliptical orbit. *Journal of Guidance, Control, and Dynamics*, 25(1):60–67, 2002.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A critical incident in the cockpit: Analysis of a critical incident interview using the Leximancer™ tool

Katie L. Plant and Neville A. Stanton
The Transportation Research Group, School of Civil Engineering and the Environment, University of Southampton, Highfield, Southampton, SO17 1BJ. UK

+44(0) 233 8059 3148 - k.plant@soton.ac.uk

Keywords: *Leximancer™, Critical Decision Method, Perceptual Cycle Model, Schema*

Abstract

It is argued that in order to establish a causal explanation of human error in the cockpit, it is necessary to understand why a pilot's actions and assessments made sense at the time. The Perceptual Cycle Model and the role of schemata are proposed as a way to explain decision making in the cockpit. In an attempt to elicit schematic processing, the Critical Decision Method was employed in a case study design. The interview was thematically analysed using a coding scheme based on the principles of the Perceptual Cycle Model. This paper presents the findings of importing the transcript into Leximancer™ to establish how Leximancer™ would categorise the data and whether it was comparable with the manual analysis. Results indicate that there is consistency between Leximancer™ and the manual analysis, although the order of importance of themes differs. Further data exploration is required to validate these findings.

1 Introduction

Human error is consistently implicated as a major contributor to accidents in safety critical systems [1; 2]. Whilst incidents usually arise from a combination of technical, systemic and human factors, what stands out is the human element, usually as there is a desire to ascribe blame [3]. The emphasis on the human element however is not without reason; human error is considered to be the principle threat to flight safety [4] and consistently attributed in seventy to eighty per cent of aviation accidents [1; 5]. Whilst it is acknowledged that human error in the aviation domain has received much attention in the literature [2; 6], traditional methods of research have tended to focus on *what* happened [7]. Dekker [7] argues that human error demands an explanation, rather than it being an explanation of failure.

Human error research should focus on understanding how people's actions and assessments made sense to them at the time i.e. *why* the error occurred [7]. A fully causal understanding of error will only occur if error is considered within the context of the system it occurred in [8]. In the cockpit, pilots are usually the final layer in the system and act to detect deficiencies in other layers. They do however,

add to these failures with their own mistakes and to answer the question of *why* an error occurred, an understanding of the interaction between the blunt and sharp ends of a system is essential [7].

The authors propose that the Perceptual Cycle Model, which places a heavy emphasis on Schemata, offers a theoretical framework for understanding how the sharp and blunt ends of a system interact and the reasons why operators perform erroneous actions [30].

1.1 The Perceptual Cycle Model

The Perceptual Cycle Model [11] presents the view that that human thought is closely coupled with a person's interaction in the world, both informing each other in a reciprocal, cyclical relationship. As illustrated in Fig. 1 world knowledge (Schemata) leads to the anticipation of certain types of information; this then directs behavior (action) to seek out certain types of information and provides a way of interpreting that information. The environmental experience (world) results in the modification and updating of cognitive Schemata and this in turn influences further interaction with the environment.

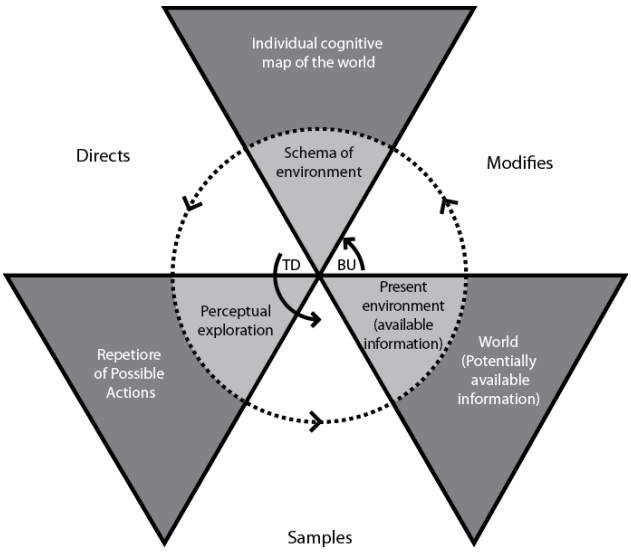


Fig. 1 The Perceptual Cycle Model (adapted from Neisser [11])

Schemata are a fundamental part of the model, these are defined as organized mental patterns of thoughts or behaviors to help organise world knowledge [11]. Walker et al. [9] define Schemata as mental templates which are neither new behaviour nor a repetition of old behaviour.

Dekker [7] argues that to understand error it is essential to start with an appreciation of the work situation (context) and to acknowledge that people change a situation by the actions they take, but that this evolving situation can also change behavior, this view of how we can understand and explain error is fundamentally the Perceptual Cycle Model. From this perspective errors can be understood by understanding the situation someone was working in, how this situation changed over time and how the evolving situation caused actions to be made that were relevant at the time.

A Strength of the Perceptual Cycle Model is that it provides a human-in-the-system approach of understanding error; individual schemata are considered in the context of the system in which they occurred. This middle stance of considering the interaction between person and environment is highlighted in the literature as the focus error research should take [7; 12]

1.2 Human error and Schema Theory

It is argued that deficiencies in mental representations need to be understood as they are potential causal factors in accidents [13]. Previous research has resulted in schema-based classifications of error [14] and the application of the principles in the Perceptual Cycle Model have been applied to accidents in a variety of safety critical domains, including rail [15] and road transport [2].

Schema theorists would argue that Schemata come into play when performing highly skilled behavior which occurs in the context of practiced routines based on stored actions and perceptual patterns, piloting an aircraft would fall into this category of behavior. Aviation accident research has consistently found that skill-based errors are implicated in a high proportion of accidents [16; 17]. The high representation of skill-based errors in accidents

suggests that applying a theory suited to skilled behavior is appropriate.

1.3 Schema Elicitation

From the literature presented above, it would appear that there is value in exploring the principles in the Perceptual Cycle Model in an attempt to establish causal explanations of error in the cockpit. There are however inherent issues with eliciting and understanding schemata, as Walker et al. [9:879] state, "*gaining insight into mental representations…is experimentally and conceptually challenging*". Schemata are metal constructs and therefore can not be directly measured but only inferred through the empirical consequences of them. It is argued however, that Schema are no more 'fuzzy' concepts than other mental inferences, such as attitudes and belief systems, and their existence is not disputed [18]. Dekker [19] argues that a lot of what is recorded in the error literature is the result of what *can* be observationally measured rather than what necessarily *should* be measured.

This paper briefly describes the results of a study in which the Critical Decision Method [CDM, 10] was employed in an attempt to elicit schemata. The focus of this paper however is on the use of the Leximancer™ software [23] and how this software aids the understanding of the representation of schemata. Additionally a comparison between a manual analysis and the Leximancer™ analysis will be made.

2 Method

2.1 Design

A single case study design was employed using the extended version of the CDM [20] in a retrospective semi-structured interview. Literature suggests it is acceptable to learn from samples of one, especially at the preliminary and exploratory stages of investigation [21; 22]. Additionally it is argued that critical incidents are a useful way of obtaining rich data and enhancing the understanding of safety as it is only a slight change in circumstance that results in an incident becoming an accident [21].

2.2 Materials

2.2.1 Critical Decision Method

The CDM [10] is a knowledge elicitation tool borne out of a need of a method that takes the knowledge of experts into account in naturalistic settings. This is achieved by the use of cognitive probes to understand decision making during non-routine incidents. The creators of the method [10] listed three criteria for the development of the method; *(1)* the need to address the competencies of highly skilled experts, *(2)* the applicability to field as opposed to laboratory conditions and *(3)* applied value for training or system design. Understanding the schematic processing of pilots also requires this information, i.e. research is focusing on highly skilled experts, in a naturalistic environment and there needs to be applied value, in this case in the form of an increased understanding of error.

The interview schedule consisted of a page of demographic questions and the cognitive probes provided by the CDM. These probes cover a range of decision making areas including situational awareness and expectancy.

2.2.2 Leximancer™

Leximancer™ is a software tool for performing rapid conceptual analysis of qualitative textual data in a six stage process. It can provide both automatic and customized analysis [23]. The approach has been employed in a variety of domains, for example to analyse factors associated with eating disorders [25], themes in the discourse of high achieving students [26] and to define themes of an annual education conference [27]. More specifically, the approach has been previously applied in the transport domain, to explore and understand mental representations held by car and motorcycle riders [9] and human error in the context of maritime operations [28].

2.3 Participant and Procedure

The participant was a thirty-nine year old male helicopter pilot recruited through an advert placed on the British Helicopter Association website. The participant was a relatively experienced pilot (three thousand flying hours) and recalled an incident that occurred when flying an AW139 during a Search and Rescue training exercise. The participant was asked to provide a high level over-view of a critical incident they had been involved with in which they were the primary decision maker. A critical incident was defined as being "a non-routine or un-expected event that was highly challenging and involved a high workload". Twenty-three CDM probes were asked, the participant was instructed to recall what happened rather than what he thought he should have done. The interview lasted approximately thirty five minutes.

2.4 Data Analysis

The interview was transcribed and the initial method of analysis, referred to as manual analysis, was to thematically analyse the data according to the categories in the Perceptual Cycle Model. The coding Scheme and examples are provided below:

- Schema: cognitive representation of present environment (current/existing knowledge), informed by experience and expectation. Example: "*…can pretty much guarantee one on most flights…the expectation is that there will be some sort of glitch"*
- Action: Perceptual exploration, locomotion and action. Example: "*...entered destination into the navigation system"*
- World: potential or available information, physical states and conditions. Example: "*...the power was on...*"

The data was modeled into the Perceptual Cycle Model following a structure of timeline of events during the incident. The phases of the incident were as follows; onset of problem (blank screens), immediate safety action, diagnosis 1 (circuit breakers), diagnosis 2 (emergency checklists), diagnosis 3 (electrical check) and decision (return to base).

Two colleagues who were not involved in the study coded segments of the interview transcript to asses for inter-rater reliability. The results of this coding were compared to that of the original analyst and a percentage of agreement between the raters was derived. The second phase of data analysis was to import the data into the Leximancer™ program. Five main stages are performed in order to transform transcripts into concept maps [9]:

1. Conversion of raw (text) data: definition of sentence and paragraph boundaries
2. Concept identifications: keywords are automatically extracted based on proximity, frequency and other grammatical parameters
3. Thesaurus learning: clusters are formed based on the extent to which collections of concepts appear together throughout the text
4. Concept location: blocks of text are tagged with associated concept names
5. Mapping: a visual representation of the semantic network is produced showing how concepts link to each other

The analyst has the option to define parameters of analysis such as how many sentences analyzed per block. As this paper is intended as an exploratory analysis a variety of editing options were tested. The outcome of this testing was to keep the analysis parameters as automatic (default options) as possible. This will result in a more objective analysis and will increase repeatability and speed of analysis which are suggested as reasons for using Leximancer™ in the first place [29]. The following changes were made to the Leximancer™ default options, in line with guidelines in the user manual for how to analyse transcript data [23]:

- Prose test threshold set to 0 (because the text is a colloquial transcript)
- Auto-paragraphing was switched off (because the transcript had no spacing for paragraph boundaries)
- Merge word variant was switched on (this will merge headwords such as look to include looking and looked)
- Merge similar words (these were manually defined as electrical/electronics, problem/problems and circuit/breakers)
- Name-like words were removed (for example Lee-on-Solent) as they were not relevant to the analysis

3 Results

3.1 Inter-rater reliability

Literature suggests that an acceptable level of agreement for inter-rater reliability can generally be defined as eighty per cent or above when reliability is calculated as the number of agreements divided by total number of segments [31]. Of the thirty two segments, agreement between the two coders and the criterion coder averaged eighty-six percent.

3.2 Thematic manual analysis

Below is a synopsis of the critical incident from the overview provided by the pilot:

"…Finished winching… then going to cliff winch… put in destination into the navigation system…typed in the three digit code… all the screens, well the 4 primary screens, went blank. Fortunately it was a clear day. So…put it into a safe configuration…went to autopilot… standard thing that we do. We initially started looking for circuit breakers... if any had popped. Couldn't see any, went to the flight cards…there wasn't one for that… not the sort of thing that is expected…agreed we would head back …looked at the electrical power and the generators to make sure the power was on …We do get lots of electrical problems and can pretty much guarantee one on most flights...the expectation is it's going to come up with some sort of glitch. Headed back discussing what electrics could have caused the problem…"

For each phase of the incident the data were modeled into the Perceptual Cycle Model, Fig. 2 provides an example of the manual data analysis for the summary of the whole incident.

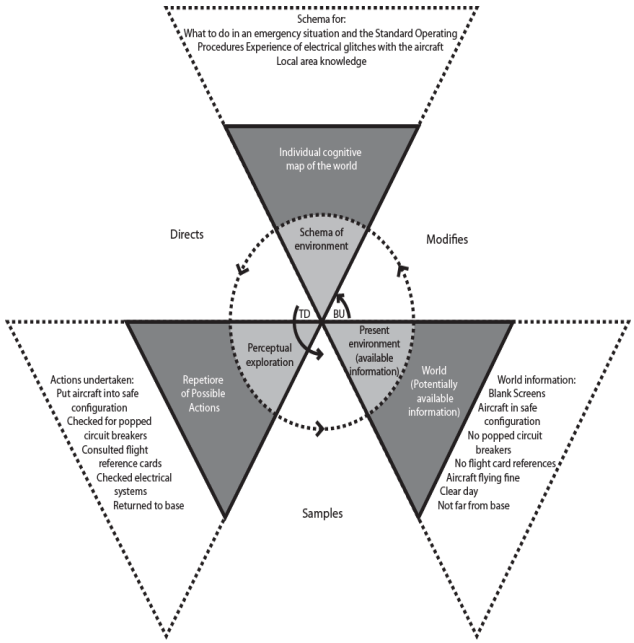


Fig. 2 Summary of the critical incident modeled into the Perceptual Cycle Model [11]

In the manual analysis 49% of the transcript was classified as schema, 29% as action and 22% as world. The fundamental principle behind Neisser's [11] original model is that information does not move in a simple linear flow but rather a cyclical interactive process. When dealing with this incident the pilot's perspective was compounded by a host of issues that are associated with the pilot's schematic representation about the aircraft. The pilot's diagnostic assumptions were based on his experience and expectations of the aircrafts' electrical system, which in turn led to his interaction in the world (fault diagnosis) being focused on the electrical system, rather than taking other factors into consideration. The situation in the world assisted his decision making (i.e. had it not been a clear day returning

to base may not have been an option) but the conditions in the world did not impact his individual cognitive representation of the situation (i.e. nothing changed his assumption that there was an electrical fault. It turns out that an electrical fault was not to blame, instead the loss of screens was due to software malfunction).

3.3 Thematic Leximancer™ analysis

The Leximancer™ output defines concepts as collections of words that 'travel together' throughout the text. These are weighted according to their frequency of occurrence in 'concept containing' sentences and non-concept sentences. The analysis generated twenty two concepts. Figure 3 displays the top ten concepts annotated with percentage relevance (100% being the most frequently occurring concept). Concepts are contextually clustered on the map such that if they frequently appear together or in similar situations in the text, they will be close together on the map. For the CDM data analyzed here, *problem* and *electronic* were the most common (frequently occurring) concepts (100% relevance each). Interestingly, *experience/knew*, *information* and *flying* all featured in the top ten concepts. These concepts are akin to the coding scheme; schema, world and action used in the manual analysis and will be discussed in more detail in the following section. Leximancer™ also displays concept connections with other concepts (not necessarily in the top ten). For example, the concept of *information* is highly linked to *screens* (86% level of co-occurrence) and *conditions* (67%).

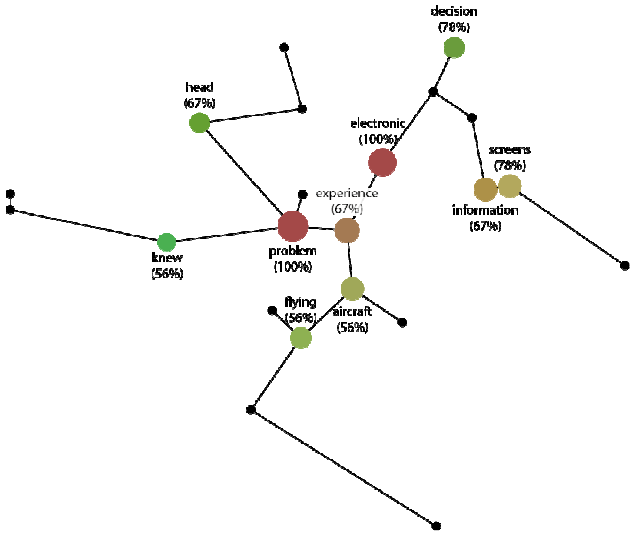


Fig. 3 Top ten concepts (with percentage relevance) from Leximancer™ output

Concept groupings are referred to as 'themes', these increase the analysis from individual items to broader, highly connected clusters. As with concepts, themes are assigned a relevance value (relative to the most connected theme which is always 100%). Figure 4 shows the top ten themes and percentage relevance. In this analysis, *problem* is the most relevant theme. As with the concepts, themes have also been identified that are akin to the coding scheme used in the manual analysis, including; *information* (38%), *flying* (30%) and *knew* (12%).

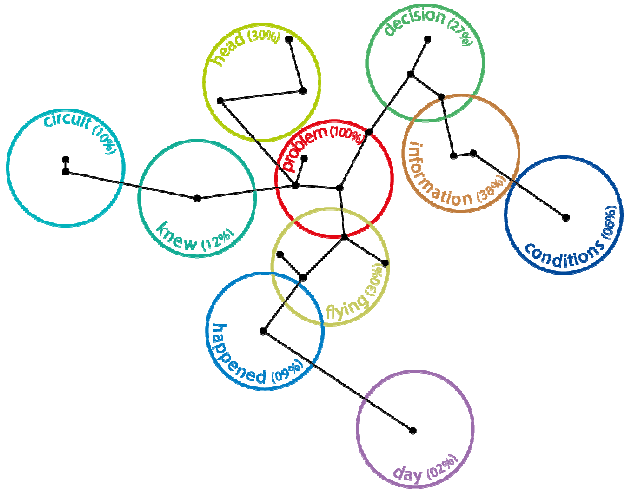


Fig. 4 Top ten themes (with percentage relevance) from Leximancer™ output

4 Discussion

Using the manual analysis approach, all data were able to be coded according to the principles of the Perceptual Cycle Model and at the highest level it would appear that the data fits with this model. It does appear that by using the CDM approach, information was elicited that can be understood in terms of Schematic principles. For example, during this incident the pilot's individual Schema of the situation was that there was an electrical fault with the aircraft. As the Schema literature suggest, this was heavily influenced by his experience of flying that aircraft as it often had electrical faults and therefore his expectations were set accordingly. It is clear to see how influential schemata are in decision making and subsequent actions during this critical incident in the cockpit. Whilst the pilot attempted a variety of actions (for example, checking circuit breakers and consulting flight reference cards) the actions always came back to looking for an electrical fault, as this is where his Schema was directing him, information in the world did not modify that Schema (i.e. nothing told him otherwise).

The purpose of this exploratory analysis was to compare the Leximancer™ output to that of the manual analysis which used a coding scheme based on the principles of the Perceptual Cycle Model, namely; schema, action and world. Leximancer™ identified twenty two individual concepts and ten key themes, this is far more than the three codes used in the manual analysis. This however, can be expected when using an automated computer program that has no pre-defined constraints on the number of concepts or themes it generates. As the purpose of this paper was to explore the Leximancer™ output in terms of its relevance to the Perceptual Cycle Model the main focus of the discussion will be on the concepts and themes relevant to elements of the Perceptual Cycle. There appear to be many similarities and over-laps between the manual and Leximancer™ analyses, suggesting that the elements of the Perceptual Cycle are a suitable way to model and understand processing in the cockpit during a critical incident.

Table 1 shows the top ten concepts along with their percentage of relevance.

Order of importance	Relevance (%)	Concept name
1	100	*Problem*
		Electronic
2	78	*Screens*
		Decision
		Experience
3	67	*Information*
		Head
		Aircraft
4	56	*Flying*
		Knew

Table 1. Top ten concepts with percentage relevance

It makes sense that *problem* and *electronic* were the two most relevant concepts as the critical incident transcript discussed the decision making process that occurred when trying to solve what the pilot thought was an electrical problem. The same applies for the relevance assigned to s*creens* and *decision*, as the context of the transcript data was decision making due to the loss of screens in the cockpit. It is interesting to note that in the top ten concepts *experience*, *information*, *flying* and *knew* feature as these are comparable with the manual analysis categories. *Experience* and *knew* are akin to the schema code used in the manual analysis, whilst *information* relates to world and *flying* to action. Leximancer™ allows the concept data to be explored in terms of which concepts are connected with other concepts (also visible in connection lines in Fig.3, i.e. the closer concepts are together, the more closely they are associated). When looking at the primary concept of *problem*, *experience* is the next most connected concept. This suggests that in the context of problem solving, schemata (prior experience of a situation) are the most prevalent sources of information to assist with the process. Whilst the concept data provides an overview of individual elements of the transcript, it is the themes (clusters of concepts) that represent a higher level of analysis of the overall situation and these will be discussed in further detail.

Table 2 shows the top ten themes along with their percentage of relevance.

Order of importance	Relevance (%)	Theme name
1	100	*Problem*
2	38	*Information*
3	30	*Flying*
		Head
4	27	*Decision*
5	12	*Knew*
6	10	*Circuit*
7	9	*Happened*
8	6	*Conditions*
9	2	*Day*

Table 2. Top ten themes with percentage relevance

It is to be expected that p*roblem* would be the highest ranking theme as the transcript was in the context of problem solving a critical incident and therefore this can be seen as the umbrella to subsequent themes. Similarly, *decision* features in the theme list, interview questions were constructed around the process of decision making and therefore the term decision features highly in the answers. As with the similarities in the concept list between the Leximancer™ output and manual coding, several themes were identified, including; *information*, *flying* and *knew* that are akin to the manual coding themes. There appears however to be a stark contrast between the Leximancer™ and manual analysis and the amount of relevance placed on each theme. In the manual analysis nearly half the data (49%) was classified as Schema. However the only theme relating to schema in the Leximancer™ output is *knew* (12% relevance to the over-arching theme of *problem*). It should be noted that the percentages represented in the manual and Leximancer™ outputs are not comparable insomuch that percentages relating to the manual output are an actual reflection of the total amount (i.e. the higher the amount the more it was represented), whereas the Leximancer™ percentages refer to connectivity to the primary theme. It can however be assumed that if the two data analysis method outputs were identical, then the themes given the most prominence in Leximancer™ would be the ones most represented in the manual analysis. This however was not the case; much more of the transcript in the manual analysis was identified as being relevant to schema. For example there are no stand alone themes relating to experience and expectation and recurring references were made to these in the transcript. *Experience* however was identified as a concept in Leximancer™ and at a concept level was the most connected concept to *problem* but not at a theme level.

For each theme, Leximancer™ provides a summary of the most relevant transcript extracts. For the theme of *information* the extracts relate to physical sates and pieces of information that were coded as world in the manual analysis. Two examples are presented below:

"the physical information available was blank screens, small artificial horizon and I could see out the window"

"…the four primary ones [screens] went blank…"

Information is the second most relevant theme in the Leximancer™ output, however only 22% of the transcript was classified as world in the manual analysis. Furthermore, *conditions* and *day* were identified in Leximancer™ as separate themes. Inspecting the transcript extracts associated within these themes reveals extracts that were coded as world in the manual analysis and also overlap between some of the extracts in the *information* theme which is why the themes are closely related on the concept map. Whilst this example highlights how three of the themes Leximancer™ generated (*information*, *conditions* and *day*) are akin to the manual coding scheme of world, it also shows a further difference between the Leximancer™ output and manual coding. In that Leximancer™ defines much narrower and specific themes than the manual coding.

The final theme in the manual analysis was action which was allocated to 29% of the data. As with the other manual themes, action was represented with the Leximancer™ output, although under a different name; *flying*. The *flying* theme received a more prominent representation than its equivalent (action) in the manual analysis as flying was the third most relevant theme identified by Leximancer™.

Furthermore, the Leximancer™ themes of *head* (the term head is misleading, here it refers to 'heading north' or 'heading back to base' and is therefore more of a navigation theme) and *happened* (in the context of 'thinking about what happened') included extracts that were coded as action in the manual analysis.

This exploratory analysis has compared the output of a manual thematic analysis with an automated analysis process. There are similarities between the two output methods insomuch that elements of the Perceptual Cycle which were used as the coding scheme for the manual analysis (schema, action and world) are represented in the automated analysis. The names may be different but the content (text extracts) that informs each theme are the same. This finding is encouraging when using the Perceptual Cycle Model as the foundation of critical incident and error research in the cockpit. The fact that the automated and more objective analysis method included the elements of the Perceptual Cycle in its classification of themes suggests that there is a reliance on Schema and the interaction of action and world information when dealing with critical incidents in the cockpit, at least in the case study presented here. The key difference between the manual and automated analyses was the level of importance each theme received. The Leximancer™ analysis appears to be driven by 'world' constraints whereas the manual analysis appears to be driven by 'cognitive' constraints. This distinction is highlighted by Vicente [32] in his explanation of ecological versus cognitive approaches to design. In the former, the mind needs to adapt to designs informed by world constraints (for example, laws of physics), whereas in the latter, existing mental models inform the design. This distinction however may be too absolute; in Schema Theory the world-mind boundaries are somewhat blurred and permeable [11].

This initial exploratory analysis has opened up interesting avenues to explore in future work Firstly, it would be interesting to manually define concepts within the three themes and assess the similarity to the specific concepts generated in Leximancer™. The three principles used in the manual coding are akin to Leximancer™ themes (i.e. a cluster of concepts). Each theme could be broken down into the concepts within them, for example, the manual coding theme of schema would include the concepts of experience, expectation and knowledge. Secondly, Leximancer™ has an 'edit concept' option in which concepts and themes can be manually defined and concept words assigned to each theme. This function was not employed here because the results of a purely objective analysis were sought. It would be interesting to assess the output if Leximancer™ was given increased guidance and whether there would be more overlap with the manual analysis. The results of this case study need to be replicated with other CDM interviews to validate the findings made here.

5 Conclusion

To conclude, it would appear that from a comparison of a manual thematic analysis and the output of the automated Leximancer™ analysis there are similarities in the themes generated. This provides support for using the Leximancer™ tool in future research which will result in a reliable, repeatable and rapid analysis process. Further exploratory work however is suggested to finalize the editable functions of Leximancer™ to ensure the most appropriate analysis is generated. Although not the primary focus of the study, the results here provide further support for using the Perceptual Cycle Model as a foundation to explain decision making and error in the cockpit as it would appear that the core elements of the Perceptual Cycle were represented in a more objective method of data analysis.

References

1. Amalberti, R., "The paradoxes of almost totally safe transportation systems". *Safety Science*, Vol. 37, 2001, pp. 109-126.
2. Stanton, N. A., & Salmon, P. M. "Human Error Taxonomies applied to driving: generic driver error taxonomy and its implications for intelligent transport systems". *Safety Science*, Vol. 47, 2009, pp. 227-237.
3. Woods, D.D., Dekker, S., Cook, R., Johannesen, L., & Sarter N., *Behind Human Error.* Aldershot: Ashgate Publishing. 2010.

4. Harris, D. & Li, W-C., "An Extension of the Human Factors Analysis and Classification System for use in Open Systems". *Theoretical Issues in Ergonomics Science*, Vol. 12, No. 2, 2011, pp. 108-128
5. Baksteen, B., "Flying is not safe". *Safety Science*, Vol. 19, 1995, pp. 287-294
6. Salmon, P., Regan, M. & Johnston, I., "Human error and road transport: Phase One – Literature Review". *Monash University Accident Research Centre Report No. 256,* 2005.
7. Dekker, S. *The Field Guide to Understanding Human Error*. Ashgate: Alddershot, 2006.
8. Reason, J., *Human Error*. Cambridge: Cambridge University Press, 1990.
9. Walker, G. H., Stanton, N.A. & Salmon, P. M., "Cognitive compatibility of motorcyclists and car drivers", *Accident Analysis and Prevention*, Vol. 43, No. 3, 2011, pp. 878-888
10. Klein, G.A., Calderwood, R., & Macgregor, D. "Critical Decision Method for Eliciting Knowledge". *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 19, No. 3, 1989, pp. 462-472.
11. Neisser, U., *Cognition and Reality*. San Francisco: W.H.Freeman and Company, 1976
12. Fedota, J. & Parasuraman, R., "Neuroergonomics and human error". *Theoretical Issues in Ergonomics Science* (iFirst), 2009, pp. 1-20.
13. Richardson, M. & Ball, L. J., "Internal representations, external representations and ergonomics: Towards a theoretical integration". *Theoretical Issues in Ergonomics Science*, Vol. 10, No. 4, 2009, pp. 335-376
14. Norman, D. A.., "Categorization of action slips", *Psychological Review*, Vol. 88, 1981, pp. 1-15
15. Stanton, N. A., & Walker, G. H., "Exploring the psychological factors involved in the Ladbroke Grove rail accident". *Accident Analysis and Prevention* (in press)
16. Wiegmann, D., & Boquet, A.., "Human Error and general aviation accidents: A comprehensive, fine-grained analysis using HFACS". (*Report number DOT/FAA/AM-05/24*). Washington DC: Federal Aviation Administration, 2005.
17. Lenne, M. G., Ashby, K., & Fitzharris, M., "Analysis of general aviation crashed in Australia using the Human Factors Analysis and Classification System", The *International journal of Aviation Psychology,* Vol. 18, No. 4, 2008, pp. 340-352.
18. Lodge, M., McGraw, K. M., Johnston Conover, P., Feldman, S., & Miller, A. H., "Where is the schema? Critiques", *The American Political Science Review*, Vol. 85, No. 4, 1991, pp. 1357-1380
19. Dekker, S. A., "Illusions of Explanation: A critical essay on error classification", *The International Journal of Aviation Psychology*, Vol. 13, No. 2, 2003, pp. 95-106.
20. Ericsson, K. A., & Smith, J., *Toward a general theory of expertise*. Cambridge: Cambridge University Press, 1992.
21. March, J.G., Sproull, L.S., & Tamuz, M., "Learning from samples of one or fewer", *Quality and Safety in Health Care*, Vol. 12, 2010, pp. 465-471.
22. Hancock, P.A., Hancock, G.M., & Warm, J.S., "Individuation: The N = 1 revolution", *Theoretical Issues in Ergonomics Science*, Vol. 10, No. 5, 2009, pp. 481-488.
23. Leximancer. "*Leximancer user Manual: From words to meaning to insight, v3.5*". Retrieved September 5, 2011, from www.leximancer.com
24. Marques, J.F., & McCall, C., "The application of inter-rater reliability as a solidification instrument in a phenomenological study", *The qualitative report*, Vol. 10, No. 3, 2005, pp. 439-462.
25. Hepworth, N., & Paxton, S.J., "Pathways to help-seeking in Bulimia nervosa and binge eating problems: A concept mapping approach", *International Journal of Eating Disorders*, Vol. 40, No. 6, 2007, pp. 493-504.
26. Grimbeek, P., Bartlett, B. & Loke, K-K., "Using Leximancer to Identify themes and patterns in the talk of three high-distinction students", *Second Annual Conference on Cognition, Language, and Special Educational Research*. Queensland, Australia, 2004, pp. 122-128.
27. Grimbeek, P., Bryer, F., Davies, M. & Bartlett, B., "Themes and patterns in three years of abstracts from the "International Conference on Cognition, Language, and Special Educational Research" identified by Leximancer analysis", In: Bartlett B., Bryer F. and Roebuck D, editors. Stimulating the ''action'' as participants in participatory research Brisbane. Australia: Griffith University, School of Cognition, Language, and Special Education, 2005. pp. 101–113.
28. Grech, M., Horberry, T., & Smith, A., "Human error in maritime operations: Analyses of accident reports using the Leximancer tool". *Proceedings of the Human Factors and Ergonomics Society Annual Meetings*, USA, 2002, pp. 1718-1722.
29. Smith, A.E., "Automatic Extraction of Semantic Networks from Text using Leximancer", *Proceedings of HLT-NAACL 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: Companion Volume*, 2003, Edmonton: ACL, pp. Demo23-Deom24
30. Plant, K.L & Stanton. N.A., "Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model", *Safety Science* (in press)
31. Marques, J. and McCall, C., "The Application of Inter-rater Reliability as a Solidification Instrument in a Phenomenological Study ", *The Qualitative Report*, Vol. 10, No. 3, 2005, pp. 439-462.
32. Vicente, K.J. *Cognitive Work Analysis: Toward Safe, Productive and Healthy computer-based work*, Lawrence Erlbaum Associates: Mahwah, NJ. 1999.

TUESDAY OCTOBER 25th



CEAS 2011

Propulsion

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Integration of CFD Tools in Aerodynamic Design of Contra-Rotating Propeller Blades

S. Béchet and C. A. Negulescu
Airbus Operations S.A.S. - Aerodynamics Department, Toulouse, France

V. Chapin
ISAE DAEP, Toulouse, France

F. Simon
ONERA DMAE, Toulouse, France

Keywords: *Contra-Rotating Propellers, Open Rotors, Blades, Aerodynamic Design, CFD*

Abstract

A strategy for Contra-Rotating Open Rotors blades design is presented. It is based on several analytical or Computational Fluid Dynamics (CFD) tools with increasing levels of accuracy. A preliminary design is made with a lifting-line code. These first geometries are then validated and improved with more advanced CFD methods. Steady CFD computations on a blade passage using the mixing-plane technique permit to perform an accurate sensitivity analysis of the main blade geometrical parameters. The key parameters driving the propulsive efficiency are thus identified.

Modeling rotor-rotor interactions require more costly unsteady simulations such as chorochronic computations on a blade passage, or chimera/sliding-mesh techniques on the full model to assess the installation effects. At the end of the design process, these unsteady methods are used to validate the robustness of the advanced designs to those interactions.

The relevance of this strategy is discussed through its results and limitations.

Nomenclature

ρ	air density [kg/m^3]
γ	air adiabatic coefficient
a	celerity of sound [m/s]
N	front propeller rotational speed [rev/s]
ω	angular speed [rad/s]
D	front propeller diameter [m]
r	radial location [m]
R	propeller tip radius [m]
J	front propeller advance ratio
V	infinite inflow velocity [m/s]
p	static pressure [Pa]
M_∞	infinite Mach number
Cth	thrust coefficient
Cp	power coefficient
η	propulsive efficiency
$._F$	variable related to front rotor
$._R$	variable related to aft/rear rotor
$._{SEC}$	variable related to blade section

1 Introduction

Contra-Rotating Open Rotors (CROR) are highly efficient aircraft propulsion systems [1]. This concept promises significant reductions of fuel consumption and pollutant emissions. Its interest is highlighted by the current context of economic and environmental concerns. After a

period of fruitful developments and tests performed in the US during the '80s in response to the oil crisis of that period, interest in CROR decreased and they finally never reached the commercial aviation.

Current economic and environmental objectives have led aeronautical industry to put new research efforts into that concept which is a candidate for next generation short-range aircraft propulsion systems. Emphasis is put on the development of experimental and numerical methods dedicated to better understand and predict CROR aero-acoustic behavior.

The CROR propeller design is a complex task as it involves multidisciplinary aspects like aerodynamic, acoustic and structural concerns. Already for an isolated engine the strong front/rear rotors aerodynamic wake and potential field interactions cause unsteady flow on the blades and interaction noise. These interactions are even more significant in installed conditions on the aircraft. Indeed, depending on the chosen configuration and because of the presence of the pylon, wing, and due to aircraft incidence, the propeller inflow is not uniform. Therefore the flow on the blades presents even more unsteadiness and will impact not only the aerodynamic performances, but also the acoustic emissions and the structural integrity of the propeller.

This paper presents the development of a CFD-based strategy for aerodynamic design of CROR blades. Nowadays several types of numerical methods and levels of modeling are available. Fast techniques such as lifting line methods have been widely used up to now [2]. However these methods are not sufficient any more, even though they remain valid for preliminary design. A higher level of accuracy is sought and can be found in more advanced CFD methods able to simulate the complete three-dimensional viscous flow and even its unsteadiness.

Two kinds of CFD methods need to be considered. Firstly, because the design process requires many iterations there is a strong need for a flexible mesh generation and short-time response CFD. Secondly, a deeper analysis is performed with costly unsteady computations in order to validate a specific configuration.

The following paragraphs present these different methods and show how they are used to conduct a complete design process. Besides, results are presented that can be expected at each step.

2 Preliminary Design

The first step of CROR blades design aims to define a preliminary set of reference geometries respecting global aerodynamic performance, installation, structural and acoustic constraints.

2.1 Operating conditions

Two conditions are considered in the design process:

- Cruise: Mach 0.75 / RPM 795 / Alt 35kft
- Take-off: Mach 0.20 / RPM 1032 / Alt 0ft

2.2 Performances

The different designs have to fulfil the cruise and take-off propeller thrust requirements. That sets the global thrust coefficient:

$$Cth = \frac{Thrust_{GLOBAL}}{\rho.N^2.D^4} \qquad (1)$$

Engine considerations also fix the power ratio between the front (F) and rear (R) rotors:

$$\frac{Cp_F}{Cp_R} = \frac{Power_F}{Power_R} = \frac{Torque_F.\omega_F}{Torque_R.\omega_R} \qquad (2)$$

These requirements will be satisfied during the iterative design process by adapting the front and rear propeller pitch angles β_F and β_R.

The aerodynamic performance of each design is given by the propeller propulsive efficiency:

$$\eta = J.\frac{Cth}{Cp} = \frac{V}{N.D}\frac{Cth}{Cp} \qquad (3)$$

The objective of the design process is to optimize this efficiency.

2.3 Reference propeller dimensions

For a 150-seat single aisle short range aircraft with a rear-mounted pusher engine configuration, installation concerns limit the diameter of the propellers to 14 feet. A high value is preferable both for performances and acoustics. For acoustic reasons, the rear rotor diameter is lowered by 10% in order to reduce its interaction with the front rotor tip vortex. To maintain the aft propeller blade surface its blade chord is increased.

Front and rear rotor blade numbers, which are the results of a compromise between aerodynamics and acoustics, have been chosen as 11 and 9 respectively. For the same reasons, the distance between front and rear rotors planes is set to 950 mm.

2.4 Definition of the blade shape

The shape of the blade is defined by the following radial blade laws distributions:

- Airfoil type
- Chord (m)
- Relative camber (% of chord)
- Relative thickness (% of chord)
- Twist angle (°)
- Sweep angle (°)
- Dihedral (m)

Structural considerations associated to reference geometries found in literature permit to define a relative thickness (decreasing from a minimum foot value to a minimum tip value) and chord distribution. At the same time, an airfoil type is selected as a compromise between the different operating conditions of the propeller (e.g supercritical airfoil).

The first lifting line computations with the code LPC2 from ONERA [3] are now performed at high and low speed conditions with imposed target thrust and power ratio with a convergence on pitch angle settings. Some iterations are necessary to determine camber and twist distributions that give at high speed the target load distribution (e.g. Goldstein load distribution [4]). Fig. 1 shows the limitations of the modeling of the load distribution (especially a root and tip) using this lifting line code.

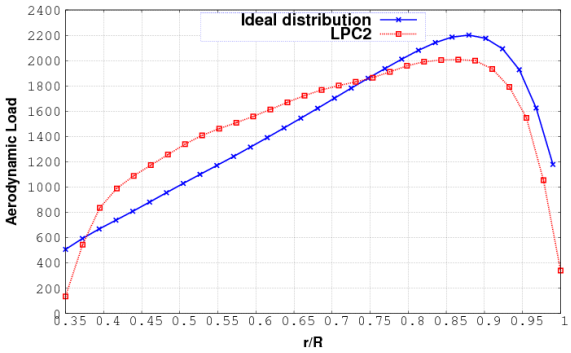


Fig. 1 Aerodynamic load distribution computed with lifting line method compared to ideal distribution

The final step is made through a new series of lifting line computations to find a more adequate sweep distribution in order to reduce the compressibility losses and increase the efficiency. For structural considerations, one has to be careful that the blades remain balanced around their pitch change axis.

Even though the quality of the lifting line solution is questionable, this first step is important because the design choices made at the beginning may condition the further results obtained with more accurate methods.

2.5 Reference configuration

This approach has led to the AI-PX7 11x9 CROR blades design. The blades are mounted on a representative nacelle dummy (Fig. 2). This configuration will be used as reference for further steps of the design. This is also a current reference configuration for the European SFWA/CleanSky project[1].

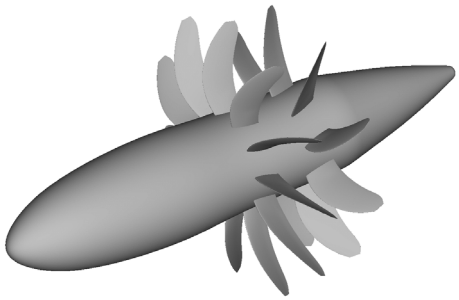


Fig. 2 AI-PX7 CROR Configuration

[1] http://www.cleansky.eu/

The upstream infinite Mach number was adjusted in order to have M=0.75 in the propeller plane (i.e. aircraft installed conditions).

3 Detailed design

In a second phase the goal is to enhance the aerodynamic performances of the proposed CROR. This is done through an analysis based on more accurate CFD tools which are able to simulate the 3D steady flow on the blades.

3.1 Mixing plane technique

The first approach consists in performing steady RANS computations on a limited-size grid using the mixing plane technique in the ONERA code *elsA* [5].

Each rotor is restricted to one inter-blade passage. The flow is solved in the rotating frame. At the interface between front and aft rotor grids, data is azimuthally averaged along a mixing plane. On the lateral boundaries periodic boundary conditions are set.

In contrast to a full unsteady modeling the steady-state mixing plane approach does not permit to simulate the interactions between the propellers, as well as any incidence or installation effects. Only steady load is known and this restricts any future acoustic evaluation to steady load and shock noise assessment.

However, such a method is well adapted for an aerodynamic design purpose, as it gives access to the 3D blade flow within a short computation time.

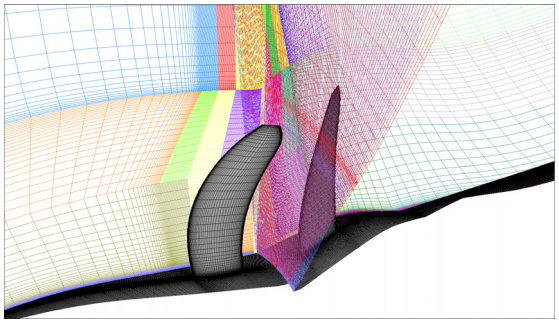


Fig. 3 Structured mesh on AI-PX7 configuration (mixing plane technique)

3.2 Automatic process

The interest of this kind of intermediate method is its ability to compute a high number of configurations. This can be even more efficient if the CFD solver is embedded inside an automatic process also including CAD and mesh generation tools. This design strategy is adopted here.

This automatic process is very important, as the extension of this work will be the integration of the parametrical design process into an aerodynamic and finally an aero-acoustic optimization process.

3.2.1 Blade parameterization

The first step of such an automatic process is to generate the blade shape. All the parameters previously defined in Section 2.4, as well as the pitch angles of the front and aft blades, are controlled. The goal is to respect with the highest possible accuracy the target radial blade laws distributions by an adequate set of control sections (Fig. 4).

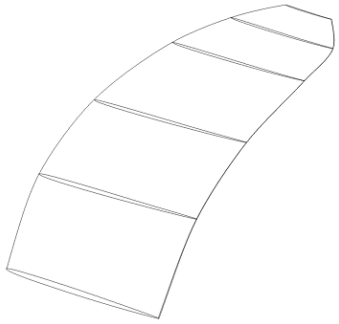


Fig. 4 AI-PX7 Front Blade Parameterization

The in-house tool *PADGE* (Parametrical And Differentiated Geometrical Engine) [6] uses these control sections to reconstruct the complete blade surface. The generation of the shape is ensured by the use of *NURBS* (Non-Uniform Rational Basis Splines) calculated between sections with tangency constraints at the sections location. One strong interest of this method is the capacity to control and modify the shape with the relevant parameters defining the blade shape.

3.2.2 Initial mesh

The geometry obtained through this parameterization tool is mounted on the nacelle and the structured multi-block mesh is generated, with a boundary layer mesh both on blades and nacelle. It has been shown that a 4 million nodes grid for the front and rear inter-blade passage is sufficient to perform an aerodynamic study of the CROR.

3.2.3 Mesh deformation

This reference mesh is an input for the initialization of the blade shape parameterization tool. From this point on, any modification of a parameter (shape or pitch angle) is treated through the in-house tool *VOLDEF* (VOLume mesh DEFormation) [7] as a displacement of the surface nodes and propagated in the volume mesh until the boundary of the domain. The displacement of the boundaries can be limited or even forbidden. This is particularly important for periodic sub-faces and the mixing plane.

3.2.4 RANS computations

Computations are run on the *elsA* CFD solver from ONERA. The study is performed with the Sparlart-Allmaras one-equation turbulence model.

3.2.5 Post-processed data

Propeller performances analysis requires essentially post-processing of pressure and friction data on the surface of the blades. Integration gives access to the aerodynamic forces and moments and so to the front, aft and global performances coefficients *Cth*, *Cp* and η.

For local analysis of the flow on the blades, the pressure coefficient taking into account the relative velocity due to the propeller rotation is defined as:

$$Kp_{ROT} = 2\frac{(p_\infty - p)}{\gamma . p_\infty . M_{rel}^2} \qquad (4)$$

with :

$$M_{rel}^2 = M_\infty^2 + M_{rot}^2 = M_\infty^2 + \frac{r^2 . \omega^2}{a_\infty^2} \qquad (5)$$

The radial distributions of lift and drag coefficients are defined as:

$$Cl(r) = 2\frac{Fl_{SEC}}{\gamma . p_\infty . M_{rel}^2(r) . c(r)} \qquad (6)$$

$$Cd(r) = 2\frac{Fd_{SEC}}{\gamma . p_\infty . M_{rel}^2(r) . c(r)} \qquad (7)$$

with Fl_{SEC} and Fd_{SEC} the radial section forces (taking into account the pressure and friction contributions) respectively normal and tangent to the profile chord whose length is *c*.

3.3 Pitch angles adaptation

As previously stated, the blades have to provide the required thrust and power ratio. As the RPM is fixed, the only way to respect those specifications is to find the adapted pitch angle settings on both rotors. These settings cannot be known directly. For each rotor a minimal and a maximal pitch angles must be defined such as the corresponding thrust and power ratio ranges will always include the target values for different blade shapes. As explained in Fig. 5, four computations are needed and adapted pitch angles are found by a linear interpolation.

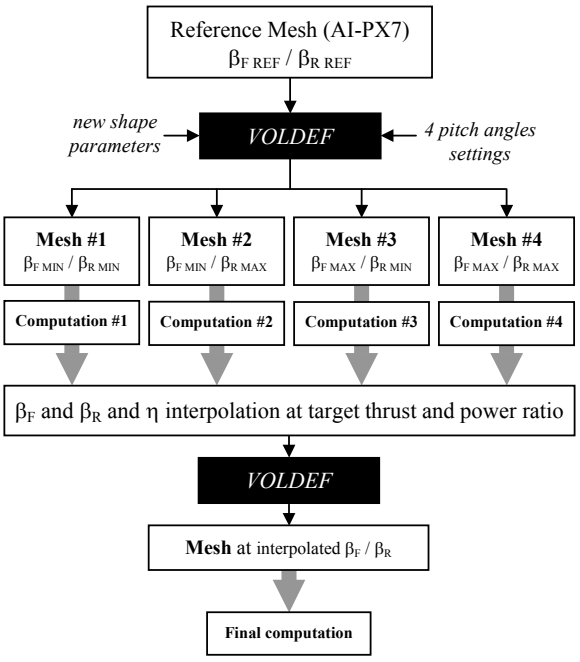


Fig. 5 Aerodynamic Computations Process

For simplicity reasons, the same extreme pitch angle values are taken for front and aft rotors. These have been estimated by the lifting line method with ±5% variations on the target thrust and power ratio.

A final computation is then done with the interpolated pitch angles giving thus access to the blades aerodynamic at the required operating conditions.

3.4 Sensitivity analysis

In the following paragraphs some results of a sensitivity analysis are presented. Shape modifications are performed on the AI-PX7 reference configuration. The objective is to identify some geometrical modifications that influence significantly the aerodynamic propeller performances (efficiency) at high speed.

The choice of each modification is either a consequence of the analysis of the reference configuration, for which weaknesses and potential improvements have been identified, or it is due to an external constraint such as structural or acoustic.

It has to be mentioned here that the level of impact of some of the parameters depends strongly on the reference design layout.

3.4.1 Effect of chord modification

A first interesting test is to scale the front and aft chord distributions by a global factor. Table 1 highlights the potential efficiency increase that can be obtained with a smaller surface, confirming the fact that the blades are not working at the optimal L/D ratio.

Table 1 Efficiency evolution for chord modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
+10% on global chord	-0.40%	-0.19°	-0.11°
+5% on global chord	-0.19%	-0.10°	-0.05°
-5% on global chord	+0.18%	+0.11°	+0.06°
-10% on global chord	+0.34%	+0.23°	+0.12°

3.4.2 Effect of sweep modification

Sweep angle is now increased at the tip of front and aft blades to reduce the local Mach number and thus the shocks intensity. As it appears on Table 2, no significant effect is noticed on the propeller efficiency for an 8° or a 14° tip sweep increase.

Table 2 Efficiency evolution for tip sweep modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
+8° on tip sweep	-0.04%	-0.03°	+0.03°
+14° on tip sweep	-0.04%	-0.01°	+0.08°

In a second step, the sweep distribution is increased all over the blade except at the tip. This is done keeping 0° sweep at r/R=55% and increasing the negative blade root sweep by -4°. Table 3 reports an insignificant effect. Apparently, increasing the sweep (locally or globally) has no major effect on the efficiency. The reference blade has obviously already a sufficient sweep level and therefore no strong shocks (Fig. 6).

Table 3 Efficiency evolution for sweep modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
Scaled sweep	+0.02%	+0.05°	+0.03°

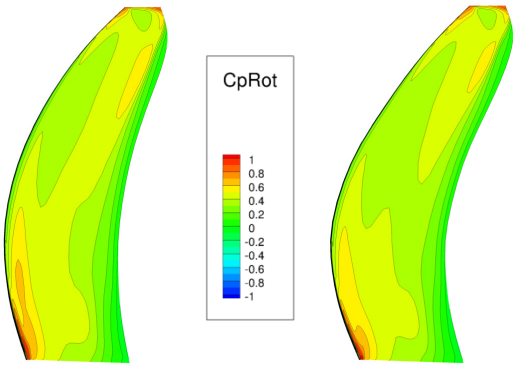


Fig. 6 Front blade suction side pressure coefficient distribution (reference vs. scaled sweep)

3.4.3 Effect of twist modification

A direct way to modify the load distribution is to change the root and tip blade twist. Two

tests are performed here: first the tip twist is decreased by 3.2°, secondly the root twist is diminished by 2.3°. Both cases correspond to a local load reduction. A third case combines these two modifications.

Effects on efficiency are presented in Table 4. It appears that a lower twist angle at blade tip increases the efficiency. On the other side the blade root twist modification decreases the efficiency. This can be explained by a better matching or respectively by a deviation from the optimal blade loading distribution. The combination of these two modifications has consequently no effect on the efficiency.

Table 4 Efficiency evolution for twist modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
-3.2° on tip twist	+0.28%	+0.05°	+0.11°
-2.3° on root twist	-0.41%	+0.12°	+0.17°
-3.2° on tip twist and -2.3° on root twist	-0.10%	+0.17°	+0.28°

3.4.4 Effect of camber modification

Front blade camber level has been defined during the preliminary design step to fulfil take-off thrust requirements. The margin was significant, so the camber is likely to be diminished. A -30% factor is applied on the front camber distribution to assess the potential benefits.

As presented in Table 5, the loss of lift is recuperated by a large front pitch angle increase. Decreasing the blade camber level results in a profile drag reduction and so an increase of efficiency (Fig. 7).

Table 5 Efficiency evolution for camber modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
-30% on front camber	+0.42%	+0.64°	-0.05°

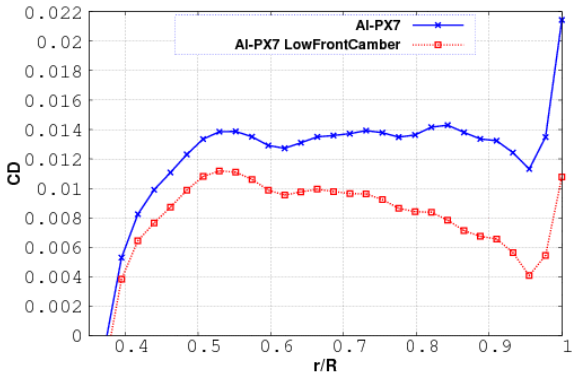


Fig. 7 Front blade sectional drag coefficient radial distribution (reference vs. lower camber)

3.4.5 Effect of thickness modification

Manufacturing or structural constraints are likely to limit the blade thickness to a minimal value. Two tests are therefore performed to assess the impact of such a modification. In the first case the blades are regularly thickened along the whole span to reach 1% more relative thickness at blade tip compared to the reference geometry. In the second case the blades are thickened starting from r/R=80% to the tip where there is a 0.5% increase of relative thickness.

Table 6 shows the deterioration of the performances in the first case, mainly caused by a drag increase.

Table 6 Efficiency evolution for thickness modification

Case	Efficiency (vs. reference)	β_{FRONT}	β_{AFT}
Global thickness increase (up to +1% relative thickness)	-0.47%	+0.04°	-0.01°
+0.5% tip relative thickness	-0.06%	< 0.01°	< 0.01°

In the second case the impact is very low. Actually, the reference blade tip is already very thin, and this result demonstrates that there is a margin to increase tip thickness without impacting the efficiency.

4 Unsteady computations

As a last step of the aerodynamic design process the robustness of the blade design to

flow non-homogeneities due to front/rear rotor interactions and installation effects (here incidence effect) has to be assessed.

The objective is to check that the performances of blades previously designed with the RANS iterative design process are still satisfactory under these unsteady flow conditions. This analysis is here performed by unsteady computations on the full model of the AI-PX7 configuration using the elsA solver.

4.1 Chimera technique

One suitable way is to perform URANS computations with the Chimera mesh technique [8]. Here, a background grid plus two cylindrical rotating grids for each rotor are linked through overlap regions as depicted in Fig. 8. Variables exchange between the different grids is done by interpolations in the overlap regions.

The interest of this technique is the ability to use a well-adapted grid with an adequate refinement on the blades without impacting the density in the background mesh. Moreover new geometries can be quite easily assessed replacing only the Chimera grids.

First applications of this technique on CROR have been published by Stuermer and al. [9][10][11]. The work performed on the Airbus AI-PX7 propeller is an application of recent developments and studies made by François and al. [12] who have defined best practices for such computations with the elsA solver.

4.2 Computations

The elsA solver is used to perform unsteady simulations on a 53 million nodes mesh with boundary layer meshing. For the performances assessment, the choice has been made to use a time step equivalent to 2 degrees of rotation, in order to simulate the main fluctuations associated to rotor-rotor interactions and incidence effect.

Starting from a converged solution (obtained after three rotations), an additional rotation is computed to get a converged cycle.

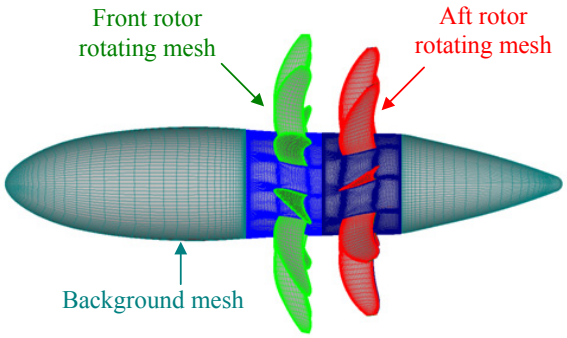


Fig. 8 Chimera mesh on AI-PX7 Configuration

4.3 Propulsion performances

As for the RANS computations, the first check is made on the performances. Such simulations provide reliable results on forces fluctuations encountered by a blade, at least for the first harmonics of the incidence effects and rotor/rotor interactions [12]. However Fig. 9 highlights, on the example of the thrust coefficient, the balancing that appears when taking into account the complete propellers. Only negligible rotor-rotor interaction fluctuations can still be observed. The result is the same for fluctuations linked to incidence effect, for which a slight increase of mean thrust appears (only 0.3% with a 2° angle of attack).

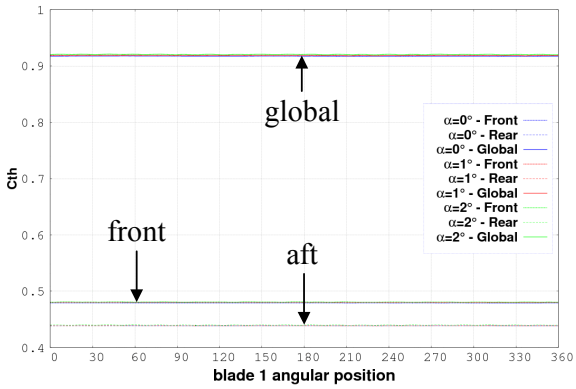


Fig. 9 Front, aft and global thrust coefficient fluctuations for 0°, 1° and 2° incidences

4.4 Aerodynamic field on blades

In addition to global performances it is interesting to observe the local unsteady aerodynamic phenomena occurring on blades. This can help in identifying parts of the blades that are not robust and are source of too much

deterioration of performances in unsteady conditions.

The analysis of the pressure field gives a good description of the flow. Fluctuations related to incidence effect can be observed on Fig. 10. Front blade 75% radial section pressure coefficient cuts are plotted for incidences 0, 1 and 2 degrees for the angular positions 120° and 290° which approximately correspond to the minimal and maximal load positions. For zero-incidence there is only a slight fluctuation, which is caused by the potential field interactions with the aft rotor. For 1° and 2° incidences, the rotor-rotor interaction fluctuations are weak compared to those related to the incidence effect whose period is one propeller rotation.

Those unsteady effects can also be observed on the aerodynamic load distributions (Fig. 11). In the present case, a good robustness of the blade can be noticed because there is no drop in the load distribution. Of course these fluctuations may have a significant impact on structural integrity or acoustic emissions and should be mastered but this is not the scope of this paper.

For the aft blade (Fig. 12 and Fig. 13), the same kind of fluctuations can be noticed. However, the amplitude related to the incidence effect (once per rotation) seems to be a bit less pronounced. This is due to the front rotor, which straightens the flow. Once again no unfavorable behavior is observed.

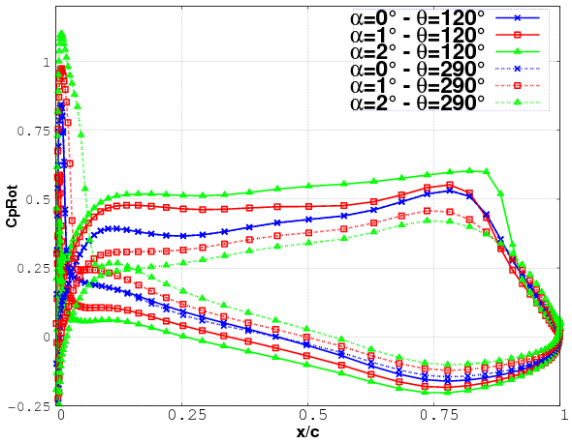


Fig. 10 Front blade pressure coefficient distribution (75% relative radial location) at 120° and 290° angular position (incidences 0, 1 and 2°)

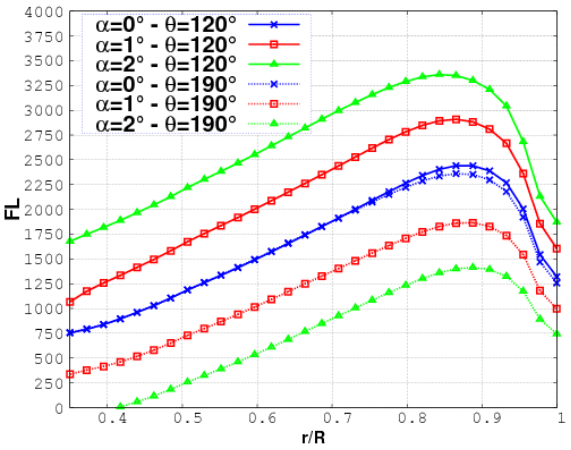


Fig. 11 Front blade radial aerodynamic load distributions at 120° and 290° angular position (incidences 0, 1 and 2°)

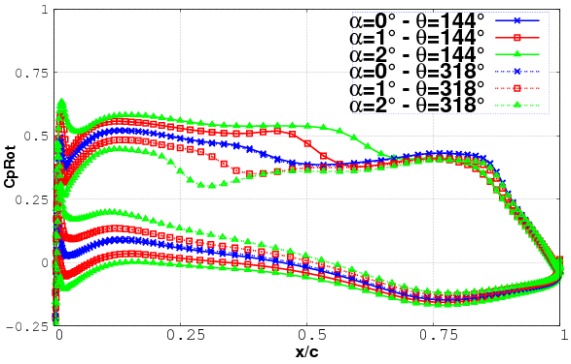


Fig. 12 Aft blade pressure coefficient cuts (75% relative radial location) at 144° and 318° angular position (incidences 0, 1 and 2°)

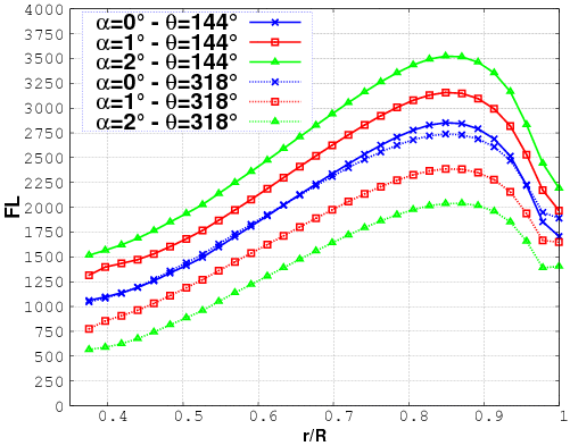


Fig. 13 Aft blade radial aerodynamic load distributions at 144° and 318° angular position (incidences 0, 1 and 2°)

A stronger impact of the incidence effect can be seen at the front blade root where the shock intensity is strongly increasing for 2° for the down-going blade (Fig. 14). Nevertheless no flow detachment is observed.

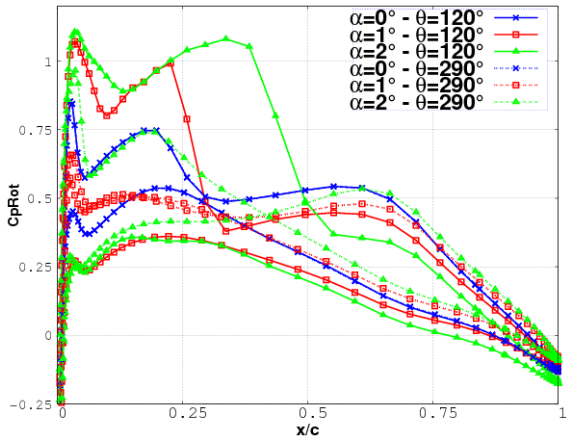


Fig. 14 Front blade pressure coefficient distribution (40% relative radial location) at 120° and 290° angular position (incidences 0, 1 and 2°)

5 Conclusion

A complete aerodynamic CROR blades design process has been presented. The lifting line approach permits to define a reference geometry. The performed RANS sensitivity study gives an idea about the main blade shape parameters influencing the propulsive efficiency. The unsteady computations analysis permits to focus on critical blades areas.

Extensions of this work will be the integration of the parametrical design process into an aerodynamic and finally an aero-acoustic optimization process as Marinus and Roger [13] or Pagano and al. [14] did on single-rotating propellers.

Acknowledgements

This work is carried out in the frame of a PhD thesis funded by Airbus. Authors would like to thank the Airbus Aerodynamics Methods and Tools Team for their work on CFD for CROR applications and tools dedicated to shape parameterization and mesh deformation.

References

[1] Hager R.D., Vrabel D., "Advanced turboprop project", *NASA Technical Report*, 1988.

[2] Saito S., Kobayashi H., and Nasu K., "Performance Calculation of Counter Rotation Propeller". Proceeding in 16th AIAA/CEAS Aeroacoustics Conference, 2010.

[3] Bousquet J. M., "Theoretical and experimental analysis of highspeed propeller aerodynamics". Proceeding in *AIAA/ASME/SAE/ASEE 22nd Joint Propulsion Conference*, 1986.

[4] Goldstein S., "On the Vortex Theory of Screw Propellers", *Proceedings of the Royal Society A,* 1929, pp. 440-465.

[5] Cambier L. and Gazaix M., "elsA: an efficient object-oriented solution to CFD complexity" Proceeding in *40th AIAA Aerospace Science Meeting and Exhibit*, 2002.

[6] PADGE User Manual.

[7] VOLDEF User Manual.

[8] Steger J., Dougherty F.C., Benek J.A., "A chimera grid scheme, Advances in grid generation". Proceeding of the *Applied Mechanics, Bioengineering and Fluids Engineering Conference*, 1983.

[9] Stuermer A., "Unsteady CFD Simulations of Contra-Rotating Propeller Propulsion Systems". Proceeding in *44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, 2008.

[10] Stuermer A. and Yin J., "Low-Speed Aerodynamics and Aeroacoustics of CROR". Proceeding in *15th AIAA/CEAS Aeroacoustics Conference*, 2009.

[11] Stuermer A. and Yin J., "Aerodynamic and Aeroacoustic Installation Effects for Pusher-Configuration CROR Propulsion Systems". Proceeding in *28th AIAA Applied Aerodynamics Conference*, 2010.

[12] François B., Costes M. and Dufour G., "Comparison of Chimera and Sliding Mesh Techniques for Unsteady Simulations of Counter Rotating Open-Rotors". Proceeding in *20th ISABE Conference*, 2011.

[13] Marinus B.G. and Roger M., "Aeroacoustic and Aerodynamic Optimization of Aircraft Propeller Blades". Proceeding in *16th AIAA/CEAS Aeroacoustics Conference*, 2010.

[14] Pagano A., Federico L., Barbarino M., Guida F. and Aversano M., "Multi-objective Aeroacoustic Optimization of an Aircraft Propeller". Proceeding in *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, 2008.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Turbojet Steady State and Transient Analytical Model Development and Validation

G. Medici, S. Chiesa and S. Corpino
Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, Italy

M. Balbo and A. Sbuttoni
Alenia Aeronautica, Corso Marche, 41, Torino, Italy

Keywords: *Simulation; UAV – Unmanned Aerial Vehicles; Turbojet; Transient Model;*

Abstract

This paper describes the modeling and validation procedure of a single spool Turbojet engine. This work was conducted during the collaboration between Politecnico di Torino and Alenia Aeronautica. The cooperation was carried out thanks to a Ph.D scholarship program (Fig. 1).

There was the need to define the working conditions of the turbojet (in particular of the turbine/nozzle stages), in order to achieve this target a steady state 1-D model was developed and validated. Starting from the matching algorithm of the steady state 1-D model, a transient 1-D model was developed and validated too. In order to verify the models, synthetic data (original engine deck), engine test bench data, and flight data was used. A graphical user interface (GUI) was developed on both models in order to give a better and more effective experience to the user.

1 Introduction

Since 2010 Alenia Aeronautica (joined with Politecnico di Torino) is offering an Industrial Ph.D scholarship program, in order to work on some of the company's strategic research fields (ie. [2],[4]). In particular in order to study a thrust vectoring Unmanned Aerial Vehicle (UAV) application, an initial survey on a turbojet turbine and nozzle performances was conducted. The turbojet engine studied pushes the Alenia Aeronautica's Sky-X UAV, a technological demonstrator (Fig. 2). In literature there are many examples of turbojet steady state modeling methods [5], and [8], some of them have been developed in order to calculate transients too [11]. The models presented in this paper are based on such methods; real turbojet engine data is used to validate the models. In particular: synthetic (calculated through a computer deck software) data, test bench data, and flight data are used.

This section introduces the research; in the second part of the paper the steady state model is described, while the third section contains the transient model. In the last part validation results are reported and commented.



Fig. 1 Ph.D Scholarship program

2 Steady State Model

In the first phase of the development a steady state analytical model was designed. The model provides a more detailed engine deck.



Fig. 2 Sky-X UAV

The engine to model features: single shaft, three stage axial compressor (fixed geometry, without any Inlet Guide Vanes IGV), single stage turbine, annular combustion chamber, fixed nozzle, pure turbojet. The simple configuration of the gas turbine, allows to model many components maps (that are unknown) [8], and [10]. Since many components lack of detailed information few simplifications were introduced in the model.

2.1 Steady State Model Layout

The software developed models a steady state, variable RPM and boundary conditions turbojet. The modeling platform is based on Matlab programming language. The routine plots properties values for each engine stage (thus is a 1-D model).

Matlab programming language is used due to its fast implementation yet powerful routines and matrix management tools; it provides a set of functions to build customized graphical user interfaces (GUI) with ease. During the calculation the isentropic coefficient, specific heat coefficient, and combustion chamber Temperature rise values depend upon Temperature and fuel/air ratio. This choice affects the algorithm process flow (it introduces few iterations), but even in a consumer laptop (Dual Core 2GHz Processor, 4 Gb DDR3 RAM) the model returns the calculation results within few seconds. The model can be controlled through a GUI (Fig. 3), that contains all the boundary conditions parameters the operator can set: altitude, and temperature relative to ISA standards, Mach (Ma), Fuel lower heating value (H_u), extracted power (P_e), percent of bleed air flow ($\beta_{\%}$), and air intake isentropic efficiency (η_{in}). Once every parameter is set and the user pushes the calculate button, a routine verifies that the analysis point requested lies inside the engine envelope (Fig. 4).

The modular nature of the software and the boundary condition initialization procedure allow this tool to be very flexible and customizable. It is possible, for example, to modify the operating envelope, by introducing the new conditions in a separate text file.

Other engine internal parameters, such as efficiencies, and iteration parameters, can be adjusted in a separate input file as well.

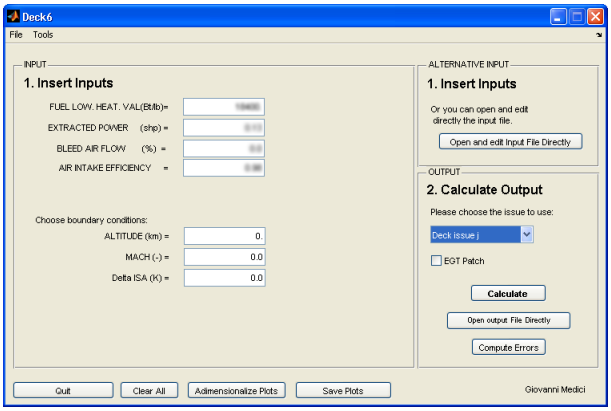


Fig. 3 Steady State Model GUI

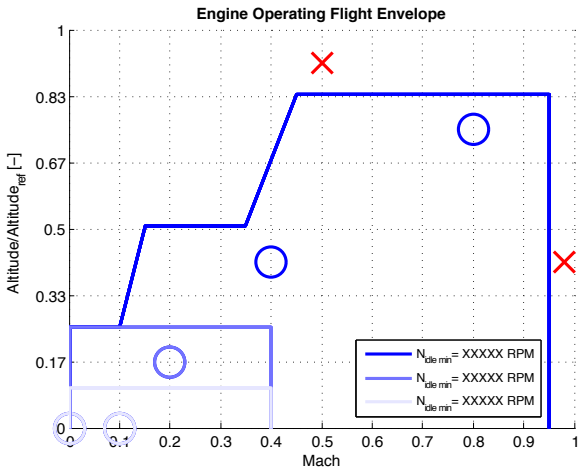


Fig. 4 Engine Operating Flight Envelope

By means of this approach the user can access rapidly to: compressor, turbine, combustion chamber, and mechanical efficiencies.

2.2 Steady State Model Algorithm

The model utilizes the method described in [9]. It is a 1-D (one dimensional) model, meaning that the turbojet engine was cut stage by stage in the longitudinal plane (Fig. 5), and every property is calculated as a mean value in each stage.

a. Ambient.
0. Intake inlet.
1. Intake outlet, compressor inlet.
2. Compressor outlet, combustion chamber inlet.
3. Combustion chamber outlet, turbine inlet.
4. Turbine outlet, nozzle.
5. Nozzle outlet.

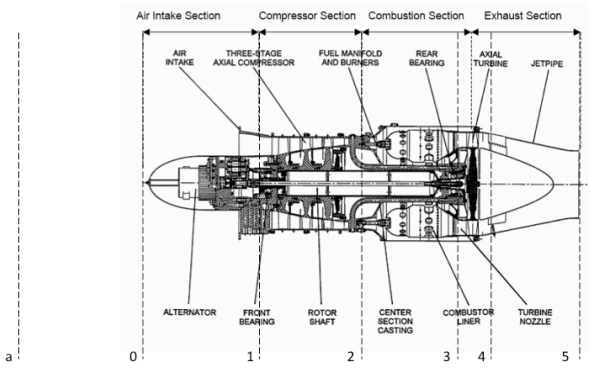


Fig. 5 Turbojet stage numbering

The particular case of a single spool turbojet steady state model leads to the following steps:

(a) The user sets the boundary conditions (altitude, Mach, and variation from the ISA day ΔT_{ISA}), inlet conditions are then calculated, by means of the air intake isentropic efficiency (η_{in}). In particular air intake inlet total temperature (T_0^0) is calculated through (2), where (γ_0) is the isentropic coefficient in the air intake inlet section, (Ma) is the Mach number at air intake inlet section and (T_1^0) is the total temperature at air intake outlet section (the flow starting at infinite and going through the air intake does not exchange heat and work thus the total temperature remains constant):

$$\eta_{in} = \frac{T_{1,is}^0 - T_a}{T_1^0 - T_a} \qquad (1)$$

$$T_0^0 = T_1^0 = T_a \cdot \left(1 + \frac{\gamma_0 - 1}{2}\mathrm{Ma}^2\right) \qquad (2)$$

There are total pressure losses due to the isentropic efficiency of the air intake; in the stages 0-1; thus total temperature and intake efficiency gives air intake's outlet total pressure (p_1^0) (3):

$$p_1^0 = p_a \cdot \left[1 + \eta_{in}\left(\frac{T_1^0}{T_a} - 1\right)\right]^{\frac{\gamma_0}{\gamma_0 - 1}} \qquad (3)$$

(b) Then picking a compressor constant-speed-line on the compressor characteristic, a data set of pressure ratio p_2^0/p_1^0, compressor corrected airflow $m\sqrt{T_1^0}/p_1^0$, compressor efficiency (η_c) is defined.

(c) The compressor total temperature ratio can be calculated through the (4):

$$\Delta T_{1,2}^0 = \frac{T_1^0}{\eta_c} \cdot \left[\left(\frac{p_2^0}{p_1^0}\right)^{\frac{\gamma - 1}{\gamma}} - 1\right] \qquad (4)$$

(d) The nozzle presents a single line characteristic map, and since the turbine characteristic do not change considerably with the RPM, a single line characteristic map can be used for this component too.

(e) This simplification allows an easier calculation of the complete turbojet cycle. In particular for a given compressor pressure ratio, p_2^0/p_1^0, the pressure ratio of the gas-generator

turbine can be found. This approximation effectively fixes the values of corrected turbine mass flow, $m\sqrt{T_3^0}/p_3^0$, and turbine temperature ratio, $\Delta T_{3,4}^0/T_3^0$.

(f) In this way it is no longer required to iterate both on the compressor and on the turbine side, single iteration is necessary to find the equilibrium running point on each compressor constant speed line.

(g) The compressor and turbine sides are linked by the (5) and (6):

$$\frac{\Delta T_{3,4}^0}{T_3^0} = \frac{\Delta T_{1,2}^0}{T_1^0} \cdot \frac{T_1^0}{T_3^0} \cdot \frac{c_{p,12}}{\eta_{mech} c_{p,34}} \qquad (5)$$

$$\frac{m\sqrt{T_3^0}}{p_3^0} = \frac{m\sqrt{T_1^0}}{p_1^0} \cdot \frac{p_1^0}{p_2^0} \cdot \frac{p_2^0}{p_3^0} \cdot \frac{\sqrt{T_3^0}}{\sqrt{T_1^0}} \cdot \frac{m_3}{m_1} \qquad (6)$$

(h) Where the iteration matches the $\Delta T_{3,4}^0/T_3^0$ value.

Once the method defines compressor and turbine's equilibrium points, every property of each turbojet's stage can be calculated.

The Matlab program algorithm contains few enhancements such as a fast iteration on each isentropic exponent and on the constant pressure heat coefficients (c_p). The combustion chamber properties are defined by means of a lookup table that links the combustion temperature rise with the fuel/air ratio and the inlet temperature. The original compressor map shows only few constant speed lines (8 going from 67% to 101%). A compressor map scale routine was developed in order to refine the compressor characteristic map by adding 28 scaled constant speed lines (4 constant speed lines inside each interval). The plot of a property with respect to the RPM results smoother, due to the scaling routine.

During the development of this Matlab program many turbojet's unknowns had to be estimated. The general map of a component (compressor, turbine, nozzle) was usually available, or could be calculated, but the components' efficiencies had to be estimated.

By choosing accordingly the components efficiencies, using reference engine examples, a close match with the experimental turbojet data is achieved.

2.3 Steady State Model Output

In particular the steady state model allows calculating a total of 56 parameters on five stages of the turbojet, while the original deck could only calculate 12 parameters. The source code can be accessed and adjusted as required (in order to adapt it to a different engine). The model provides various plot tools that allows creating graphics on multiple simulations too. For example it is possible to check in which way the turbojet cycle changes due to the variation of air intake efficiency, Mach, or altitude.

In Fig. 6, and Fig. 7, and Fig. 8 a comparison between two versions of the original deck and the actual steady state model are shown. The figures show some of the most important turbojet engine's properties: Thrust, Airflow and Fuel flow. The boundary conditions used in the comparison are described in Table 1. Results are presented with respect to the reference value of each property. The plots show a similar trend; results of the steady state model are close to the one calculated using the original decks. It should be reminded that the steady state model can not calculate the working points under 67% corrected RPM values, since the compressor map (constant speed-line) for such low speed is not known.

Table 1 Sea level standard boundary conditions

H_u	**42800 [kJ/kg]**
P_e	**0.1 [kW]**
$\beta_{\%}$	**0 %**
η_{in}	**0.98 [-]**
h	**0 [km]**
M_a	**0 [-]**
ΔT_{ISA}	**0 [K]**

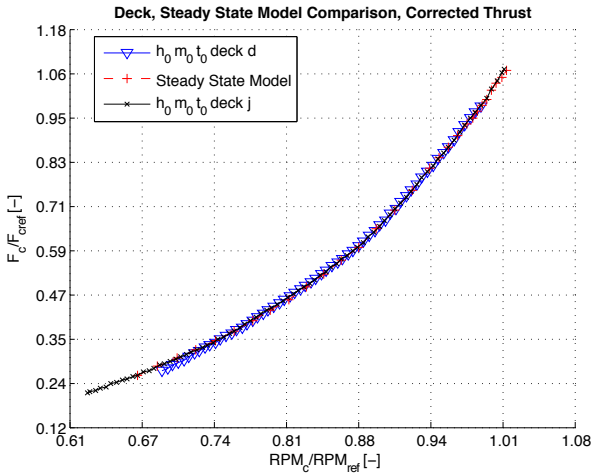


Fig. 6 Sea Level Static Corrected Thrust Comparison

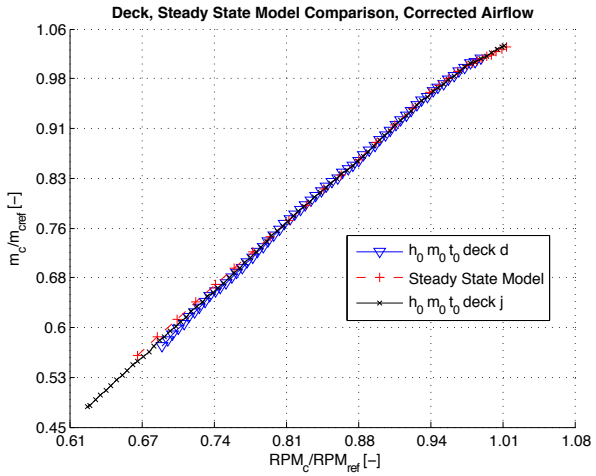


Fig. 7 Sea Level Static Corrected airflow Comparison

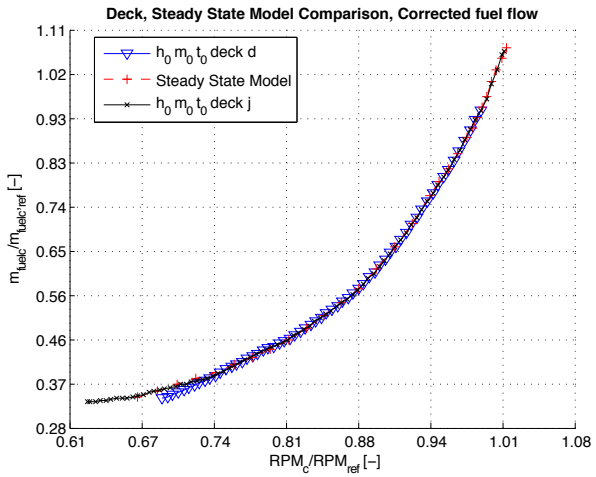


Fig. 8 Sea Level Static Corrected fuel flow Comparison

The boundary condition (input) data, and calculated one can be saved in a Matlab file. Data can be managed with ease, while in the previous deck text files were used as input and output, and data management could be performed only by a parsing routine on the text file, or manually by the user (that could give rise to errors due to particular text formatting of the output file).

3 Transient Model

During the second phase of the analysis, the transient model was developed. The transient model calculates Engine's dynamics, (properties variation vs. time) by dropping the work compatibility equation, and introducing the angular acceleration and inertial moments. The model uses the same core Matlab function developed in the steady state model to calculate every stage's property, but the time depending dynamics are managed by the Matlab Simulink tool. The model does not include any information on the wall temperatures (thus no inertial wall temperature model has been included). It is 1-D, meaning that every property is calculated as a mean value on each stage. The stages numbering is the same of the steady state model, and is shown in Fig. 5.

3.1 Transient Model Algorithm

Off-design equilibrium running calculations are based on satisfying the flow and work compatibility between the components. In the transient model while the flow compatibility is kept, the work compatibility between compressor and turbine drops; the net work between the two components cause the turbojet spool to accelerate or decelerate. During the setup of the model the first problem was to calculate the torque excess. The Newton's Second Law of Motion can be used to relate the acceleration of the compressor rotor and the excess of torque (ΔG):

$$\Delta G = J \cdot \dot{\omega} \qquad (7)$$

In the last equation (7) the angular acceleration of the rotor is ($\dot{\omega}$). The polar moment of inertia (J) was estimated using a dry crank test data. Initially the turbojet engine is accelerated to a known RPM speed. Then the torque (C_{known}) is removed and the engine's spool keeps rotating, by knowing the time the

engine takes to stop, its polar moment of inertia can be estimated (8).

$$C_{known} = J \cdot \dot{\omega} = J \cdot \frac{d\omega}{dt} \approx J \cdot \frac{RPM_{\text{dry crank}}}{t_{\text{to stop}}} \qquad (8)$$

The torque excess can be written considering turbine and compressor loads (9):

$$\Delta G = G_t - G_c \qquad (9)$$

Now by considering that the turbine and compressor power can be written as (10), and (11):

$$P_t = \eta_m \left(\dot{m}_a + \dot{m}_f\right) \cdot c_{p,34} \cdot \Delta T^0_{3,4} \qquad (10)$$

$$P_c = \dot{m}_a \cdot c_{p,12} \cdot \Delta T^0_{1,2} \qquad (11)$$

The torque excess can be calculated by using the (12):

$$\Delta G = \frac{\eta_{mech}\left(\dot{m}_a + \dot{m}_f\right) c_{p,34} \Delta T^0_{34}}{2\pi N} + \\ - \frac{\dot{m}_a c_{p,12} \Delta T^0_{12}}{2\pi N} \qquad (12)$$

Where ($c_{p,12}$) and ($c_{p,34}$) are the constant pressure heat coefficients related to the compressor and the turbine; the airflow mass and fuel flow mass are ($\dot{m}_a$) and ($\dot{m}_f$). Now by knowing the polar moment of inertia and the time step of the simulation is possible to calculate the transient running points of the turbojet. In particular by choosing accordingly the time step in Simulink, turbojet spool acceleration can be considered locally constant. Some assumptions can be made due to the simple geometry of the engine and relative slow dynamics (67% 100% slam takes some 6 – 8 seconds): in particular inside each component constant volumes are assumed, so that temperature and pressure changes instantly and mass conservation is assumed.

3.2 Transient Model Engine Control Unit

The engines dynamics are now defined, but the real turbojet contains an Engine Control Unit (ECU), that cuts fuel flow when some constraints thresholds are reached. By analyzing the engines time histories an elementary Engine Control Unit was modeled in Simulink. The ECU model changes the fuel flow in order to keep the engine within its structural and thermal thresholds:

1. Mechanical: when the RPM reaches 98% and the engine is accelerating, the ECU dumps the fuel flow ramp.
2. Thermal: when the Exhaust Gas Temperature (EGT), or its derivative, overcame a threshold the fuel flow is cut by 10%.

3.3 Transient Model GUI

The transient model shares with the steady state model the same Matlab sub-routines. The layout of the model is slightly different, as it must provide the feature of changing the boundary conditions during the simulation. For this reason Matlab Simulink interface was used (as shown in Fig. 9).

The Matlab Simulink environment is particularly effective for the simulation of programmable logic controllers and logics, such as Engine Control Unit [1], [3], [6], and [7].

The Simulink model uses some customized icons, this helps the user to understand what are the subsystem blocks related to (eg. turbojet routine, ECU…).

Once the user starts the simulation a figure, containing the engines compressor map shows up (Fig. 9 top). In the compressor's map the red star, indicating the current turbojet running point, moves in agreement to the actual engine's mechanical dynamics, leaving behind a trail (gray dashed line). Once the simulation ends a Matlab structure (with the same nomenclature as the steady state model) is saved to the hard disk, so that post-processing analysis can be performed. The post-processing tools allow to plot every property with respect to the simulation time; comparison between multiple data sets are also allowed. The validation technique will be described in the next section; by the way, in order to enable models results comparison with real flight test data, the

transient model contains two different configurations: classic simulation, and time history simulation. In the former: the operator changes the boundary conditions in real time during the simulation.

In the latter a flight test data file is used to provide boundary conditions (Altitude, Temperature, Pressure and RPM Command vs. time history) readings.

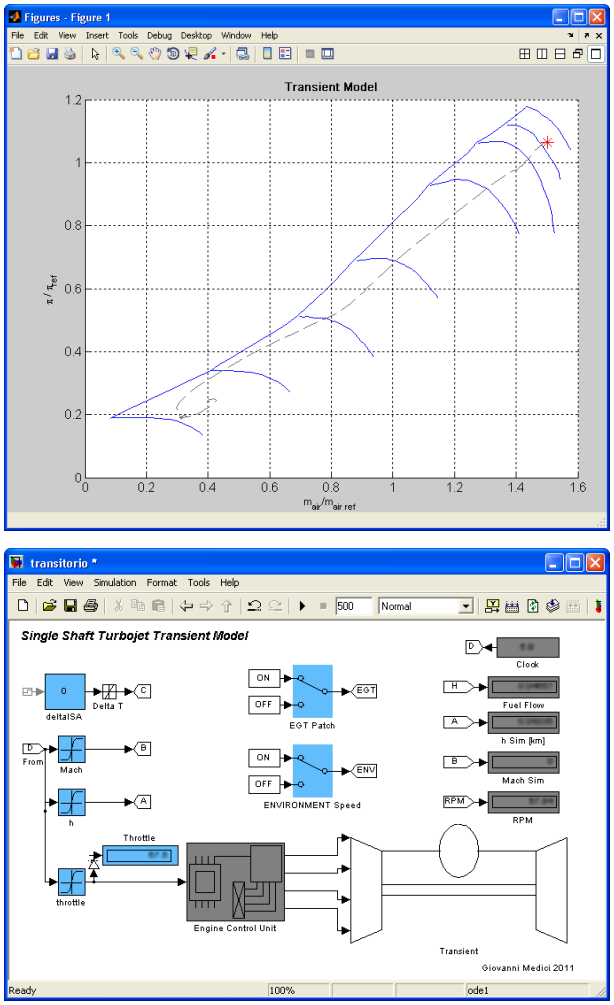


Fig. 9 Transient Model GUI

The transient model then calculates the engine properties according to the boundary conditions. Simulation result can be the plotted compared to real flight data properties (only few are collected by the Flight Control Computers: such as corrected RPM, exhaust gas temperature, and bleed pressure…).

The transient model developed allows to test the turbojet engine under limiting (critical) conditions (near surge-line), and outputs a complete data-set for each stage. The tool developed allows a more precise data analysis. Through the use of this tool is possible to develop software controllers that can establish critical engine conditions (such as critical EGT change rate, or off envelope intake working points…).

4 Validation and Comparison

Great accuracy has been found during the validation phase, with an estimation error module lower than 4% inside the flight envelope, for the steady state 1-D model (as depicted in Fig. 6, Fig. 7, and Fig. 8 for the Sea Level Standard case). Various tests were performed investigating standard cruise conditions and envelope boundaries. The results' general trend lies well within the engines scattering band (3 – 4 %). The transient model has been validated, as well as the steady state model, by using the flight data. The transient model's error (time delay) is of less than one second (as shown in Fig. 10 during the flight data comparison). The plot depicts the time history of the engine throttle command (dashed black line), the real engine flight data (solid blue line), and the model result data (solid red line). The initial gap between the two time histories is due to the fact that the compressor map used does not contain a constant speed line for low RPM speeds (under 67 % corrected RPM values). After the first slam the two time histories are almost overlapped. There are only minor differences in the RPM time history, and in the compressor pressure plot (Fig. 10, and Fig. 12). Exhaust gas temperature plot (Fig. 11) shows a different behavior since the internal thermal inertia of the components (and in particular of the real engine mounted EGT probe time constant) is not modeled. The EGT sensor captures data on the nozzle cone (just after the turbine rotor); some thermal paint tests have shown that the exhaust gas temperature gradient in the nozzle cross-section can not be neglected, and in particular the mean temperature of the flow is lower than the nozzle skin's temperature. The max error with the respect to the RPM is 4% with mean value of 0.19% (as shown in Fig. 13).

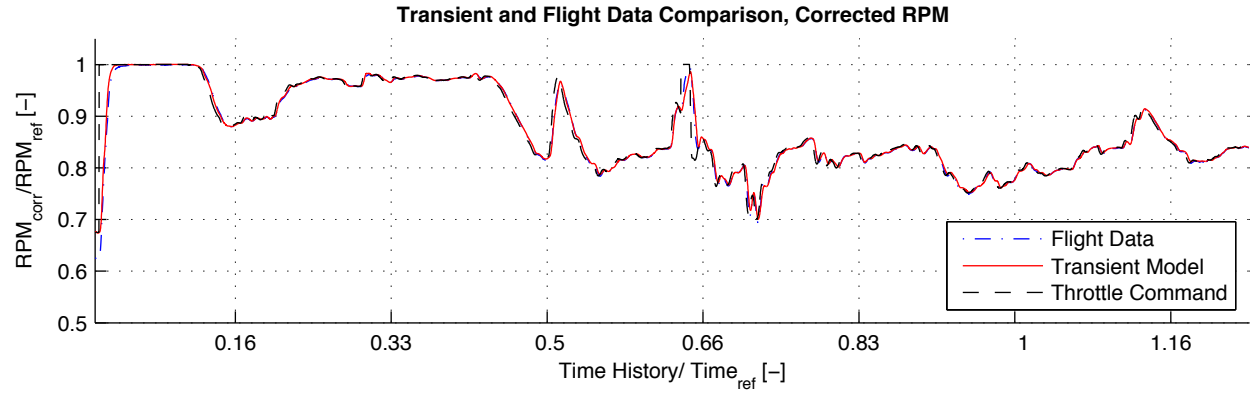


Fig. 10 Transient Model Corrected RPM Comparison

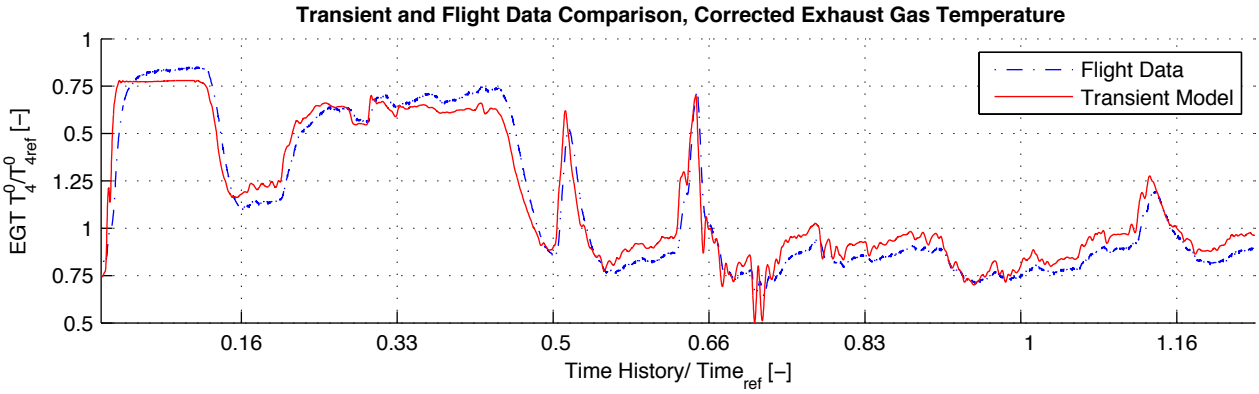


Fig. 11 Transient Model Exhaust Gas Temperature Comparison

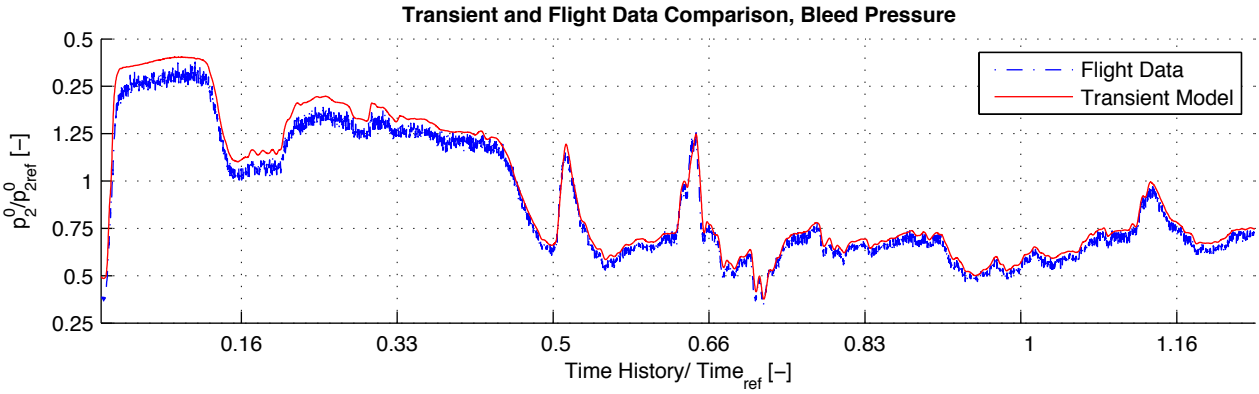


Fig. 12 Transient Model Compressor Bleed Pressure Comparison

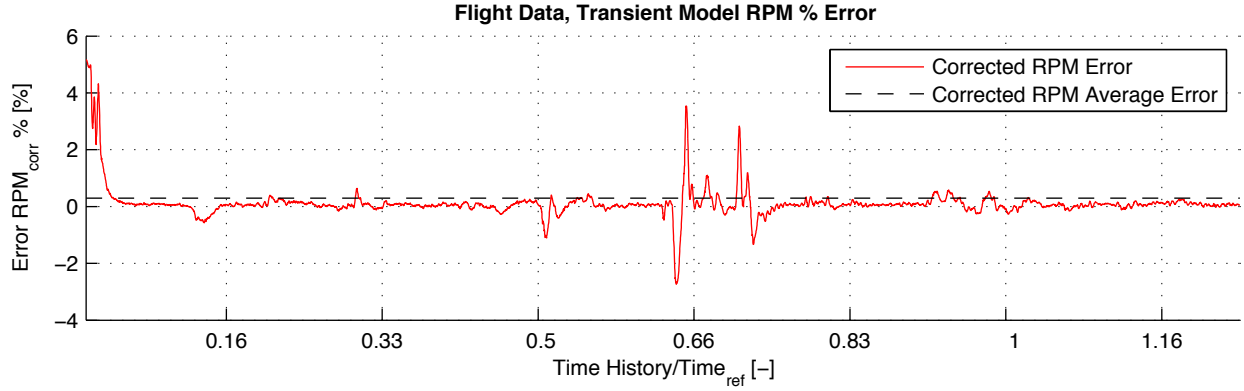


Fig. 13 Transient Model RPM error

In order to verify that the transient model calculates the same properties as the steady state model, a comparison has been made. By introducing a slow accelerating RPM command in the transient model, a quasi-equilibrium simulation can be performed (Fig. 14, Fig. 15, and Fig. 16). By plotting the desired properties with respect to the corrected rpm (neglecting the time history), the two models are then compared. In general the models are almost identical, there is a little region (with low RPMs) where the mass of fuel does not perfectly overlap, this may be caused from the custom ECU developed in the transient model.

5 Conclusion

Through the development of those models, much information on the engine dynamics has been found. In particular many components' efficiencies have been estimated. Using the analytical development described in this paper it was possible to define in a modular approach a well-defined analysis path:

1. Components characteristics estimation.
2. 1-D steady state model, development and validation.
3. 1-D transient model, development, validation and comparison with the steady state model.

This paper deals with the research made on a pure Turbojet engine. A detailed analytical engine deck has been developed. The analysis started with a minimum set of known values (very limited in terms of stages and properties), and led to a detailed 1-D steady model. By using the same model (neglecting the compressor turbine work balance, while introducing the angular acceleration), a 1-D transient model has been developed, and validated.

The steady state model calculates 52 parameters, while the previous turbojet deck provided only the 12 most important ones. Engine transient performances can be calculated and critical engine operation envelope can be defined with a better degree of detail (in particular during transients). The transient model error lies inside the engine performances scattering band.

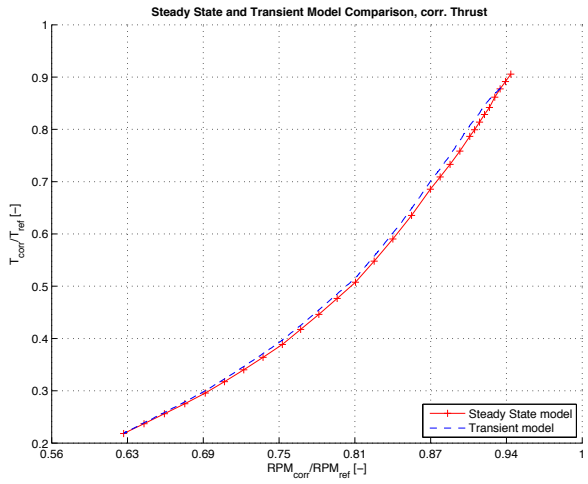


Fig. 14 Transient, steady state model comparison, corrected thrust

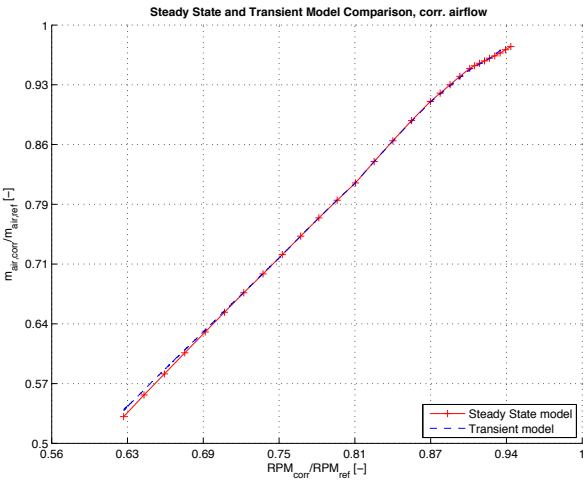


Fig. 15 Transient, steady state model comparison, corrected airflow

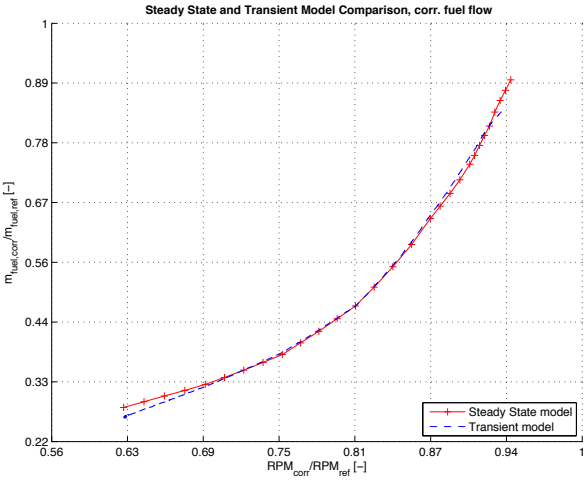


Fig. 16 Transient, steady state model comparison, corrected fuel flow

Acknowledgements

The present work has been performed through a close cooperation with Alenia Aeronautica S.p.A. and through the use of Confidential Information and Data, property of Alenia Aeronautica S.p.A., which remains the sole owner of all such relevant IP rights. The result of the present work shall be, therefore, property of Alenia Aeronautica S.p.A.

Authors wish to thank Eng. Renzo Bava, Eng. Marco Ferrero and Eng. Aldo Tonon.

References

[1] Chiesa, S., Corpino, S., and Medici, G., "System Programmable Logic Computer Aided Development Procedure", *8th ACD2010 European Workshop on Advanced Control and Diagnosis*, Ferrara, 2010, pp. 229-234.

[2] Chiesa, S., Farfaglia, S. and Viola, N., "Design of all electric secondary power system for future advanced MALE UAV", *27th International Congress of the Aeronautical Sciences, International Congress of the Aeronautical Sciences*, Nice, 2010, pp. 1-10.

[3] Davidson, C.M. and McWhinne, J., "Engineering the control software development process", *Factory 2000 - The Technology Exploitation Process, Fifth International Conference,* 1997, Cambridge, Vol. 1, Pag 247-250.

[4] Farfaglia S., Tranchero B., and Chiesa S., "The SAVE project: hypothesis and investigation strategies for alternative energy based systems for MALE UAV", *20th National Congress AIDAA*, Milan, 2009, pp. 1-18.

[5] Fielding D., Topps J. E. C., "Thermodynamic Data for the Calculation of Gas Turbine Performance", Aeronautical Research Council Reports and Memoranda, 1954.

[6] Gilberl, J.G. and Diehl, G.R., "Application of Programmable Logic Controllers to Substation Control and Protection", IEEE Transactions on Power Delivery, 1994, Vol. 9, pp. 384-388.

[7] Keller, J.P., "Interactive control system design", *Control Engineering Practice*, 2006, Vol. 14, pp. 177-184.

[8] Kurzke J., Riegler C., "A New Compressor Map Scaling Procedure For Preliminary Conceptual Design of Gas Turbines", *ASME TurboExpo,* 2000, Munich, 2000-GT-0006.

[9] Saravanamuttoo, H. I. H., *Gas Turbine Theory*, 4th ed., Pearson Prentice Hall, Harlow, 2009.

[10] Sieros G., Stamatis A. and Mathioudakis K., "Jet engine component maps for performance modeling and diagnosis", *Journal of Propulsion and Power,* Vol. 13, 1997, pp. 665-674.

[11] Sog-Kyun Kim, Pilidis P., and Junfei Yin., "Gas Turbine Dynamic Simulation Using Simulink®", *Society of Automotive Engineers*, San Diego, 2000, ISSN 0148-7191.

Optimization of a 100 daN Small Turbojet

Jeni A. Popescu, Valeriu A. Vilag, Cleopatra F. Cuciumita, Georgel Vizitiu and Valentin Silivestru
COMOTI Romanian Research and Development Institute for Gas Turbines, Romania

Keywords: *turbojet, compressor, CFD, numerical simulations, optimization*

Abstract

The paper presents the evaluation and the optimization methods for a small turbojet designed and manufactured at COMOTI Romanian Research and Development Institute for Gas Turbines in the 1980's – MTR 100. In the context of the existing small turbojets designated for drones and single seat jet aircrafts, MTR 100 is evaluated regarding its technical specifications and performances as they have been considered in the design stage of the project. Due to unsuccessful experimentations in the past, a reconfiguration of the turbojet is required, first considered step being the redesigning of the compressor using modern methods such as numerical simulations in CFD environment. The results are to be validated in the future through test bench experimentation.

1 Introduction

Several decades ago a series of small gas turbines were developed intended for equipping drones, small aircrafts, moto-sliders or even ground vehicles.

Built in the early 1970's the Budworth Puffin was a product of Budworth gas turbines LTD, a British company. This unit was clearly designed as a small propulsion unit. The design however proved troublesome and was modified later on. This engine was also used as a turbo shaft unit. The Puffin engine served as a gas generator which drove a power turbine. One application for this unit was as a blower for a hover craft [1].

The Noel Penny Turbines (NPT) 100 series engine was developed from the Budworth Puffin. NPT took over Budworth in the 1970s and went on to produce the 300 and 400 series engines. The 300 series engine was produced for the propulsion of target drones and missiles. Noel Penny 301 was developed to power recoverable drone type un-manned aircraft and UAVs. Two units were also fitted to a prototype four seat business jet, Leopard Jet, in the late 1980s. The NPT301 is an advanced simple cycle high reliability turbojet. The engine has a multiple mission role for use primarily in Remotely Piloted Vehicle applications. The NPT301 can be likened to a very large model aircraft power plant, it is similar mechanically, although slightly bigger, to the Mircroturbo TRS18 engine. Noel Penny Turbines planned to produce a whole family of small turbojet engines and also a turbo-shaft unit. Perhaps one of the UKs most powerful turbojets in its class, the Noel Penny Turbines 301 engine was the last design to be produced before the company folded in 1990 [1].

Rover/Lucas TJ125 (CT3201) Mini-Jet, built around 1975, the gas generator section of the Rover 2S series APUs and car engine was used as a turbo-jet engine for UAV and drone propulsion. The engine was only 22 kg and produced around 50 daN of thrust at full power.

The engine featured an innovative toothed belt drive to transmit power to the fuel pump and engine mounted electrical generator. This engine was fitted to the Epervier high performance rocket-launched reconnaissance drone used by the Belgium air force [1].

Williams WR24 Turbo-Jet Engine is a small turbojet engine with similar rotating components to the Williams WR9 APU. This relatively common engine could be considered a rival to the British built Lucas/Rover CT3201 engine [1].

The TJ 100 engine, developed by the Czech company PBS Velká Bíteš was designed for wide range of civil and military applications, such as UAV's, UCAV's or target drones. Civil version can be used as a drive unit for gliders, ultralight or experimental aircraft. The TJ 100 is certified for civil applications according to JAR-E. The main advantage of the engine is the compact design, small size, excellent thrust to weight ratio and complete equipment for operation including FADEC control system. Integrated 800 W generator and consumption 1.2kg/daN rank this unit among top products in this category. The newly developed TJ 100 C jet engine is intended for motorized gliders and experimental aircrafts [2].

2 MTR 100 small turbojet

The MTR 100, designed and manufactured at COMOTI Romanian Research and Development Institute for Gas Turbines in the 1980's, regarded as a suitable solution for small aircrafts such as moto-sliders or UAV's.

The choice of constructive solutions and functional parameters and the elaboration of fuel, lubrication and electrical schemes were based, from lack of reference models, on the experience of the researchers in previous project connected to the 50 daN Lucas and 100 kgf TRS18 units.

The utilization of this type of engines on sliders, apparently non-economical from the point of view of fuel consumption, was justified by the low weight, low frontal resistance, utilization of cheaper fuels and easier maintenance.

Table 1 Small turbojet engines

No.	Turbojet	Speed [rpm]	Thrust [daN]
1	Rover-Lucas TJ125-CT3201	65,000	50
2	Noel Penny 100	40,000	80
3	Budworth Puffin	40,000	89
4	Williams WR24	52,000	89
5	MTR 100	45,500	100
6	TJ100	60,000	100
7	KHD T117	45,000	105
8	Microturbo TRS18	47,000	133
9	Noel Penny LTD NPT301	40,500	133

MTR 100 is a simple cycle single-shaft turbojet with an almost classic mechanical layout including an annular axial flow intake, a single stage radial compressor, a single stage radial turbine, a reverse flow annular combustion chamber with fuel supplied through eight injection nozzles and a high speed jet nozzle. The compressor and the turbine are placed back-to-back on a common shaft, on the same side of the transmission allowing the bearings to operate in the cold area of the engine and simplifying the lubrication.

The inflow case and the transmission box consist of one part, leading to a compact transmission and reducing the length and the weight of the engine. The lubrication system is completely integrated in the engine without apparent pipes.

The starter, initially electrical, consists now of three nozzles in the compressor's case to connect with the pneumatic start system.

At 100 – 120 [daN] maximum thrust at sea level, 1,3 [kg/daNh] specific fuel consumption, 45,500 [rpm] speed and fuelled by kerosene, its resource has been initially evaluated at 500 hours [3].

The small gas turbine has been tested at COMOTI's experimental facility two decades ago, the measured performances showing a non-concordance with the initial design theme. The existing documentation presents experimental

data corresponding to operating regimes of up to 72%.

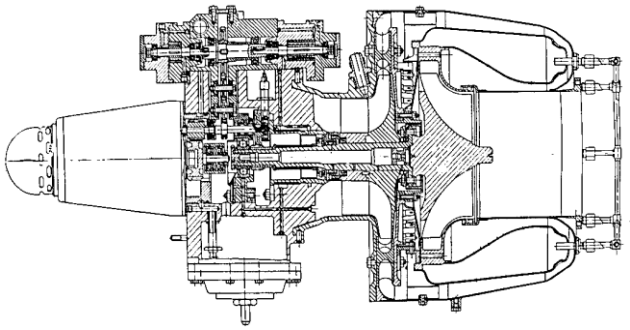


Fig. 1 The MTR 100 layout

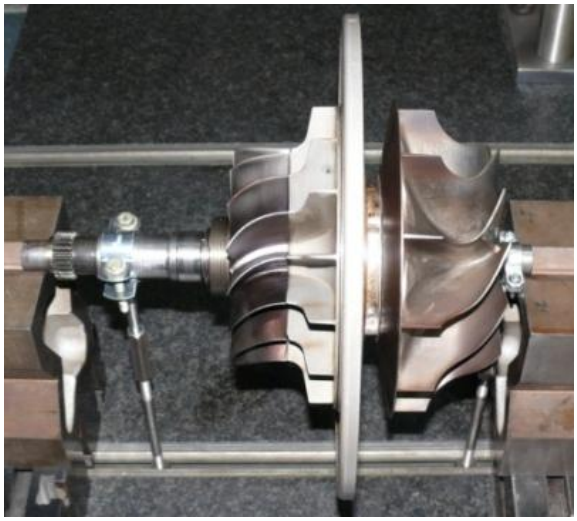


Fig. 2 MTR 100 rotor in present configuration

Research activities have been conducted along several years in order to adjust the performances with no satisfactory results.

3 Numerical analysis of the compressor

The project has been recently reactivated, mainly for didactic purposes. The analysis of old experimental data has led to the idea of improving the processes in MTR 100, particularly in the compressor and turbine sections. The existing compressor (single stage radial impeller with double-stage diffuser) and turbine are currently evaluated using numerical simulations in CFD environment.

In order to evaluate the performances of the existing components, the measurement of the parts was the first step for creating tri-dimensional models and then fluid models for simulating the processes in the turbojet with modern methods, in CFD environment.

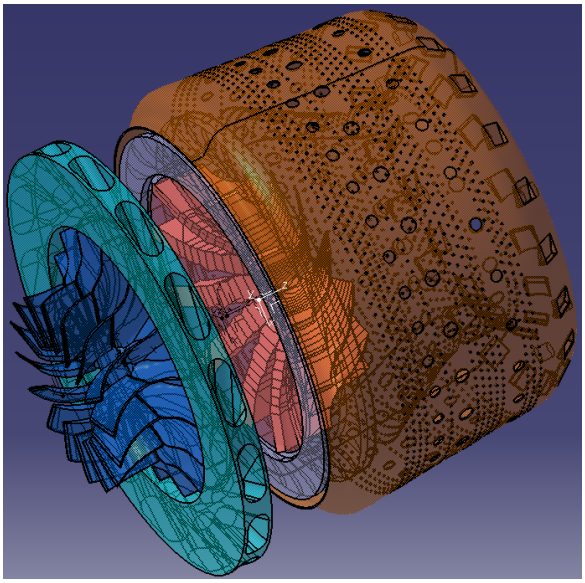


Fig. 3 3D models of the main gas-dynamic components of MTR 100

The paper presents the first phase of the evaluation, consisting in the numerical simulation of the flow process in the existing compressor of the MTR 100 turbojet [4].

A fluid volume has been modelled based on the symmetry hypothesis for one of the 16 blades of the radial impeller and the regions have been defined in order to apply the boundary conditions necessary for the numerical simulation of the process. A numerical grid of approximately 1 million volumes, with greater fineness in the blade area, has been created in ANSYS Workbench.

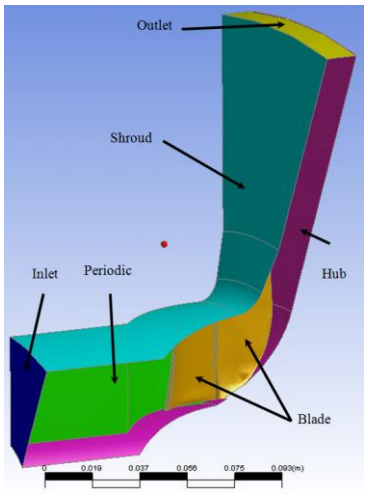


Fig. 4 Fluid model with regions definition

The numerical case presented has been defined for the steady operating regime of 72%:
- total pressure on inlet: 1 [bara]
- temperature on inlet: 288 [K]
- mass flow rate on outlet: 1.45 [kg/s] (per machine)
- rotating speed: 34,500 [rpm]
- turbulence model: SST
- fluid: air ideal gas

The numerical results consist in tabulated results, blade loading charts for different spans, streamwise charts for pressure, temperature, Mach number, alpha and beta angles and entropy, spanwise charts for Mach number and velocity on leading edge and trailing edge and blade, blade-to-blade, meridional and circumferential plots. The figures below present several results in pressure.

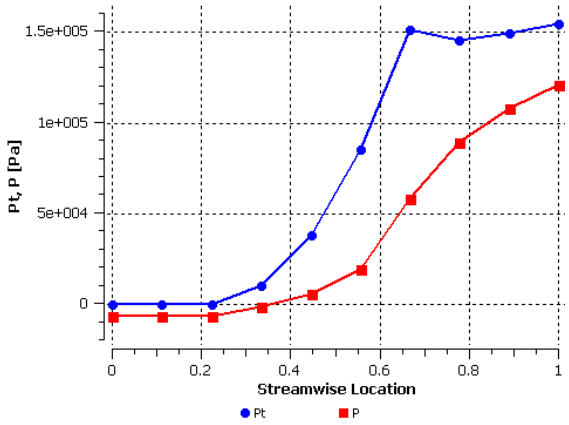


Fig. 5 Streamwise plot of pressure and total pressure on rotor

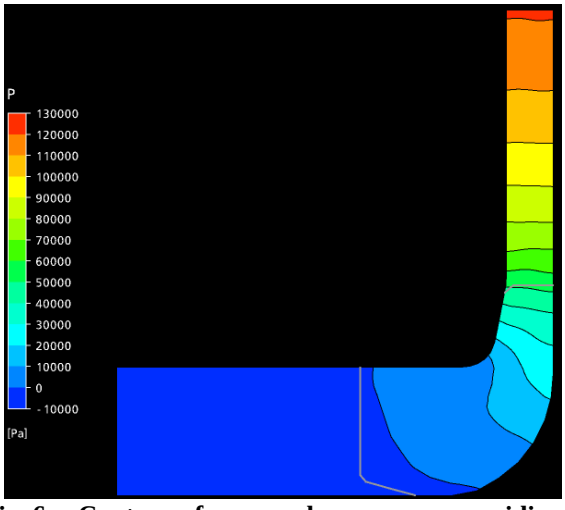


Fig. 6 Contour of averaged pressure on meridional surface of rotor

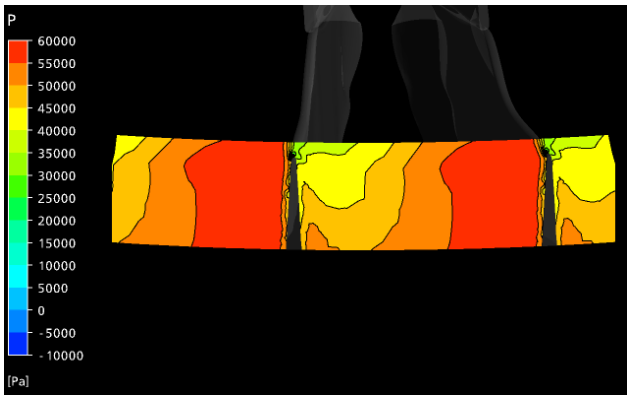


Fig. 7 Contour of pressure at blade trailing edge

The intermediate results obtained for the simulation of the flow process in the compressor's rotor:
- total pressure: 3.56 [bara]
- total temperature: 399 [K]
- compression ratio: 2.53
- isentropic efficiency: 78.9%

are only slightly different from the experimental ones, obtained for a regime of approximately 76% of the nominal speed:
- total pressure: 3.25 [bara]
- total temperature: 412 [K]
- compression ratio: 2.25

According to [3], the gas-dynamic pulsations occurring to high operating regimes are due to turbulence in the diffuser's area.

4 Redesigning the compressor

In the mean time the compressor is redesigned to match the initial imposed performances: pressure ratio of 3.8 at the maximum speed of 45,500 [rpm] and an air flow rate of 2 [kg/s] [4]. Additionally the alteration of the working channel is required in order to adapt the inlet and outlet sections to the existent geometry of the turbojet.

Analysing the existing geometry of the compressor, it may be noticed that the rotor consists in two parts, one axial and one radial, each with 16 blades, fastened at a very low radius, area where high losses are to be expected.

The iterative process of defining a new compressor's geometry consists in creating calculation models for finite volume numerical methods, setting the boundary conditions in the

inlet, outlet and intermediate sections of the compressor, analyzing the obtained performances and readjusting the channel and blade geometry in order to avoid detachments of the current and high pressure losses.

4.1 First stage redesigning

The main geometric elements to be modified in the first stage are: the number of the blades on rotor, the outlet diameter, the width and form of the canalization between the inlet section of the rotor and the outlet section of the diffuser, the number and the shape of the blades on the axial diffuser. The present radial diffuser may be replaced with a bladed one.

The definition of the new geometry has been made with the help of analytical programs allowing the construction of the calculation models for CFD in order to verify the performances.

Over 20 numerical cases [5] have been analysed for reaching relevant conclusions regarding the modifications to be applied. The paper presents the significant one defining the geometry of the rotor: 7 blades and 2 rows of splitters.

A numerical case with three domains has been created: a rotating domain corresponding to the section including a blade and two splitters, and including the intake channel; a stationary domain corresponding to one blade and one splitter of the bladed radial diffuser, including part of the working channel between the two diffusers; a stationary domain corresponding to a section including three blades of the axial diffuser, including the working channel between the compressor and the combustion chamber.

The first numerical cases have used a fluid model consisting in a rotating domain for the intake and the compressor's rotor and a stationary one corresponding to the radial diffuser. With the help of this model the geometry of the rotor and the corresponding working channel has been refined due to the advantage of a lower calculation effort. The model including the three domains has been used for optimizing the geometry of the axial diffuser.

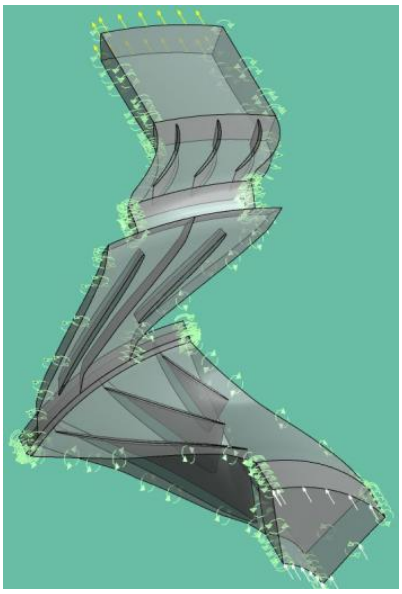


Fig. 8 Final fluid model with 7 blades and 2 splitters

The geometrical solution with blades and splitters has been used since the first version of the rotor, but it has been modified resulting the final one with 7 blades and 2 splitters between two blades assuring an improvement of the rotor's performances.

The peripheral speed in the outlet section of the rotor reaches a high value, 470 [m/s], due to the requested pressure ratio of 3.8. This fact, along with the ratio between the inlet diameter and the outlet diameter at the rotor's blade tip lead to the occurrence of an area with supersonic speed, approximately 1.6 Mach. The shape of the aerodynamic profile has been chosen to minimize the effects of this phenomenon.

Considering minimum alterations on the turbojet, the three domain fluid model has added the axial diffuser in the present configuration. The obtained results have been unsatisfactory; therefore the decision of redesigning it has been taken. The working channel has been modified in the bending area and the number of the blades has been reduced from 60 to 54. The performances of the compressor have been improved and the instable operation due to detachments has been eliminated.

The following boundary conditions have been set for the numerical cases:

- total pressure on inlet: 1 [bara]
- mass flow rate on outlet: 2 [kg/s] (per machine)

- rotating speed: 45,500 [rpm]
- turbulence model: SST
- fluid: air ideal gas

The numerical results obtained in the first stage of redesigning have indicated the following directions for the new compressor development:

- increasing the blending radius between the axial and the radial components of the rotor;
- eliminating the axial component of the compressor's rotor completely;
- decreasing the number of blades on the rotor and using splitters;
- increasing the diameter in the outlet section of the rotor for obtaining the desired pressure ratio and compensating the pressure losses between the two diffusers;
- eliminating the present radial diffuser, with divergent channels, and using a radial diffuser with profiled blades in order to decrease the pressure losses of the old geometry;
- redesigning the axial diffuser.

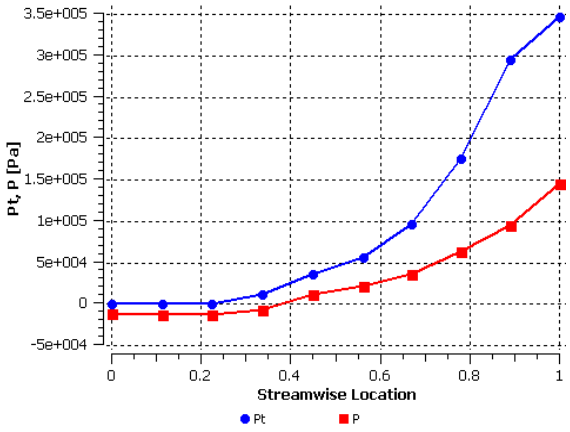


Fig. 9 Streamwise plot of pressure and total pressure on rotor

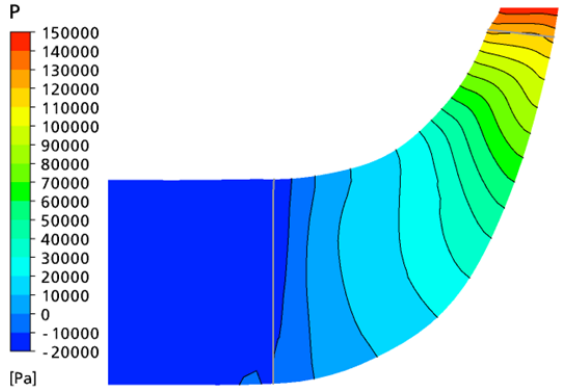


Fig. 10 Contour of averaged pressure on meridional surface of rotor

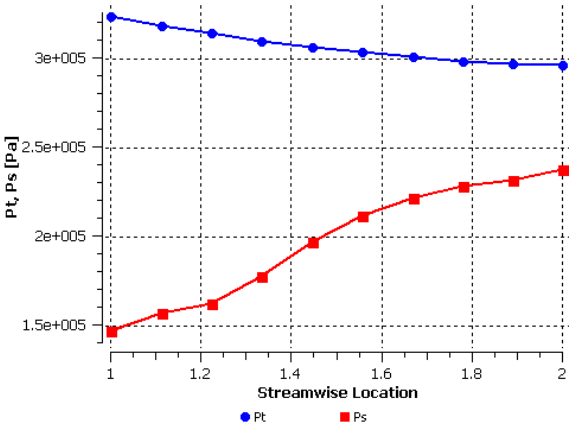


Fig. 11 Streamwise plot of pressure and total pressure on radial diffuser

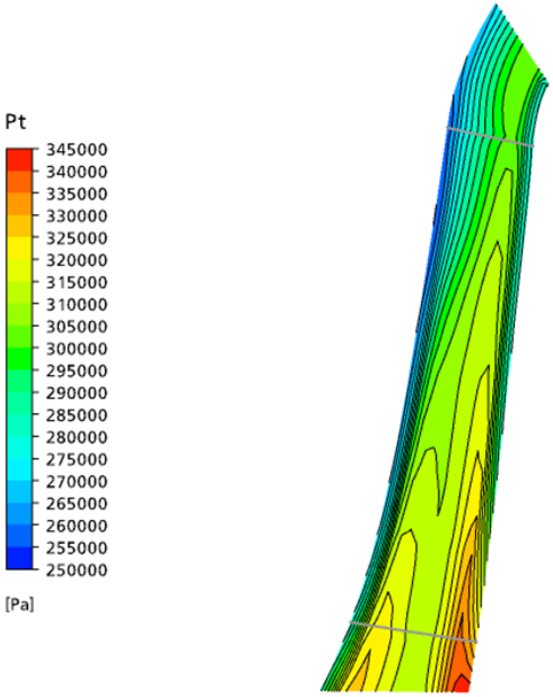


Fig. 12 Contour of averaged total pressure on meridional surface of radial diffuser

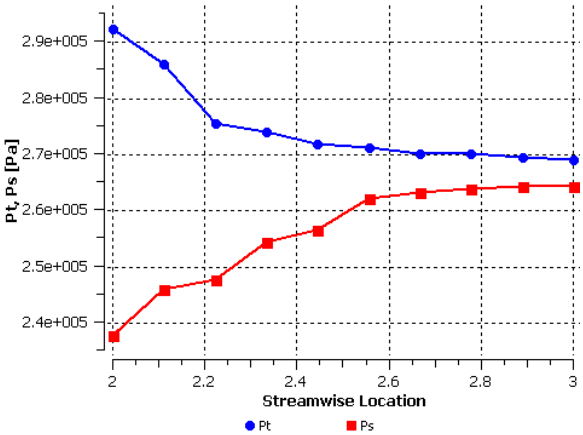


Fig. 13 Streamwise plot of pressure and total pressure on axial diffuser

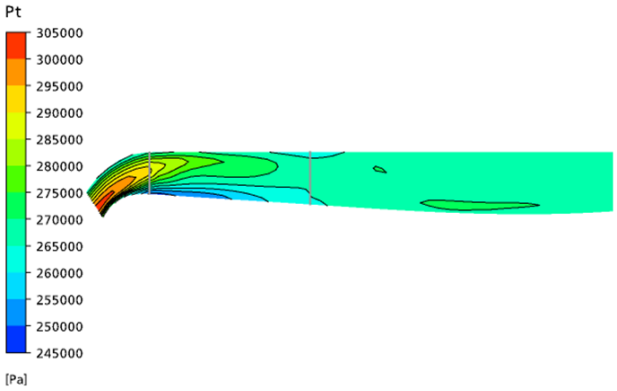


Fig. 14 Contour of averaged total pressure on meridional surface of axial diffuser

The following performances have been obtained for the imposed operating regime of the rotor:
- total pressure: 4 [bara]
- total temperature: 456 [K]
- compression ratio: 3.95
- isentropic efficiency: 80.5%

For evaluating other operating regimes, the air mass flow rate has been modified obtaining the performances in Table 2.

Table 2 Performances of the rotor at different operating regimes

Air mass flow rate [kg/s]	Compression ratio	Isentropic efficiency %
1,84	3,78	79
2	3,95	80,5
2,1	3,75	78,5

The performances calculated for the 2.3 [kg/s] air mass flow rate have shown a dramatic drop in pressure due to the occurrence of a supersonic speed area in the radial diffuser, leading to the conclusion that the component must be redesigned in order to re-establish the performances of the compressor for higher air mass flow rates.

4.2 Second stage redesigning

In the second stage of the redesigning process the following solutions have been applied [5]:
- keeping the configuration with the two diffusers – radial and axial;
- using a single blade for the radial and axial sections of the diffuser [6];
- using splitters in the diffuser bladed area;
- modifying the diameter of the rotor.

Several numerical cases have been solved aiming to minimize the pressure losses caused by the interaction of the high speed flow at the outlet of the rotor with the bladed section of the diffuser.

The diffusers have been kept distinct initially, modifying the position of the radial one along the working channel. In the same time, the position and the number of the blades of the axial diffuser have been modified. This solutions have been only partial successful.

The next step has been a slight increase in the diameter of the rotor, obtaining the desired compression ratio and a 1.5% increase in efficiency.

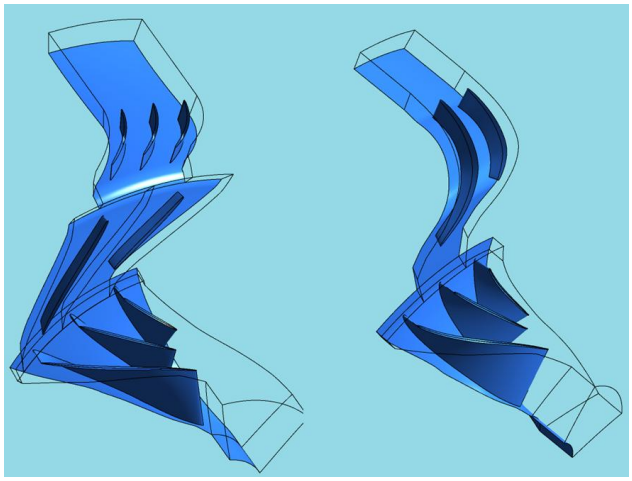


Fig. 15 Evolution of the geometric configuration – first stage and optimized solution

Figure 15 presents the evolution of the compressor's geometric configuration from the first stage of the redesigning to the result of the optimization process.

The disadvantage of the second version of the diffuser consists in a more expensive machining of the profile particularly in the bending area.

The final solution with single diffuser with 20 blades and splitters has better results regarding the global efficiency of the compressor, of at least 80%, and stable flow for a large range of mass flow rates. The choice has been confirmed by the comportment of the compressor for high values of the air mass flow rate.

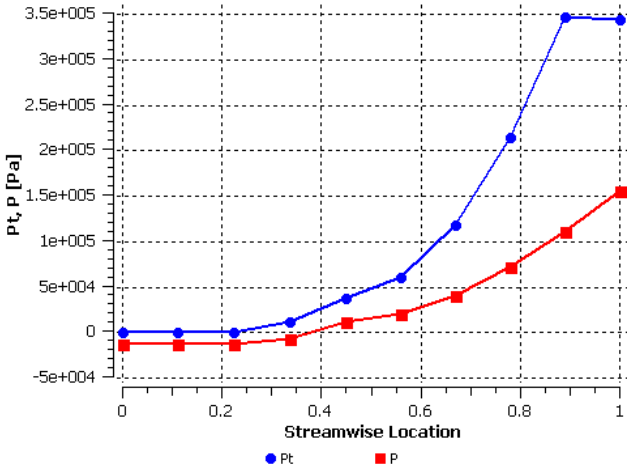


Fig. 16 Streamwise plot of pressure and total pressure on rotor

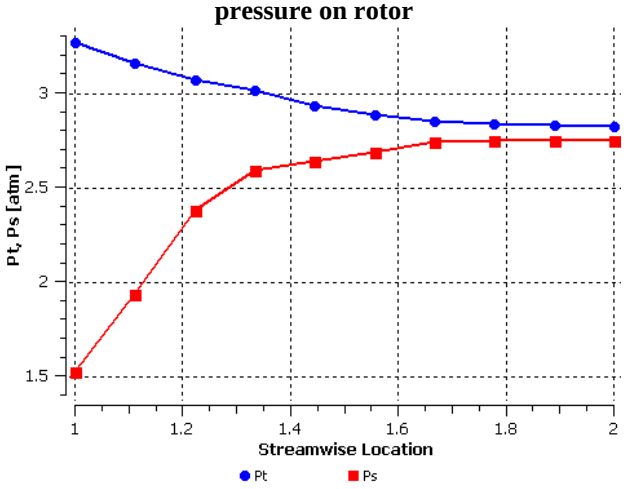


Fig. 17 Streamwise plot of pressure and total pressure on diffuser

The working characteristic of the compressor shows the variation of two parameters – the compression ratio and the efficiency – as functions of the air mass flow rate passing through at constant speed.

A partial compressor's characteristic, presented in Fig. 18 and 19, has been calculated for three operating regimes:

- nominal speed – 100% (red)
- 90% of nominal speed (blue)
- 80% of nominal speed (magenta)

12 points have been determined for each value of the speed by solving the numerical cases in the specified conditions of air mass flow rate and speed.

The diagram in Fig. 19 shows a large area of mass flow rates for which the efficiency is approximately 80%. Their width is at least 10% of the nominal air mass flow rate suggesting a stable operation of the compressor for high values of the parameters. Based on this observation, the diagram in Fig. 18 presents the desired operation line of the compressor (black) [5].

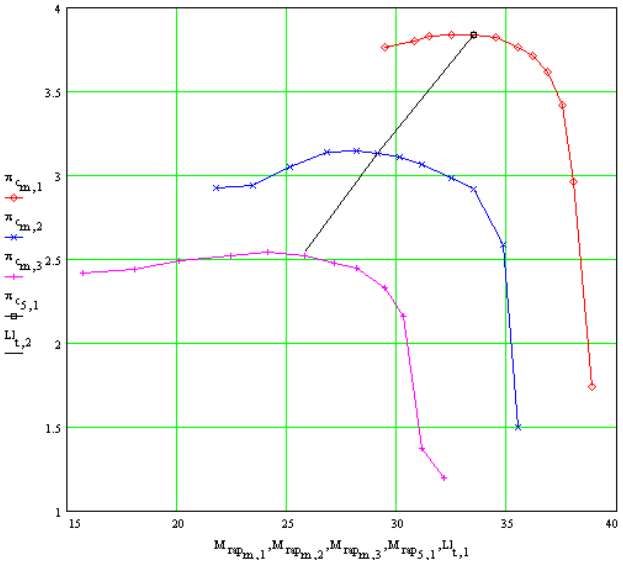


Fig. 18 Compression ratio per stage

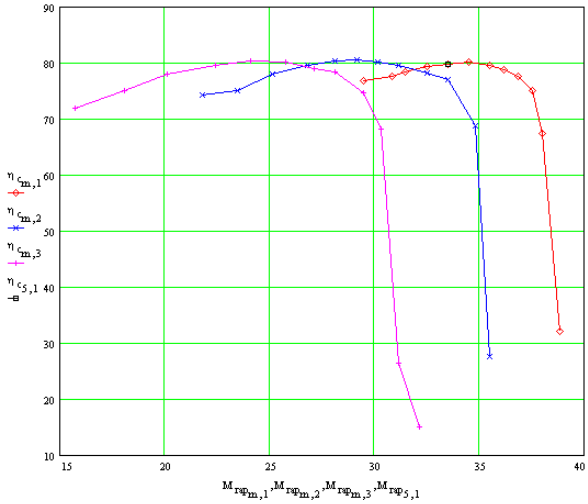


Fig. 19 Stage efficiency

5 Conclusions

The paper presents the MTR 100 turbojet, designed and manufactured at COMOTI Romanian Research and Development Institute for Gas Turbines in the 1980, similar to a series of small gas turbines developed in the same period. The unsatisfactory results of experimentation have led to a reactivation of the project for didactic purposes, the new researches aiming the redesigning of the gas-dynamic components of the turbojet.

The first results, regarding the compressor, are presented in terms of numerical simulation of the process in tri-dimensional fluid models of the existing part and the redesigned one in CFD environment.

Following the optimized solution for the compressor, presented in terms of gas-dynamics, stress calculation will be made in order to determine the final geometry of an efficient component. A similar method will be applied for the turbine.

The numerical results obtained for both existing and redesigned components are to be validated through test bench experiments in near future. The instrumentation layout and the experimentation program will be designed to obtain data in all important sections of the gas turbine in order to allow the comparison with the numerical results and the theoretical cycle analysis.

Test bench experimentations are an important stage in validating the numerical methods and establishing the final configuration of a gas turbine suitable for flight.

References

[1] www.gasturbine.pwp.blueyonder.co.uk
[2] www.pbsvb.com
[3] Lefcadus C., "100 daN Turbojet engine", COMOTI Technical Archive, 1986
[4] Popescu J., "Researches concerning the development of small gas turbines", COMOTI Technical Archive, 2009-2011
[5] Vizitiu G., " Gas-dynamic calculation and dimensioning of the compressor of the 100 daN turbojet engine", COMOTI Technical Archive, 2010
[6] Schreckling K., *Gas Turbine Engines for Model Aircrafts*, Traplet Publications, UK, 2003, Chap. 2

Piston engine cooling: an evergreen problem

Piancastelli Luca, Caligiana Gianni, Frizziero Leonardo, Marcoppido Simone

DIEM, Università di Bologna, Italy

Keywords: *Meredith effect, engine cooling, aircraft engines*

Abstract

Immediately before WWII (Wold War II), piston engine cooling has become a resource instead of a problem. The Meredith's paper defined a static engine that converted the heat dissipated by the engine into thrust.
Some solutions (see the De Havilland Mosquito, BF 109F, Supermarine Spitfire…) used hot air also to improve the efficiency of flaps.
In any case the Meredith effect proved to be effective even at the relatively low speed of F1 racing cars. The main problem proved to be the size of the device.
In recent times common-rail-turbodiesel piston engines are coming back onto aircrafts. However the experience coming from the racing cars cannot be directly applied to the flying machines for many different reasons.
First of all many racing-rules ban moving flaps, that are commonplace in aircraft. Moreover racing cars are relatively "dirty" reversed wing air machines, very different from the slender and elegant aircrafts.
Then the cooling system arrangement should be adapted to diesel engines and to high altitudes.
Finally, the high efficiency of the common rail diesel engine drastically reduces the cooling amount and consequently the Meredith effect.
This paper starts from the existing historical and new solutions and demonstrates that many improvements are still possible to improve the overall propulsion efficiency of new flying machines.

1 Introduction

The radiator pressure drop may be expressed by the formula (1)

$$\Delta p = \xi V^2 \qquad (1)$$

Where ξ is an experimental coefficient, V is the air velocity at radiator inlet, and Δp is the pressure drop.
The WWI (World War I) solution of the frontal radiator (see fig.1) proved to be prohibitively expensive in terms of power loss as aircraft velocities increased. At the beginning several radiator shapes were tested to overcome the problem.



Fig.1: Frontal radiator arrangement

On August the 14th ,1935 the British A.R.C. engineer F.W. Meredith published the memorandum *Note on the cooling of aircraft engines with special reference to ethylene glycol radiators enclosed in ducts*. In this memorandum Meredith showed how to design the "perfect radiator", not only for cooling, but also for drag reduction. Moreover, if the aircraft speed exceeds a certain value, the drag turns into a positive thrust.

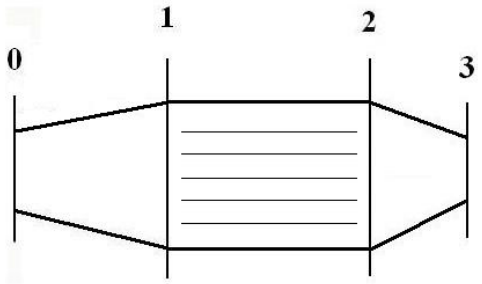


Fig.2: Shows a schematic drawing of a meredith radiator arrangement.

Figure 2 shows a schematic drawing of a Meredith radiator arrangement composed by a diffuser (section 0-1), the radiator (section 1-2) and the nozzle section (2-3). In Meredith's paper the radiator drag turns into a positive thrust over 300 mph.

The first prototype "V1" of the Bf 109 had a Rolls-Royce Kestrel with a chin radiator mounted close to the propeller (underlined in green in fig. 3). This arrangement suffered of severe cooling problem, mainly due to the propeller downwash.

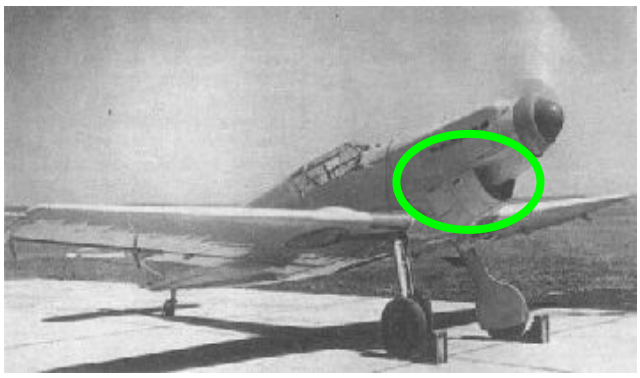
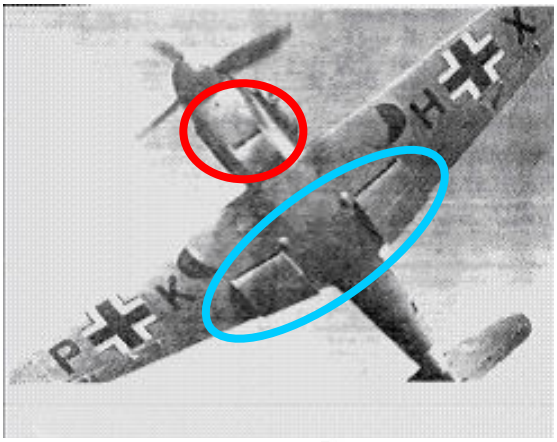


Fig.3: Bf 109 V1

With the introduction of the DB601 engine, the cooling system was redesigned. The single large radiator was replaced by a smaller oil cooler (rounded in red in figure 4), while the cooling system was obtained by two radiators which were partially included within the wings (rounded in blue in fig. 5).



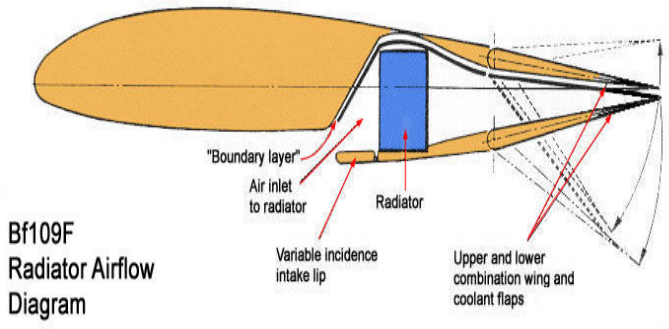


Fig. 4 and 5 - Bf 109 F cooling ducts

Cooling air was slowed down by the short diffuser and passed through the radiator, and the coolant temperature regulation was achieved by means of a combination of the rotations of the

two flaps at the aft edge of the radiator. A boundary layer duct was also added (see fig.5).

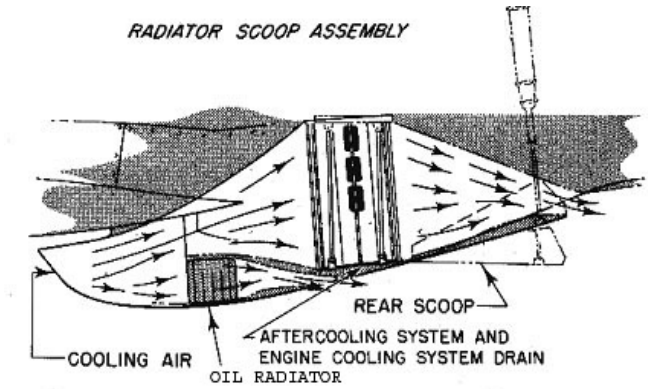




Fig. 6 and 7 - Mustang P- 51D arrangement

In the P51, the ducted radiator was located under and behind the pilot (rounded in red in figures 7). The oil radiator was separated from the main radiator, which includes the engine radiator and the aftercooler, and had an independent outlet. The main cooling regulation was given by an automatically actuated flap at the end of the duct. On the P-51D, an additional thrust of 2000 N due to Meredith effect was estimated.

However the Meredith paper didn't come from an original idea, it was simply an engineering analysis of a well known solution.

In fact, dutch engineer Jugo Junkers patented in 1915 the so-called "ducted cooling system" (see figure 7). In figure 8 the cooling duct patent issued as Reichspatent # 299799 to Hugo Junkers on Jan. 15, 1915 is depicted.

The dimensions are all there and still valid today. The JUNKERS J2 was the first and only aircraft with this installation, because the advantages in cooling and aerodynamics were of only minor importance in the early days of aviation.

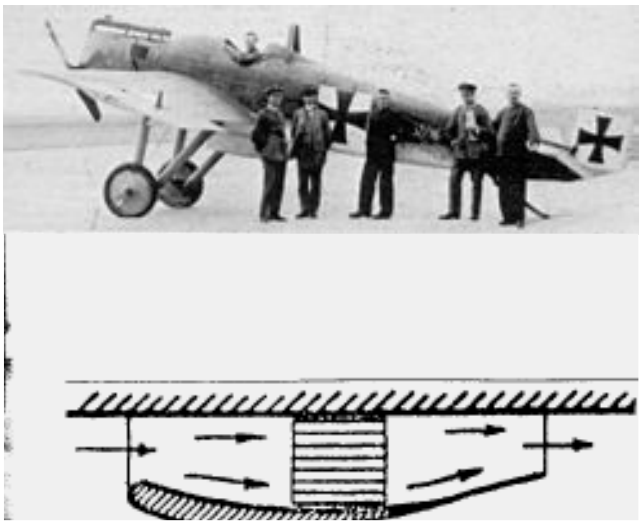


Fig. 8 and 9: the Junkers J2 "cooling duct"

During WWII, the Meredith-effect ring-radiator (figure 10) was also used by the Germans. The ring-radiator simplified the implementation of the "power pack" solution. In fact German engines had relatively short TBO (Time Between Overhaul). An advantage was the relative lightweight armour to protect the segmented frontal radiator system. In figure 11 a ring-radiator solution for a 4-in-line HALE (High Altitude Long Endurance) turbodiesel is depicted. Due to the necessity of high altitude cruise, several turbochargers, intercoolers and the aftercooler are needed. As it can be seen the piping arrangement from rear to front may be complicated. This is a shortcoming of this solution. Flaps are to be added to regulate the hot air outflow.

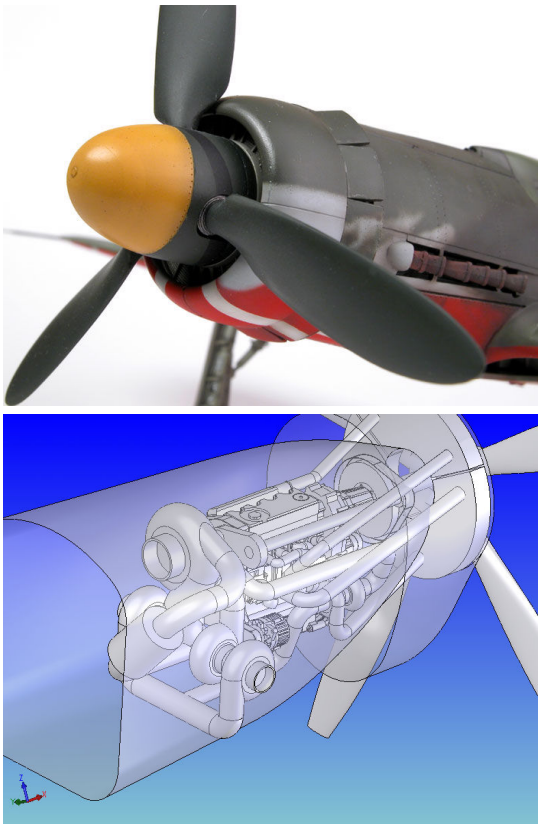


Fig. 10 and 11: The ring radiator for an high altitude diesel and for the FW190A

Again the ring radiator was not entirely new, since it was already adopted in the NACA cowling [11] and in several previous arrangement of air-cooled radial engines, again in order to achieve additional thrust from heat.
The De Havilland Mosquito arrangement combines the advantages of the Meredith effect to the improvements on the wing lift with or without flaps. See figure 12.

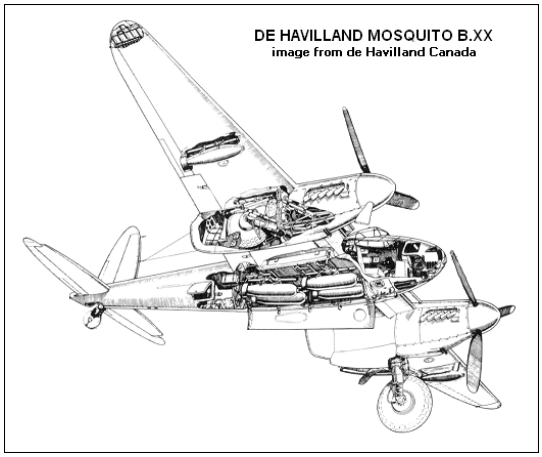


Fig. 12: The DeHavilland Mosquito

2 State of art

The NACA report 896 [4] introduces a complex relation to evaluate the radiator pressure drop starting from air velocity, temperature and density. Besides, Gothert [4] suggested to build a relatively long diffuser and to add fins to limit vortex and separation. Gothert [4] also demonstrated that at high speed there's always a drag reduction, because the radiator drag is largely compensated by the heating on the nozzle thrust, due to the increase of the pressure drop and the higher kinetic energy.

The better form available on the market [2] is a compact heat exchanger with finned flat tubes (figure 13). The BF 109 F had a radiator with symmetric airfoil tubes to minimize drag.

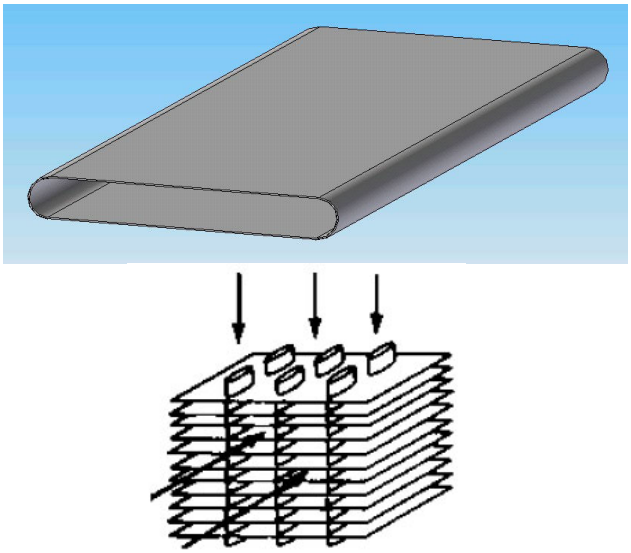


Fig. 13: Compact heat exchanger with flat tubes

The basic radiator formula is the following (2) :

$$Q = U\ S\ \gamma\ \Delta T_M \qquad (2)$$

In particular, in [2] it is possible to choose between two different types: the 9.68-0.87 and the 11.32–0.737-SR. The second one is more efficient, but involves more risk of fine dust deposit, dirtiness and breaking. Hence the first model is the better choice for reliability.

3 The "nearly optimum" solution

The optimum radiator duct is designed for an automotive conversion of the Smart CDI engine (see table 1).

Maximum theoretical power [HP]	90
Number of cylinders	3
Number of injectors for cylinder	1
Cylinder capacity [cm^3]	266
Compression ratio	15.5
Crankshaft rotation speed @max power[rpm]	4400
Total dry weight [kg]	60

Table 1 – diesel engine basic data

The first step is the design of the radiator. In this case, the worst condition is steep climbing just after the take off with maximum outside temperature. However, this approach is too conservative, in fact the radiator is excessively over-dimensioned for cruise and maximum speed. For this reason the radiator for the diesel engine is designed for cruise speed condition and then increased of 20%. With spark ignition engines the consequences of this solution are: a probable overheating in steep climbing condition and a limitation of the maximum taxiing time. Usually the overheating during climb can be controlled by reducing the climb angle, while the taxiing problems are absent in diesel up to a temperature of ISA+20. The turbocharging system should also have been water-cooled in order to reduce the inside temperature and to maximize the Meredith effect. However our automotive-derived Mistubishi turbocharger is air cooled.

3.1 Various hypothesises on the inlet section of the ducted radiator

The temperature rise in duct is assumed to be 57°C, from (ISA-0=15°C to 72°C) due to maximum coolant temperature of 100°C. However for the diesel turbocharged engine, we have a frontal aftercooler radiator that works from 15° to 35°C, then the "water" cooling radiator. For this reason our liquid cooling radiator works with a nominal temperature rise of 37°C (35 °C to 72°C) Then, given the cruise speed of 200km/h@300m, it is possible to calculate radiator frontal area with the formula (3).

$$S = Q / (\rho V C_{pa} \Delta T)*1.2 = 0.024 \text{ m}^2 \qquad (3)$$

The inlet "shark mouth shaped" under the propeller spinner seems to be the best solution, a very long diffuser can be then installed under the engine (figures 14 and 15).

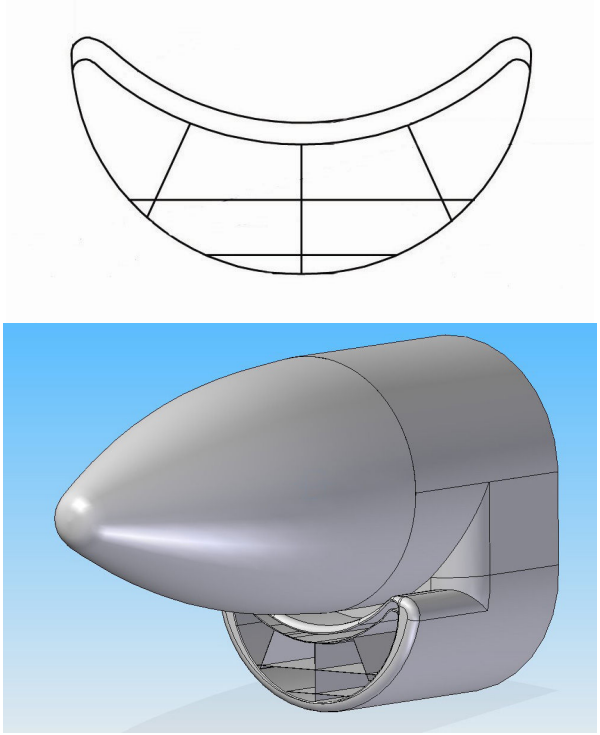
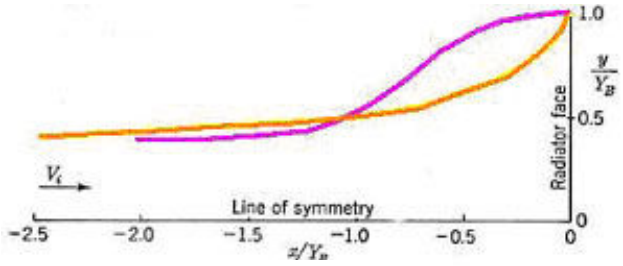


Fig. 14 and 15 – The duct inlet section

The consequences caused by the positioning of the air intake near the propeller are known and are well illustrated by [8] and [9]. The propeller increases the total pressure in front of the radiator approximately 7% of the free-stream dynamic pressure at the high speed thrust coefficient of 0.02 (i.e. at cruise speed) and approximately 45% of the free-stream dynamic pressure at the climb thrust coefficient of 0.11.

On the other hand, there's an effect of propeller operation on total pressure distribution at the face of the radiator.



ape [16]

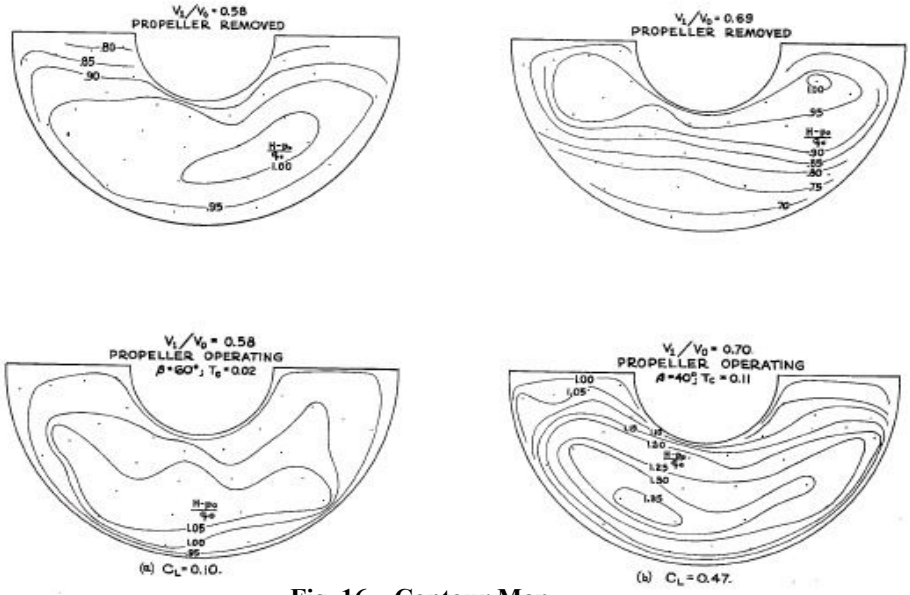


Fig. 16 – Contour Map

Another visible solution is the *inlet guide vanes* to reduce the vortex in the diffuser [8] and [9]. The *outlet guide vanes* should be added as well [9].

3.2 The diffuser

A tentative opening angle of the diffuser smaller than 7° is used.
Actually, the presence of the diffuser causes an increase in the radiator losses, because of the deviation of the airflow in front of the radiator. The optimal profile is showed in figures17-18 with an orange line: it's very similar to a trumpet. In this case, the internal losses are reduced (see figures 17-18).

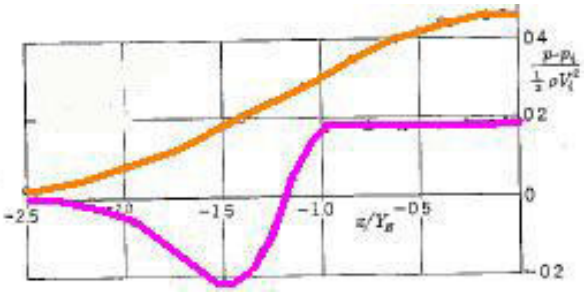


Fig. 18 Diffuser pressure rise [16]

The oil radiator is absent since the diesel engine as a water-to-oil internal heat exchanger.

3.3 The heat exchanger

The radiator tubes geometry can be modified to minimize the drag coefficient of the radiator

core, passing from the flat tube to a NACA 0014 profile (figures 19 and 20).

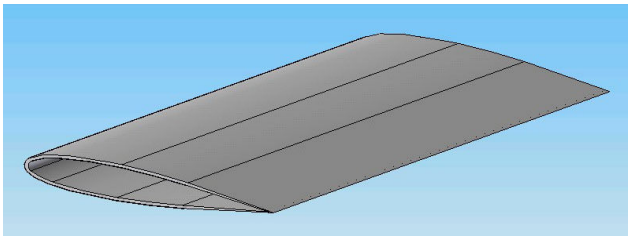


Fig. 19

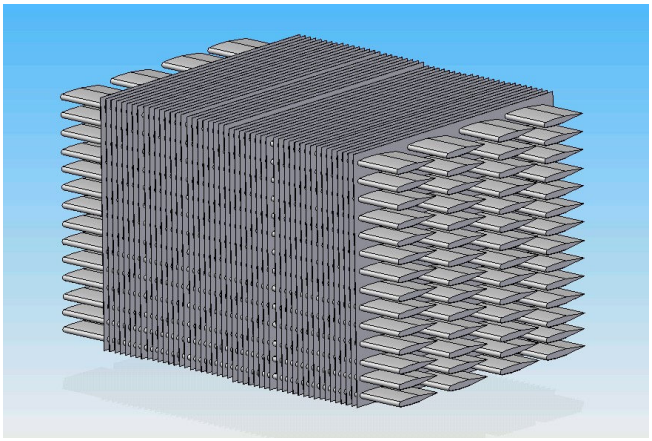


Fig. 20

With this profile the overall thermal conductance U is kept equal to the flat tube one with a significant reduction of the drag coefficient.

The presence of the fins elevates the value of the drag coefficient with an exponential law in function of the fin pitch (figure 21).

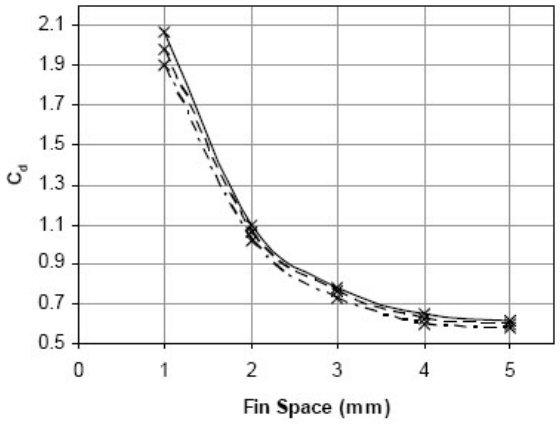


Fig. 21 Drag increase coefficient due to fins [16]

From figure 21 [16] a 93% increase in the drag coefficient of the isolated profile was estimated. The array effect is important for lift, while it's negligible (in particular for low Mach number) for drag. Finally, we considered indirectly null the angle of attack. From Hoerner's book [3]:

$$C_{Df} = \frac{2.656}{\sqrt{Re}}$$

$$C_D' = 2\, C_{Df}\, [1 + 2\, t / c + 60\, (t / c)^4]$$

$$C_D = 1.93\, n\, C_D'$$

It is then possible to calculate the pressure drop. From this calculations a further pressure increment of about 6% for the NACA 0014 profiled tubes over the flat tubes radiator can be obtained.

3.4 The nozzle

Both the distributed and the concentrated losses of the nozzle are negligible. So, the only problem is to evaluate the influence of the nozzle on the drag coefficient of the radiator.
From [8] the most important parameter in dimensioning the nozzle is the distance *a* between the outlet section and the radiator. In fact, an increase in a/H_B reduces the rise of the radiator drag coefficient due to the presence of the nozzle (figure 22).

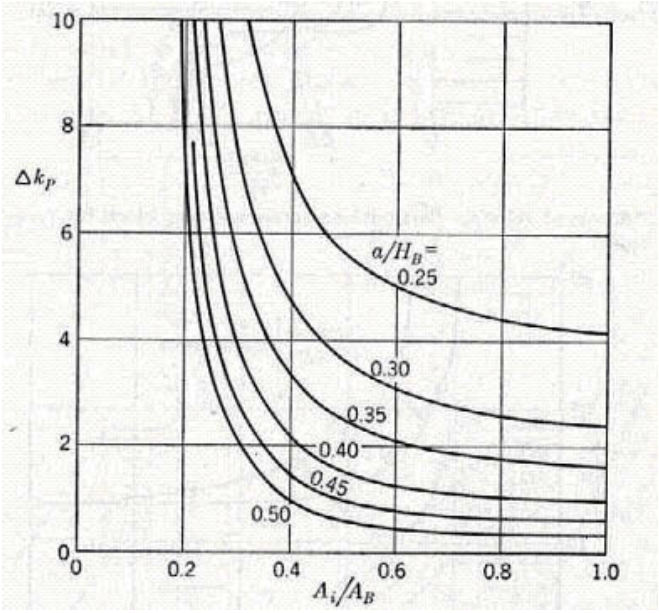


Fig. 22 Radiator Drag increase due to nozzle [16]

3.5 The Meredith's theoretical thrust force

The estimated air exit speed is about 408 km/h out of the principal duct. The result is a theoretical thrust force of about 35 N. Finally, it is possible to calculate the thrust force vs airspeed: under 160 km/h there is not additional thrust.

4 A simplified approach

Since at the very low cruise speed considered the contribution of the Meredith effect is not high, it is better to simplify the design in order to reduce weight and size.

It is important to asses that size is a key factor for low drag and, as it will be discussed afterwards, for high propeller efficiency.

A very effective solution is depicted in fig. 23. With this approach it is possible to have a very reduced radiator drag without penalties in terms of weight and height (undercarriage). The aftercooler and the radiator can be taken from a car. The aftercooler can be assembled directly upon the water radiator. For example the radiators pack of the FIAT 1300 jtd can be effectively used.

The defect of this solution is that the "water" radiator heats up the aftercooler. This is not a problem since aftercooler inertia is very low and when the aircraft runs for take-off it is readily cooled down. However for STOL (Short Take Off Landing) aircrafts it is possible to keep the "water" radiator under the engine and to move the aftercooler on one side (see figure 24).

5 Push or pull arrangement

As it is widely known the fuselage interference highly affects the propulsion efficiency. Tsagi has introduced a spinning fuselage to improve propulsion efficiency [15]. Especially pusher configuration are critical, as it can be seen in the graph 1. In this field a very confusing concept was introduced by the NASA work [12], furthermore investigated in [13].

In this NASA paper a very bulky and streamlined fuselage and a very small propeller at very low speed were investigated.

This gives the disputable conclusion of Wikipedia for thruster propellers that quotes "Efficiency can be gained by mounting a propeller behind the fuselage, because it re-energizes the boundary layer developed on the body, and reduces the form drag by keeping the flow attached." This is true for submarine or ships with very low Reynolds and very small propellers (as it is correctly specified afterward in Wikipedia), but normally false for aircrafts, where proportions and/or Reynolds number are different. Uneven pressure fields in front a propeller reduce propeller efficiency.

Tractor solutions are easier provided that the fuselage is kept as slim as possible (P-51D) or a nacelle stands behind the propeller (De Havilland Mosquito).

Push pull arrangement are an exception since the frontal propeller improves the rear propeller efficiency (SIAI Marchetti S55) by reducing suction problems.

6 Conclusion

The Meredith effect was investigated on a diesel installation of a low speed aircraft, in this case the heavy and cumbersome arrangement necessary to achieve a high static thrusts from

radiators is not justified. A simplified approach should then be adopted. A very effective solution is briefly discussed in this paper for the specific installation.

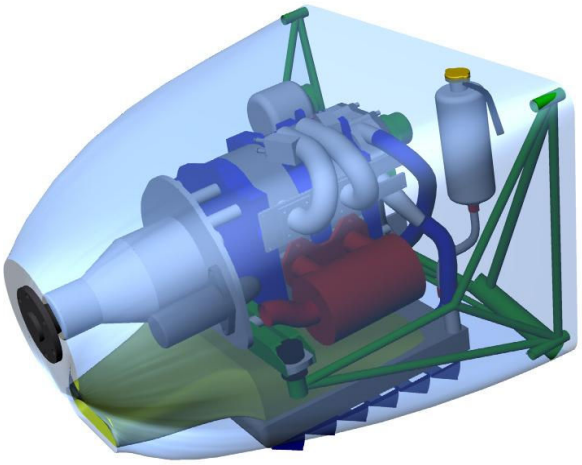
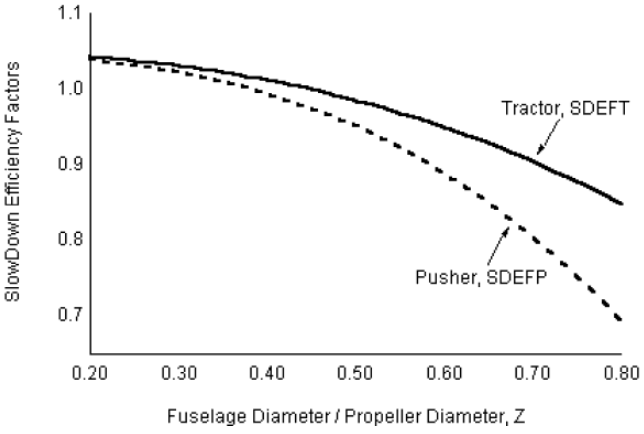


Fig. 23 – Very simplified and effective radiator installation (www.rotaryengine.net)



Fig. 24 – the aftercooler is installed on the side (www.rotaryengine.net)



Graph 1. Fuselage efficiency to be multiplied to the ideal thrust of the propeller [14]

References

[1] Kuchemann D., Weber J., Aerodynamics of propulsion, Mc Graw-Hill, 1953

[2] Kays W.M., London A.L., Compact Heat Exchanger, Mc Graw-Hill, 1955

[3] Hoerner S.F., Fluid Dynamic Drag, Hoerner, 1965

[4] Gothert B., "The Drag of Airplane Radiators with special References to Air Heating", NACA report n°896, 1939

[5] Winter H., "Contribution to the Theory of the Heated Duct Radiator", NACA report n°893, 1939

[6] Meredith F.W., "Note on the Cooling of Aircraft Engines with special references to Ethylene Glycol Radiators Enclosed in Ducts", British A.R.C. memorandum n°1683, 1935

[7] Weise A., "The Conversion of Energy in a Radiator", NACA report n°869, 1938

[8] Nelson W.J., Czarnecki K.R., Harrington R.D., "Full-Scale Wind-Tunnel Investigation of Forward Underslung Cooling-Air Ducts", NACA report n°L115, 1944

[9] Czarnecki K.R., Nelson W.J., "Wind-Tunnel Investigation of Rear Underslung Fuselage Ducts", NACA report n°L438, 1943

[10] Ames Research Staff, Equations, "Tables and Charts for Compressible Flow", NACA report n°1135, 1953

[11] Johnston, J Ford; "Review of flight tests of NACA C and D cowlings on the XP-42 airplane"; naca-report-771; 1943

[12] NASA-TN-1026

[13] AI AA-26-2693, "Aerodynamic design of Low Speed Aircraft with a NASA fuselage Wake/propeller Configuration", October 20-22 1986/Dayton Ohio

[14] Lowry, John T., "The Bootstrap Approach to Aircraft Performance", Part Two, Constant-Speed Propeller Airplanes, December 12 1999

[15] Tsagi. "Spinning the fuselage to save fuel",www.flightglobal.com/.../russian-tsagi-spinning-the-fuselage-to-save-fuel

[16] Piancastelli,L. , Pellegrini, M. "The bonus of aircraft piston engines, and update on the Meredith effect", Int. Jour. Of Heat and Technology, Vol. 25, N.2,2007.

Symbols

a distance between A_3 and A_2 [m]
c radiator tube length [m]
C_{Df}, Abbot and Van Doenhoff drag coefficient
C_D radiator drag coefficient
C_{pa} air specific heat capacity [J/kg K]
H_B radiator height [m]
n tubes number in depth radiator direction
p air pressure [Pa]
Δp pressure loss [Pa]
Q heat removal rate [W/s]
Re Reynolds number
S heat exchanger total transfer area [m^2]
t pitch of radiator tubes [m]
T air temperature [K]
U overall thermal conductance [W/ m^2K]
V air speed [m/s]
γ counter flow coefficient
ΔT air temperature rise within the radiator [K]
ΔT_M heat exch. mean log. temperature [K]
ξ concentrated loss coefficient
ρ air density [kg/m^3]

0 intake section of the duct (diffuser inlet)
1 frontal radiator section (diffuser outlet)
2 exit radiator section (nozzle inlet)
3 exit section of the duct (nozzle exit)

CEAS 2011

Collision Avoidance for Autonomous Vehicles

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Flight Demonstration of Radar-based Autonomous Non-cooperative UAS Collision Avoidance

Giancarmine Fasano, Domenico Accardo, Lidia Forlenza and Antonio Moccia
University of Naples "Federico II", Naples, Italy

Salvatore Luongo and Vittorio Di Vito
Italian Aerospace Research Center (CIRA), Capua (CE), Italy, I81043

Keywords: *unmanned aerial systems, autonomous collision avoidance, non cooperative flying obstacles, airborne radar tracking*

Abstract

In the framework of a research project carried out by the Italian Aerospace Research Center and the Department of Aerospace Engineering of the University of Naples "Federico II", an integrated radar/electro-optical system configuration was adopted to demonstrate in flight autonomous non-cooperative UAS collision avoidance. Real time data fusion and decision making algorithms were developed to estimate intruders motion and generate proper escape trajectories. The prototypical hardware/software system was installed onboard an optionally piloted flying laboratory of Very Light Aircraft category, and an extensive flight test campaign with a single intruder aircraft was carried out. This paper focuses on the results from radar-based autonomous collision avoidance flight tests. System performance is analyzed in terms of accuracy and reliability in estimating intruder position and velocity, and effectiveness in performing safe avoidance maneuvers while minimizing the deviation from nominal trajectory.

1 Introduction

Following the most important guidelines about Unmanned Aerial Systems (UAS) integration into civil airspace, the onboard avionics of these aircraft has to include a Sense and Avoid (S&A) system that is capable of replacing the human pilot in performing visual collision threat detection and avoidance [1-3]. Therefore, several experiences have been carried out worldwide in order to develop a system that can perform the above mentioned function.

In all these projects the configuration is comprised of two subsystems:

1. the Obstacle Detection and Tracking subsystem that permits to reveal flying intruders in a selected Field of Regard and to estimate their motion;
2. the Collision Avoidance subsystem that provides Conflict Detection and Resolution capabilities.

Regarding the "Sense" function, several solutions have been proposed depending on available budget and onboard resources, ranging from standalone Electro-Optical (EO) sensors [4-6] to integrated architectures which comprise airborne radars and collaborative systems based

on broadcasted information, such as TCAS and ADS-B [7-11].

A significant research activity has also been carried out about autonomous decision-making strategies for safe collision avoidance. The literature proposed approaches cover a variety of methods, including the application of the game theory [12], programming techniques [13-14], soft computing techniques (such as neural networks and genetic algorithms) [14-16], force field approach [17] and analytical geometric approaches [18-19].

Within the research project named TECVOL, the Italian Aerospace Research Center (CIRA) and the University of Naples "Federico II" developed a prototypical S&A System which relies on an integrated radar/electro-optical (EO) configuration. In particular, the sensing system is comprised of a pulsed radar, four EO sensors and two processing units for image processing and real time multi-sensor tracking. A hierarchical data fusion model was selected where the radar is the main sensor that must perform initial detection and tracking and the EO sensors are used as auxiliary information sources to increase tracking accuracy and measurement rate [20]. Then, a decision making logic was developed and integrated in the automatic flight control system to handle collision conditions in real time and perform adequate evasive maneuvers. The prototypical hardware/software system was installed for flight demonstration and performance assessment onboard a customized optionally piloted Very Light Aircraft (VLA) named FLARE (Flying Laboratory for Aeronautical REsearch). FLARE is a modified version of the TECNAM P92 plane and represents a cost-effective platform to test innovative flight technologies. Several flight tests with a single intruder aircraft in the same category of FLARE were carried out in different environmental conditions first to gather sensor data in realistic near collision conditions, then to estimate real time multi-sensor tracking performance. Indeed, given the relatively low speed and the high maneuverability of this type of platforms, autonomous collision avoidance can be tested in full safety conditions even in radar-only configuration. Thus, it was decided to carry out the first series of collision avoidance tests in this configuration.

While avoidance tests based on real time radar/EO fusion are currently in progress, this paper describes some results of the radar-based autonomous collision avoidance experiments and is organized as follows. First, the functional and hardware architecture of the developed collision avoidance system is briefly presented, including algorithms for real time obstacle tracking, collision detection, and escape trajectory generation. The architecture for flight tests execution is then summarized. Finally, some results from autonomous collision avoidance tests are presented. It is demonstrated how is the system able to react to the collision threat while preserving high reliability both in terms of situational awareness and decision making.

2 System architecture

As anticipated above and shown in Fig. 1, from a functional point of view the autonomous collision avoidance (ACA) module is based on two core algorithms. Firstly, a customized multi-sensor tracking software ensures that intruders' dynamics is properly estimated. Secondly, a decision making logic is able to handle collision conditions by commanding the autopilot to perform adequate evasive maneuvers whenever a potential conflict is detected.

The hardware architecture of the sensing sub-system comprises an airborne pulsed Ka-band radar, a panchromatic camera in visible band, a color camera in visible band (aimed at obstacle identification, which will be implemented in the future), two Infra Red (IR) cameras (aimed at obstacle detection in low light conditions), a CPU devoted to Image Processing, and a CPU devoted to Real Time Tracking by sensor data fusion. This unit is connected via CAN bus to the onboard Flight Control Computer (FCC) that performs autonomous navigation and flight control, including autonomous collision avoidance maneuvers. The system exploits data from onboard navigation sensors, such as

Attitude and Heading Reference System, Laser Altimeter, Standalone GPS, and Air Data System. Moreover, the FCC derives the proper commands for electric servo-actuators that handle the control surfaces. The described architecture is depicted in Fig. 2.

The selected radar is a customized version of the AI-130™ OASys™ (Obstacle Awareness System) model produced by former Amphitech™. It is a mechanically scanned radar operating with a carrier at 35 GHz and it has been used in several UAS sense and avoid research programs [7-8,21]. Selected carrier frequency provides a good compromise between antenna dimensions, angular accuracy and sensitivity to rain and fog. The two visible cameras are installed parallel to the aircraft longitudinal axis and work with 1280 x 960 resolution, while the IR sensors have a resolution of 320 x 240 pixels and are set slightly eccentric in order to cover the same field of view (FOV) of the visible cameras. The sensor unit is placed on the top of the aircraft wing with the radar in central position (Fig. 3). The selected configuration demonstrates the S&A capability in a reduced field of regard (120° in azimuth).

3 Algorithms

3.1 Obstacle detection and tracking

The developed tracking algorithm can work in two operating modes, such as:

1. Standalone radar tracking mode that performs current estimates of obstacle relative position and velocity on the basis of measurements of radar and inertial unit;
2. Radar/EO tracking mode where current estimates of obstacle relative position and velocity are determined also on the basis of EO sensors measurements.

The output data rate is 10 Hz for both modes, but the radar/EO mode has the potential of overcoming the radar-only performance in terms of accuracy on angles and their derivatives.

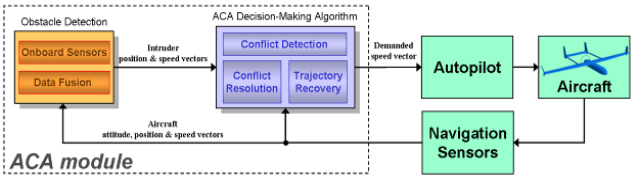


Fig. 1 ACA system functional architecture within the closed-loop control system

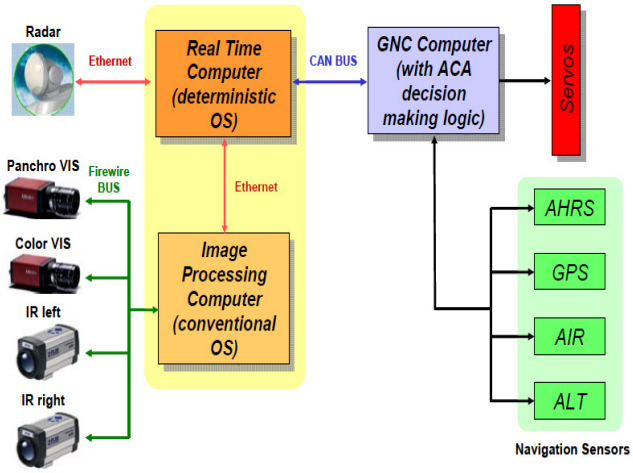


Fig. 2 S&A/GNC hardware architecture.



Fig. 3 S&A System installation onboard FLARE

While also presenting some general aspects of the developed algorithm, the following discussion will be more focused on the radar-only configuration. The reader is referred to references [11,20] for a detailed analysis of issues and solutions relevant to effective real time radar/EO fusion. The real time tracking module is needed to associate measurements of the same intruder gathered in subsequent scans, to eliminate false alarms and clutter returns, to increase output rate with respect to radar raw data (about 1 Hz), and above all to derive obstacle kinematics. In fact, this latter

information, which is not directly provided by the sensor, is of key importance for estimating whether or not it is necessary to perform an evasive maneuver [11].

The algorithm is based on an Extended Kalman Filter (EKF) with linear dynamic model and non linear measurement equations. In particular the state vector is comprised of nine components which are the target obstacle coordinates in NED (North-East-Down reference frame with origin in the aircraft center of mass) with their first and second time derivatives. The adopted dynamic model assumes that every component of target acceleration evolves in terms of a correlated noise process with given time constant and instantaneous variance [22].

As regards the measurement equation, the measurement vector has three components in case of radar detections, while since the filter works in NED the measurement covariance matrix depends not only on radar/EO specifications and performance but also on AHRS inaccuracies in estimation of aircraft attitude angles.

Navigation data are used by the algorithm at the frequency of 10 Hz so that own vehicle dynamics is properly tracked. In particular, GPS data are used in track initialization phase while acceleration and attitude measurements are used in all tracking phases. Angular velocity estimates are used to correct lever arm effects on acceleration measurements.

Important issues relevant to real time implementation are sensor latency and ground clutter filtering.

As for the first point, the radar sends its measurements at the end of each pass with a (known) delay that can be at most of the order of 1 s. To perform gating, association and track updating, the solution is to keep in memory a sliding window where navigation and track data are stored. Measurement filtering process consists in evaluating filtered state and covariance at measurement time and then modifying current state and covariance on the basis of this information.

Regarding ground clutter removal, in lack of Doppler information, the problem is faced by using single scan criteria (range-dependant altitude thresholds) before applying tentative tracking.

3.2 Autonomous Collision Avoidance

The ACA algorithm on-line accomplishes two tasks: the conflict detection and the avoidance maneuver evaluation.

The first task concerns the identification of a situation that can lead, in the very near future, to a possible conflict, meaning a reduction of the separation distance between the own aircraft and the considered intruder below a specified value (minimum separation threshold). In order to define the conflict situation, a proper volume (defined as “safety bubble”) is set around the intruder aircraft. The shape and dimensions of this safety bubble have to be set in such a way to ensure that, if the own aircraft does not breach the bubble, the minimum required separation distance between the vehicles is maintained. This means, on the other hand, that a conflict situation arises when the future projected trajectories of the intruder and of the own vehicle are expected to lead to a breach of the safety bubble.

The ACA algorithm implements a spherical safety bubble whose dimensions are:

$$R = 2R_0 = 1000\ ft \approx 304.8\ m,$$

where R is the radius of the sphere centered in the intruder aircraft, while R_0 is the FAA minimum required safety distance [7].

Conflict detection is performed by modeling the own aircraft (A/C_A) as a point-of-mass object A, with three degrees of freedom and having velocity $\vec{V}_A$. The intruder, with its surrounding safety bubble, is modeled as a spherical object centered in the point B (i.e. the point-of-mass describing the intruder), with radius R and having velocity $\vec{V}_B$. A NED reference frame with origin in the point B is used. The graphical representation of the detection problem geometry is shown in Fig. 4.

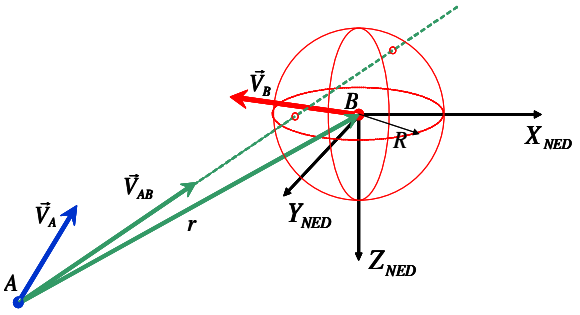


Fig. 4 Detection problem geometry.

In order to use an equivalent but easier to manage geometry, the relative velocity vector $\vec{V}_{AB} = \vec{V}_A - \vec{V}_B$ is considered, so that the above described geometry is translated in the equivalent situation where the sphere is stationary and the point-of-mass object A moves, with relative velocity $\vec{V}_{AB}$. With reference to the adopted geometrical schematization, the necessary and sufficient possible collision conditions (that in general terms can be defined as "conflict" conditions) are conceptually the following:

- the direction of the $\vec{V}_{AB}$ vector intersects the sphere,
- and the two aircraft are approaching.

This concept has been mathematically implemented in the algorithm, so that the ACA algorithm is able to detect on line the possible conflict situations, based on the current estimated relative velocity vector between own and intruder aircraft.

For what concerns the second task of the ACA algorithm, i.e. the evaluation of the avoidance maneuver, the strategy is based on the analytical solution of the following kinematic optimization problem: to find a new A/C_A velocity vector, in compliance with the vehicle envelope limitations, able to avoid a collision with the safety bubble surrounding the intruder aircraft A/C_B, while minimizing the change of the A/C_A new velocity vector with respect to the previous one.

This leads to the formulation of a nonlinear constrained optimization problem. In this problem, the main imposed constraint guarantees that the desired relative velocity vector direction does not intersect the sphere (collision avoidance constraint) and constraints are also imposed on the speed module and flight path angle envelope limitations of the A/C_A.

Concerning the methodology used in the ACA algorithm in order to solve the posed optimization problem, the first step consists in addressing a simplified problem where no dynamic and envelope constraints act on the A/C_A vehicle. Furthermore, in the search of the optimal solution only the trajectories tangent to the safety bubble are considered. In this way, the closest point of approach (CPA) of the possible avoidance trajectories is placed on the boundaries of the safety region, so the deviation of the new vehicle trajectory from the original flight plan is minimized. Based on these assumptions, the problem is reduced to a kinematic unconstrained (as far as A/C_A dynamic and envelope limitations are concerned) optimization problem.

Once solved the simplified optimization problem, the constraints are considered. Regarding the desired track angle of the own aircraft, no envelope limitations apply. The limitations in terms of desired inertial speed module V^d and desired flight path angle γ^d of own vehicle A/C_A, on the other hand, are accounted for by eliminating from the simplified problem solution process the candidate solutions where V^d and/or γ^d are outside the envelope limitations of the vehicle A/C_A. This approach is possible because, among all the possible A/C_A desired inertial velocity vectors, at least one always exists which is compliant with the envelope limitations, consisting in a rotation of the velocity vector in the horizontal plane.

In conclusion, the above described ACA algorithm above described is able to find on-line, at each evaluation step (it works in real time at 100 Hz), a maneuver able to skim the safety bubble.

For more details about the ACA algorithm, the reader is referred to Ref. 19.

4 Flight test architecture

The collision avoidance flight test campaign was carried out by exploiting the following configuration of test facilities (besides FLARE aircraft):

- A piloted intruder aircraft in the same class of FLARE equipped with GPS;
- A Ground Control Station (GCS) for real-time flight coordination and test monitoring;
- A full-duplex data-link between FLARE and GCS;
- A downlink between intruder and GCS.

As regards the GCS, it is comprised of several workstations which are connected on an Ethernet bus and they communicate by means of the UDP protocol. In particular, two workstations are dedicated to monitoring the obstacle sensing and the collision avoidance systems. These workstations send commands and receive data in real time from a computer which acts as the ground communication controller. The latter is devoted to data excange with FLARE. In fact, it communicates with an on-board communication controller by means of a full duplex Radio Frequency (RF) data link. This latter computer exchanges data with the flight control computer and the real time tracking computer by means of the CAN bus.

Correct execution of obstacle detection and collision avoidance flight tests requires that the intruder is properly synchronized with FLARE in order to realize the desired approaching geometries. Thus, the intruder is equipped, besides the GPS receiver, with a RF transmitter to download position data to the ground station, and a processing unit to store flight test data. In this case the obstacle sensing workstation is directly connected with the intruder, as shown in Fig. 5, and broadcasts intruder telemetry on the GCS Ethernet bus. Intruder GPS data (stored onboard and in the GCS computers) are also used in post processing phase, together with FLARE GPS data, to conduct error analyses.

The software tools installed in the ground workstations allow the engineers to monitor test status by means of information reported in graphical displays and numerical output, and to send commands to the flight system, including single sensors such as the radar. The graphical interfaces of the two workstations are depicted in Fig. 6 and Fig. 7.

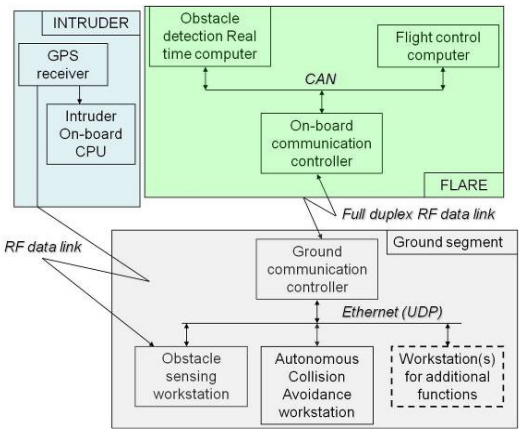


Fig. 5. Physical architecture relevant to the execution of collision avoidance experiments

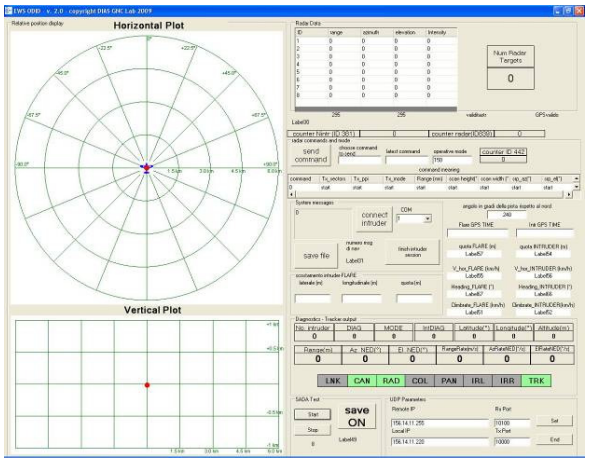


Fig. 6. Graphical interface for obstacle sensing workstation

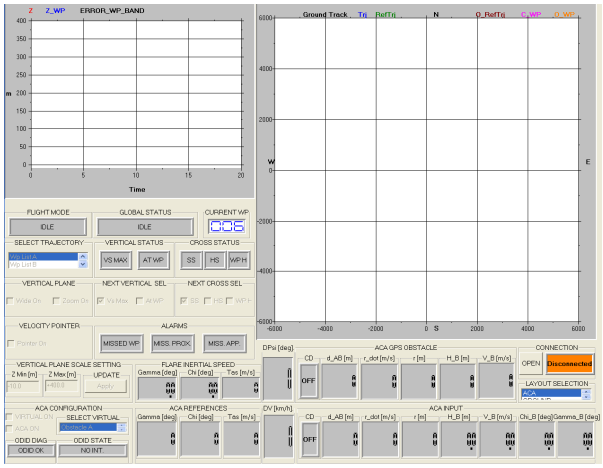


Fig. 7. Graphical interface for ACA workstation

5 Flight test results

As already stated above, an extensive flight test campaign was carried out first to gather sensor data for off-line analysis, then to estimate real time obstacle tracking performance [20,23], and finally to demonstrate autonomous collision avoidance capabilities.

Several encounters were tested in different geometries and environmental conditions, and the autonomous avoidance system performed successful and reliable avoidance maneuvers in all the tests.

For the sake of brevity, a single frontal encounter is described in detail in what follows. However, general considerations can be derived both from the points of view of tracking and avoidance maneuver execution.

In the considered encounter, the two aircraft are initially on a quasi frontal collision course. FLARE is in autonomous flight mode while the intruder aircraft is manually piloted with the aid of the information provided by the flight test engineer in the GCS. For safety reasons, an altitude separation of about 70 m is implemented, with the intruder flying at higher altitude than FLARE.

On the basis of the outcome of previous flight tests, the radar is set to work with a maximum range of 2 nm. Given the low approaching speed, this distance still allows a very large time to collision. The radar field of view is set to be 90° (azimuth) times 20° (elevation).

In the considered scenario, the intruder is detected by the radar for the first time at a range of about 2290 m (Fig. 8). After three consistent detections, firm tracking is entered at a range of about 2090 m. This distance is sometimes called "declaration range" in literature [21]. As soon as the track is declared, relevant kinematic information is passed to the collision detection logic which estimates distance at closest point of approach and time to closest point of approach (d_{CPA} and t_{CPA} [11]). Collision is declared when the range rate is negative and the d_{CPA} estimated on the basis of tracker outputs is smaller than a threshold. In order to keep into account the unavoidable uncertainties in estimation of intruder dynamics, the threshold is set equal to the bubble distance indicated by aeronautical regulations (500 ft) plus a safety margin of 500 ft, as mentioned earlier. In the considered case, because of the small value of the estimated d_{CPA}, the avoidance maneuver is initiated soon after intruder declaration. The maneuver is mostly in the horizontal plane and consists in a (very smooth) right turn.

Because of the aircraft attitude dynamics during the avoidance maneuver, a huge amount of ground clutter echoes is detected, while a few detections are useful for track update, as it can be also seen in Fig. 9 and Fig. 10, which report intruder azimuth and elevation in NED. Nevertheless, the tracking algorithm is able to maintain a satisfying accuracy, even when the intruder falls out of the radar FOV (Fig. 11) and no more radar measurements are available.

A satisfying tracking performance is also obtained considering estimates of intruder relative velocity, as it is demonstrated for example in Fig. 12 which refers to range rate.

Maneuver effectiveness is shown in Fig. 13 which clearly shows how is the autonomous system able to increase the estimated d_{CPA} thus avoiding the collision threat.

A global view in the horizontal plane of the whole performed maneuver is shown in Fig. 14, where the FAA safety bubble (the internal circle), the safety bubble increased by the safety margin (the external circle), FLARE and intruder trajectories as measured by the GPS ("true" trajectories) and the intruder trajectory as estimated by the Obstacle Detection and Tracking algorithm (measured intruder trajectory) are also indicated. The distance at closest point of approach and the maximum deviation from the nominal trajectory are also reported in the figure. Furthermore, Table 1 reports synthetic performance indexes relevant to the considered collision avoidance maneuver.

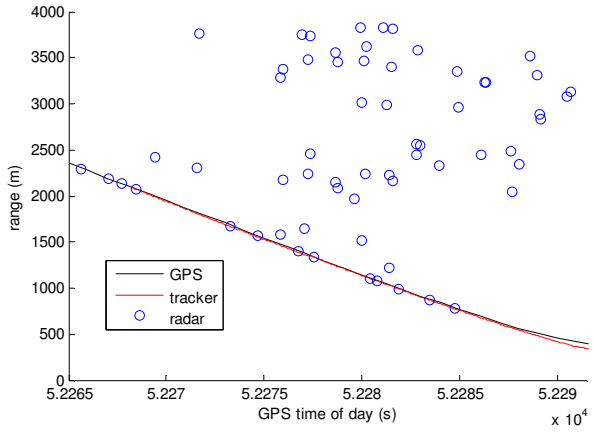


Fig. 8 Intruder Range in the considered scenario as estimated by GPS, tracker, and radar, as a function of GPS time of day

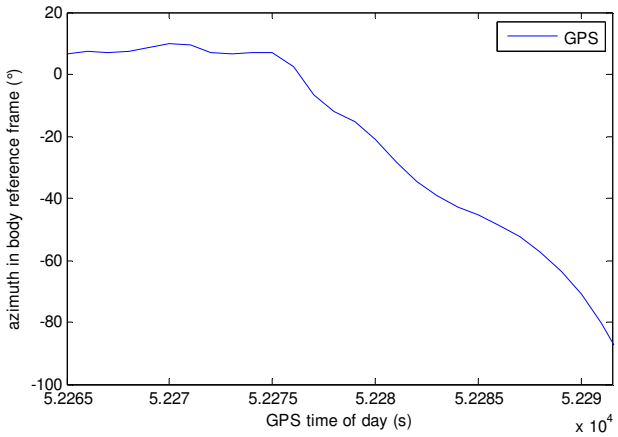


Fig. 11 Intruder azimuth in body reference frame in the considered scenario as estimated by GPS, as a function of GPS time of day

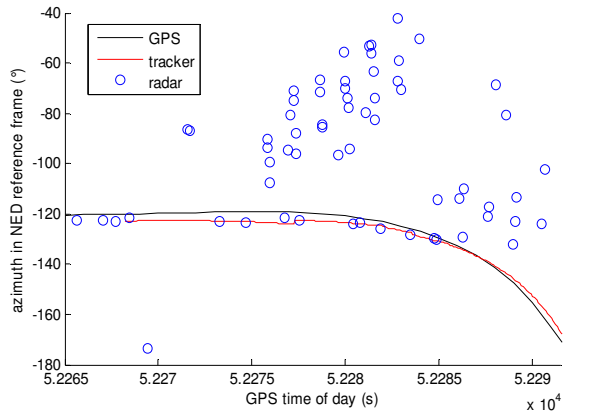


Fig. 9 Intruder Azimuth in NED in the considered scenario as estimated by GPS, tracker, and radar, as a function of GPS time of day

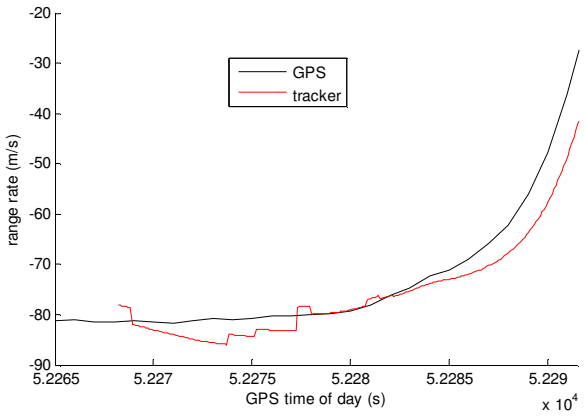


Fig. 12 Intruder Range rate in the considered scenario as estimated by GPS, and by the tracker, as a function of GPS time of day

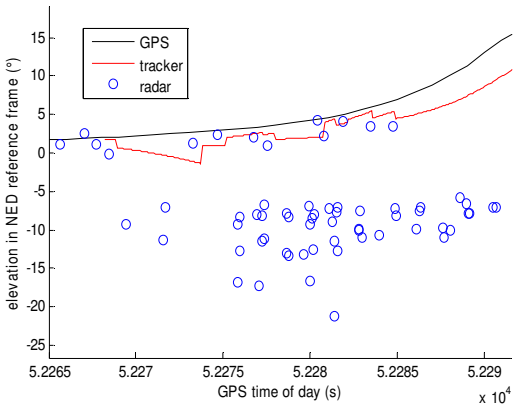


Fig. 10 Intruder Elevation in NED in the considered scenario as estimated by GPS, tracker, radar, and EO system, and estimation errors w.r.t. GPS, as a function of GPS time of day

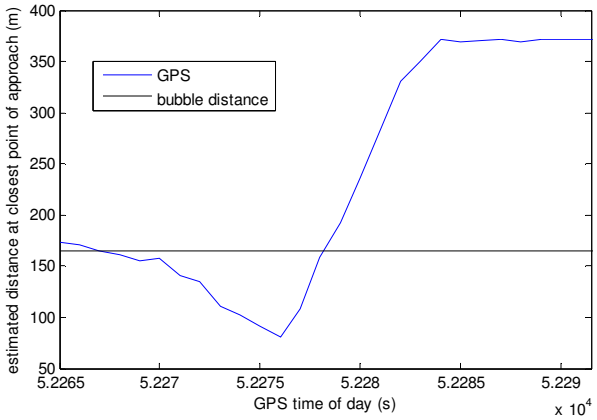


Fig. 13 Estimated distance at closest point of approach estimated by GPS in the considered collision avoidance scenario, as a function of GPS time of day

Table 1 Synthetic indexes relevant to the considered collision avoidance maneuver

Declaration Range [m]	2090
Maximum 4D deviation from nominal trajectory [m]	519.5
Rotation of the speed vector [°]	44.1
Maximum angle of attack [°]	5.6
Maximum vertical load factor [g]	1.20
Maximum angular velocity [°/s]	32.6

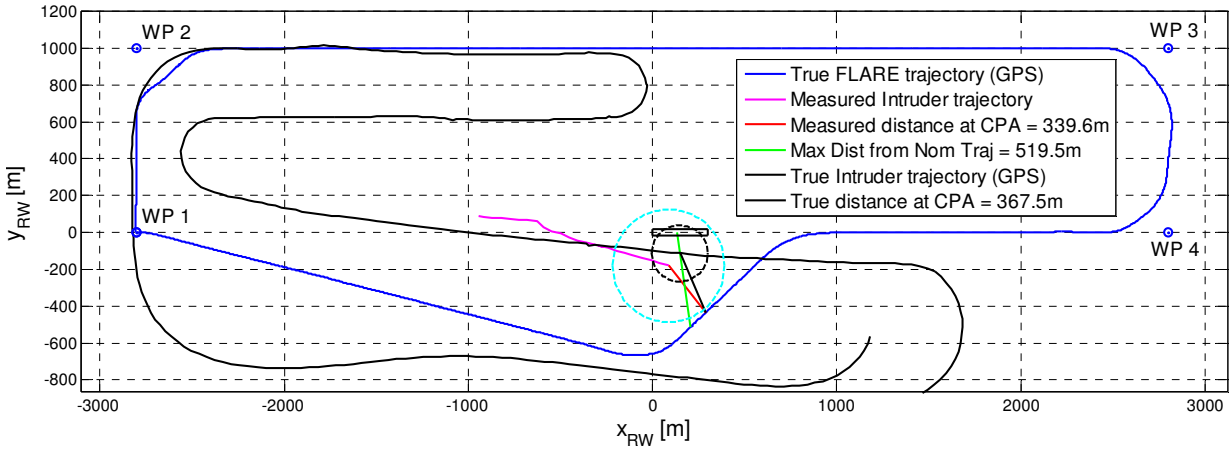


Fig. 14 Planar view of the whole collision avoidance maneuver performed in the examined flight test

All the reported results demonstrate the effectiveness of the autonomous collision avoidance system, which combines accurate and reliable airborne tracking with an adaptive maneuvering logic, thus also avoiding unnecessarily aggressive maneuvers.

6 Conclusion

This paper focused on results from flight testing of a fully autonomous non-cooperative collision avoidance system for UAS which has been developed by the Italian Aerospace Research Center and the University of Naples "Federico II". In particular, radar-based collision avoidance was discussed in detail.

In all the performed tests, the system demonstrated the capability to combine fast responsiveness and high reliability: real time intruder tracking performance allowed the autonomous flight control system to generate proper escape trajectories and execute safe and smooth avoidance maneuvers. Future works will focus on results from multi-sensor based collision avoidance flight tests.

Acknowledgements

The PhD grant of Ms. Lidia Forlenza was funded by CIRA.

References

[1] EUROCONTROL, "EUROCONTROL Specifications for The Use of Military Unmanned Aerial Vehicles as Operational Air Traffic Outside Segregated Airspace", version 1.0, doc no. EUROCONTROL-SPEC-0102, Bruxelles, Belgium, July 27th, 2007.

[2] US Federal Aviation Administration (FAA), March 27th, 2008, "Airworthiness Certification of Unmanned Aircraft Systems", Order 8130.34, Washington DC, USA..

[3] US Department of Defence (DoD), "Unmanned Aircraft Systems Roadmap 2005-2030",

Washington DC, USA, August 4th, 2005.
[4] Utt J., McCalmont J., Deschenes M., "Development of a Sense and Avoid System", *1st AIAA Infotech@Aerospace Conference*, Arlington, Virginia, Sep. 26-29, 2005, pp- 1-10.
[5] Lai O., Mejias L., Ford J.J., "Airborne vision-based collision-detection system", *Journal of Field Robotics*, Aug 2010, DOI: 10.1002/rob.20359
[6] Dey D., Geyer C., Singh S., Digioia M., "Passive, Long-Range Detection of Aircraft: Towards a Field Deployable Sense and Avoid System", *Field and Service Robotics*, Vol. 62 (2010), pp. 113-123. doi:10.1007/978-3-642-13408-1_11
[7] MTSI, "Non-Cooperative Detect, See, & Avoid (DSA) Sensor Study", July 2002, NASA ERAST technical report.
[8] Korn B., Edinger C., "UAS in Civil Airspace: Demonstrating "sense and avoid" capabilities in flight trials", *27th Digital Avionics System Conference*, St. Paul, USA, 26-30 October 2008.
[9] Shakernia O., Chen W. Z, Graham S., Zvanya J., White A., Weingarten N., Raska V. M., May 2007, "Sense and Avoid (DS&A) Flight Test and Lessons Learned", *2nd AIAA Infotech@Aerospace Conference and Exhibit*, Paper AIAA 2007-3003, Rohnert Park, CA, pp. 1-12.
[10] Portilla E., Fung A., Chen W.Z., Shakernia O., and Molnar T., May 2007, "Sense And Avoid (SAA) & Traffic Alert and Collision Avoidance System (TCAS) Integration for Unmanned Aerial Systems (UAS)", *2nd AIAA Infotech@Aerospace Conference and Exhibit*, Paper AIAA 2007-3004, Rohnert Park, CA.
[11] Fasano G., Accardo D., Moccia A., Carbone C., Ciniglio U., Corraro F., and Luongo S., 2008, "Multi-Sensor-Based Fully Autonomous Non-Cooperative Collision Avoidance System for Unmanned Air Vehicles," *AIAA Journal of Aerospace Computing, Information, and Communication*, vol. 5, pp. 338-360, October 2008, DOI: 10.2514/1.35145
[12] Tomlin C., Pappas G. J. and Sastry S., 1998, "Conflict Resolution for Air Traffic Management: A Study in Multiagent Hybrid Systems", *IEEE Transactions on Automatic Control*, Vol. 43, pp. 509–521, April 1998.
[13] Pallottino L., Feron E. M. and Bicchi A., 2002, "Conflict Resolution Problems for Air Traffic Management Systems Solved with Mixed Integer Programming", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 3, No. 1, March 2002.
[14] Christodoulou M. A. and Kontogeorgou C., 2008, "A novel algorithm for collision avoidance in commercial aircraft using neural networks and non-linear programming", *IEEE 16th Mediterranean Conference on Control and Automation*, Ajaccio, France, June 2008.
[15] Alliot J.-M., Gruber H., Joly G. and Schoenauer M., 1993, "Genetic Algorithms for Solving Air Traffic Control Conflicts", *Proceedings of the Ninth Conference on Artificial Intelligence for Applications*, 1-5 March 1993, Pages: 338–344.
[16] Frazzoli E., Mao Z.-H., Oh J.-H. and Feron E., 2001, "Resolution of Conflicts Involving Many Aircraft via Semidefinite Programming", *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 1, January-February 2001.
[17] Ghosh R. and Tomlin C., 2000, "Maneuver design for multiple aircraft conflict resolution", *Proceedings of the 2000 American Control Conference*, Vol. 1, Issue: 6, 28-30 June 2000, Pages: 672–676.
[18] Bilimoria K.-D., 2000, "A Geometric Optimization Approach to Aircraft Conflict Resolution", *AIAA Guidance, Navigation, and Control Conference*, August 2000, Denver, Colorado.
[19] Luongo S., Corraro F., Ciniglio U. and Carbone C., 2009, "An Optimal 3D Analytical Solution for Collision Avoidance between Aircraft", *IEEE Aerospace Conference 2009*, BigSky, Montana, March 2009.
[20] Fasano G., Forlenza L., Accardo D., Moccia A., Rispoli A., " Data fusion for UAS collision avoidance: results from flight testing ", *AIAA Infotech@Aerospace Technical Conference*, St. Louis, Missouri, March 2011.
[21] Chen R.H., Gevorkian A., Fung A., Chen W.Z., Raska V., "Multi-Sensor Data Integration for Autonomous Sense and Avoid", *AIAA Infotech@Aerospace Technical Conference*, St. Louis, Missouri, March 2011.
[22] Singer R.A., Sea R.G. and Housewright K.B., "Derivation and Evaluation of Improved Tracking Filter for Use in Dense Multitarget Environments," *IEEE Transactions. on Information Theory*, vol. 20, no. 4, July 1974, pp. 423-432.
[23] Fasano G., Forlenza L., Accardo D., Moccia A., Rispoli A., "Integrated Obstacle Detection System based on Radar and Optical Sensors", *AIAA Infotech@Aerospace Technical Conference*, Atlanta, Georgia, USA, April 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Strategies for Improving the Electro-Optical Object Detection SW of an Integrated Sense and Avoid System

Lidia Forlenza and Domenico Accardo
University of Naples Federico II, Department of Aerospace Engineering (DIAS), Naples, Italy

Keywords: *Sense and Avoid, Electro-optical systems, Image processing, Object detection, Auxiliary sensor*

Abstract

An aircraft is provided with "sense and avoid" capabilities is able to detect the presence of other aircrafts on a collision route in its surroundings and to perform course corrections or evasive maneuvers to prevent possible collisions. In this framework, the Italian Aerospace Research Center (CIRA) developed a project that aims at realizing a prototypal Sense and Avoid system characterized by a multisensor configuration: pulsed radar and four Electro-Optical cameras. This paper is focused on the description of the image processing algorithm for the panchromatic camera. Its main original feature is that it is optimized in order to detect objects position with higher accuracy and data rate than the radar, i.e. the main sensor in the multi sensor architecture. The algorithm has been tested on images captured during flight tests so that strategies for controlling image critical issues would have been derived from real conditions. The paper ends with the description of the assessed object detection software performance, that demonstrates that the algorithm is fully satisfactory for the integrated sense and avoid system.

1 Introduction

One of the major hurdles in realizing the potential civilian applications of Unmanned Airborne Vehicles (UAVs) is how to safely integrate them into the national airspace. The FAA requires a Detect, Sense and Avoid (DS&A) technology which assures an equivalent level of safety with respect to manned aircraft, and it is able to avoid conflict with other aircraft [1-4]. Generally, DS&A systems are comprised of two subsystems, such as the Obstacle Detection and Tracking System and the Collision Avoidance one. The first is dedicated to detect intruder aircraft position, whilst the second performs collision detection and engages the collision avoidance maneuver. Thus, the TECVOL project, carried out by the Italian Aerospace Research Center (CIRA) and the University of Naples "Federico II", aims at demonstrating a prototypal DS&A system onboard a Very Light Aircraft (VLA) experimental platform. The sensing unit is composed by four Electro Optical (EO) cameras and Ka-band pulse radar. This choice resulted adequate to support systems with a reliable full autonomy from ground [5, 6]. In particular, the radar is the main sensor because it allows for all-time all-weather use and it provides a direct estimate of range-to-obstacle. Consequently, the EO cameras are adopted as auxiliary sensors to increase the angular resolution and the data rate.

This paper presents the original approach for the development of the EO segment. In particular it discusses the image processing algorithm for the panchromatic camera, which is dedicated to the obstacle detection [7]. Therefore, after a brief description of the flying experimental platform, the most important issues about the image processing algorithm are discussed such as the adopted technique for detecting obstacles and the obstacle tracking in the image. The central part of the paper is focused on evaluating the critical issues of the image processing algorithm applied on the real images acquired during flight tests (e.g. sun in front of the camera, horizon line presence in the image), and how the false alarm rate was reduced. Subsequently, a hardware-in-the-loop system is presented [8], as a support to the real flying platform, in order to test the feasibility and the computation time of the assessed panchromatic object detection software. Finally, algorithm performance results are discussed in terms of accuracy and data rate.

2 Experimental platform architecture

The Sense and Avoid technology developed by TECVOL project is based on the DS&A system installed onboard a VLA named FLARE (Flying Laboratory for Aeronautical REsearch). It comprises an airborne pulsed radar, two visible cameras, one panchromatic and one color, two Infra Red (IR) cameras, a CPU devoted to Image Processing (IP-CPU), and a CPU devoted to Real Time Tracking by sensor data fusion (RTT-CPU). This unit is connected to the onboard Flight Control Computer (FCC) that performs autonomous navigation and flight control, including autonomous collision avoidance maneuvers.

The radar is the main sensor, since it provides intruder range, azimuth and elevation information. The panchromatic camera is used as aiding sensor of the radar, in order to increase data rate and accuracy of detected object. The color camera is devoted to aircraft identification. Finally, IR cameras have the same function as the panchromatic camera when it is not able to work, such as in dark luminosity conditions. Obstacle identification and day/night operations will be developed in future phases of TECVOL project. The entire sensor package is on the wing, radar in central position and cameras parallel to it.

The system communication is implemented on the basis of figure 1. Radar performs the initial intruder detection which is transmitted to the RTT-CPU by Ethernet connection. Subsequently the RTT-CPU provides the intruder position estimated by the radar to the IP-CPU, which runs the panchromatic camera object detection software. The Camera communicates with IP-CPU via a IEEE1394b link. If the intruder aircraft is in the Field Of View (FOV) of the panchromatic camera, the image processing algorithm provides output measurements of intruder position, which are sent back to RTT-CPU. Finally, this unit performs fusion of radar and EO data and allows for tracking the intruder aircraft.

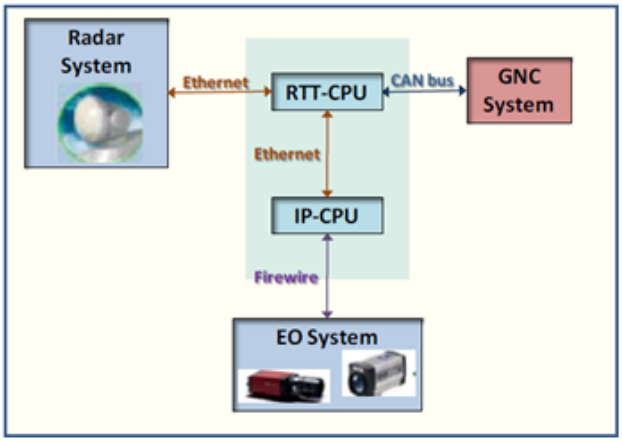


Figure 1 Onboard DS&A architecture

3 Development and critical issues of image processing algorithm

Image processing algorithm for panchromatic camera exploits radar measurements to perform the intruder optical detection and tracking, in order to increase the position accuracy and data rate. The adopted panchromatic camera is the model Marlin™ produced by Allied Vision Technologies™. Table 1 summarizes its performance in terms of FOV, resolution and data rate.

Table 1 Panchromatic Camera Performance

FOV	49.2° x 37.9°
RESOLUTION	1280 x 960 pixels
DATA RATE	7.5 fps

Each time the RT-CPU commands the IP-CPU to run the image processing algorithm, it starts searching in a window within the last acquired image to perform the obstacle detection. That window is centered on the azimuth and elevation measurements provided by radar, and its size depends on the current intruder range (fig. 2). Therefore, the image processing algorithm performs the intruder detection in that limited portion of the overall flight image, in order to increase the accuracy of intruder position, and to provide that information after a very short computation time, because of the reduced window size. Finally, the EO outputs are converted from the Camera Reference Frame (CRF) to the aircraft Body Reference Frame (BRF) by means of the procedure derived in ref. 15 and they are sent to the RTT-CPU.

3.1 Determination of the search window size

Search window definition has to take into account several aspects. First of all, window dimensions have to be set considering uncertainty sources such as radar error in NED and residual time/space registration errors. On the other hand, at relatively small range the search window has to be large enough to enclose even the largest possible obstacles. Thus, search window dimensions are defined as follows. When the estimated intruder range is larger than 350 m, the window has constant dimensions in pixels, i.e. 150 (width) x 100 (height) pixels, corresponding to an angular FOV of about 6° x 4°. Of course, linear dimensions of the search window increase for increasing range and are always large enough to enclose a typical civil aviation aircraft, such as a Boeing 737 or an Airbus A320. In order to set a lower bound for search window dimensions, when the range is smaller than 350 m the window dimensions in pixels are inversely proportional to range thus achieving constant linear dimensions. Indeed, this case is of little interest for sense and avoid applications.
Figure 2 depicts linear dimensions (in m) of the search window as a function of the estimated intruder range.

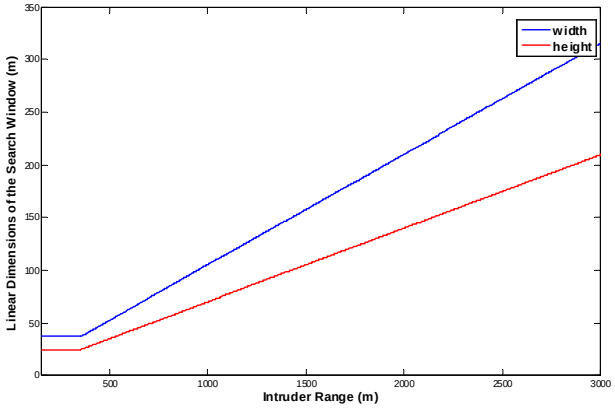


Figure 2 Search Window dimensions

3.2 Image processing technique

The object detection technique adopted in the image processing algorithm is the coupled edge detection - labeling, based on the Sobel method [9, 10]. It resulted the fastest and the most accurate technique, as already discussed in previous papers [7], and it has been customized for our applications. Moreover, among the several edge detection methods (Robert Cross, Prewitt, Canny and Sobel [11, 12]), the Sobel one is the most suitable for real-time operations with isotropic answers along the two image dimensions.
The implementation of the edge detection-labeling technique is carried out in different steps:

- By means of the Sobel edge detection technique, the search window is binarized. In order to limit the impact of background noise a relatively high threshold is considered in Sobel method, on the basis of the assumption that the intensity gradient generated by the intruder is larger compared with other background objects;
- The labeling technique connects all detected pixels in the binarized search window (if any) and it outputs a limited number of edges;

• Finally, the intruder is detected in the search window as the largest edge.

Hereinafter we show an example of edge detection-labeling implementation, applied to a real image acquired during a flight test. In particular, Figure 3 illustrates the image processing technique implemented within the search window, which encloses the intruder aircraft during a frontal encounter flight test. The instantaneous range between FLARE and the intruder aircraft is of about 1 km.

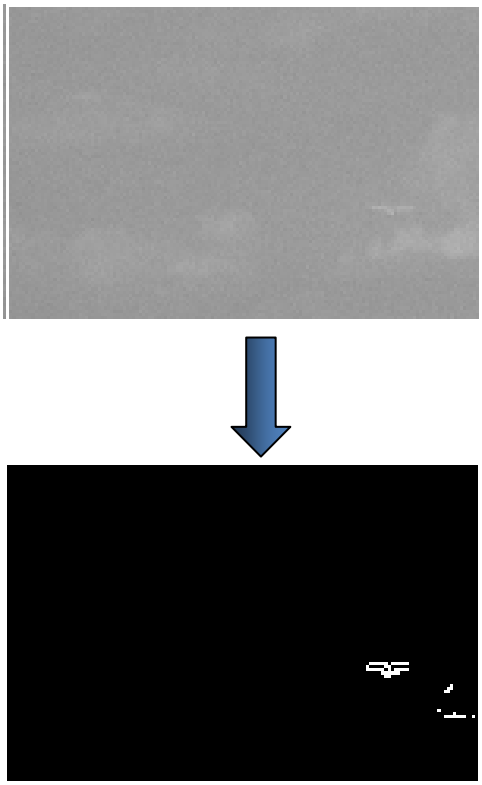


Figure 3 Edge-Detection Labeling Implementation

Moreover, the following table synthesizes the algorithm performance, in terms of false alarms, missed detections, correct detections and the detection range at the first intruder catching.

Table 2 Algorithm Performance

False Alarms	Missed Detections	Correct Detections	Detection Range
9 %	11 %	80 %	2400 m

However, the computation time has been evaluated by means of the laboratory test-bed [8], described more in detail later, and it resulted less than 100 ms, suitable for our real-time system.

3.3 Critical Issues

Let us observe that the false alarm percentage declared in table 1 is quite high from the algorithm reliability point of view; in fact that algorithm performance doesn't allow the EO system to supply the required auxiliary function to radar, but it can degrade the overall DS&A system performance.

In particular, the critical situations for false alarms risks can be classified in two main categories: presence of sun glares in the captured image, and horizon presence in the search window.

Next section aims at describing how those issues have been managed in order to reduce the false alarm rate and to realize a more robust system.

4 Improved Strategies for the development of the Image Processing technique

In the sun light in the image case, a "Sun Presence Detector" has been implemented. It evaluates the percentage of saturated pixels of the whole image and it is based on a statistical study applied on real images. From results it has been always verified that when sun is in front of the camera, more than 15% of the pixels are saturated. This percentage is the reference threshold considered by the detector; if the sun is detected, the edge detection algorithm is not run and the IP-CPU does not send any EO measurements to tracker.

An example of sun detector implementation is shown in figure 4.

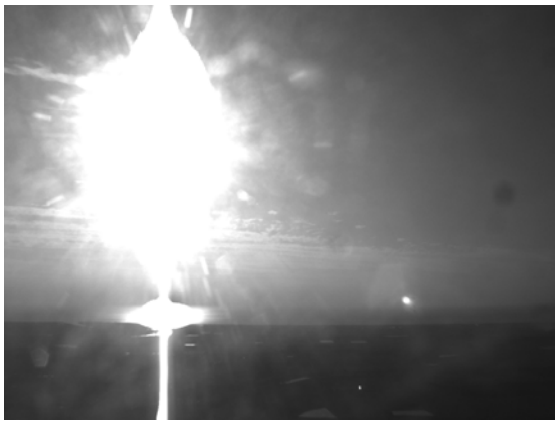


Figure 4 Sun in the image detected by the relative algorithm control

Horizon line presence in the search window is handled in different ways depending on intruder range. In particular, after edge detection-labeling, a further filter is implemented in order to compare the largest edge length to the instantaneous intruder range and to evaluate the probability that it is an intruder or background noise. The reference lengths are the wing span of a civil aviation aircraft (B737) as well as the wing span of smaller vehicles such as the TECNAM P92. They are calculated at different ranges, taking into account also the extension atmosphere and height effects, as stated in ref. 13. The concept is explained by the following equation:

$$C_R = e^{(-\sigma_v(h) * R)}$$

Where C_R is the attenuated length in percentage, R is the object range from the point of observation, and σ_v is the attenuation factor, depending on the object altitude h.

Therefore, the edge detection output is declared as intruder if it is enclosed in the range of reference wing spans. Secondly, the algorithm provides a higher-level control to reject false alarms since it must ensure that the detected edge has no comparable segments; otherwise it does not output any intruder position measurements

An example of horizon presence in the search window and of the resulting edge detection output is reported in figure 5.

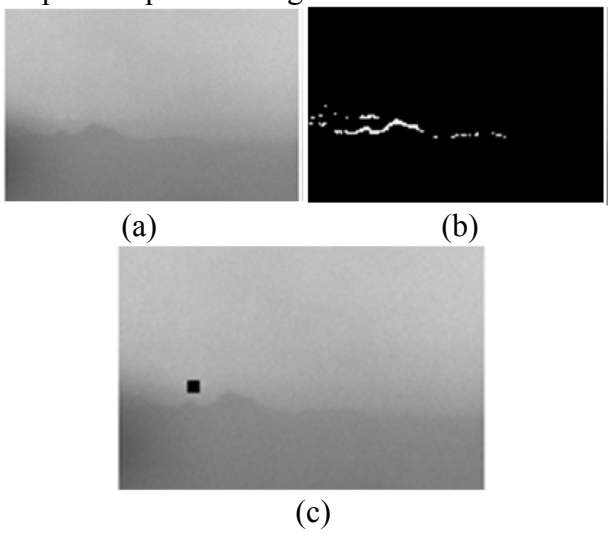


Figure 5 a) Search window enclosing horizon; (b) Edge detection-labeling implementation; (c) Image processing algorithm output.

5 Test-bed for analyzing the developed strategies

Real-time image processing algorithm performance have been tested by means of a hardware-in-the-loop (HWIL) facility realized to support development of the sense and avoid system onboard FLARE [8,14]. In particular, it aims at validating the main real time flight system operations, such as intruder tracking, image processing, and data fusion. In fact, it allows a wide range of flight scenarios to be reproduced, also those ones that are difficult to be replicated in real flights, and to compare the evaluated system performance with the expected one.

In the current configuration, the real hardware components tested are the visible EO sensor and the processing units, while the radar is replaced by a dedicated simulator which reproduces its outputs. A camera captures synthetic images of the intruder aircraft that are displayed on a LCD screen located in front of it. Figure 6 reports the overall HWIL architecture, while figures 7a and 7b depict the layout of camera, monitor and collimator on the optical bench and the black box which encloses all optical components.

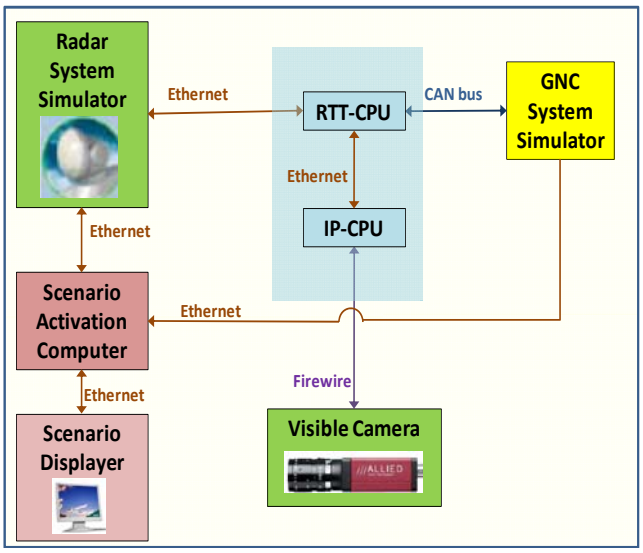
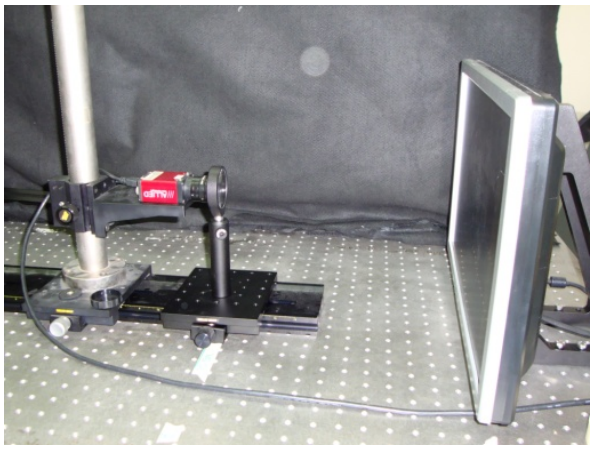


Figure 6 HWIL system architecture



(a)



(b)

Figure 7 Camera-Collimator-Monitor on optical bench; b. Black Box on optical bench

During laboratory tests data flow is the same as the one of the real flight system: RTT-CPU receives the estimates of target position from the Radar Simulator and transmits the relevant firm tracks estimates to the IP Computer which processes the displayed images as it is captured by the camera. Subsequent processing steps are identical to the one executed in real flights.

The indoor facility can operate in two operating modes: pre-recorded or simulated flight scenario. In the case of pre-recorded flight data, intruder position is output at a frequency of 1 Hz (GPS update rate), so that the scenario displayer computer performs an interpolation of intruder range, azimuth and elevation in order to generate images at a frequency of 20 Hz.

HWIL results regarding the EO obstacle detection performance relevant to tests based on flight data are reported in figure 8 and figure 9. In particular, they represent intruder azimuth and elevation in the Body Reference Frame (BRF) during a near collision encounter between FLARE and the intruder aircraft; therefore, the distance between the two aircraft decreases while they fly at almost the same altitude. The performance improvement of intruder position estimated by EO sensor compared to radar estimations can be clearly appreciated in terms of accuracy and measurement rate; moreover radar azimuth measurements are affected by a constant bias of about 1°, due to residual misalignment w.r.t AHRS. In this test the image processing system was interrogated at a frequency of 5 Hz and it exhibited a response time of less than 0.1 s, on average. Table 3 summarizes these results by comparing radar and EO camera from the point of view of accuracy and data rate of angular measurements.

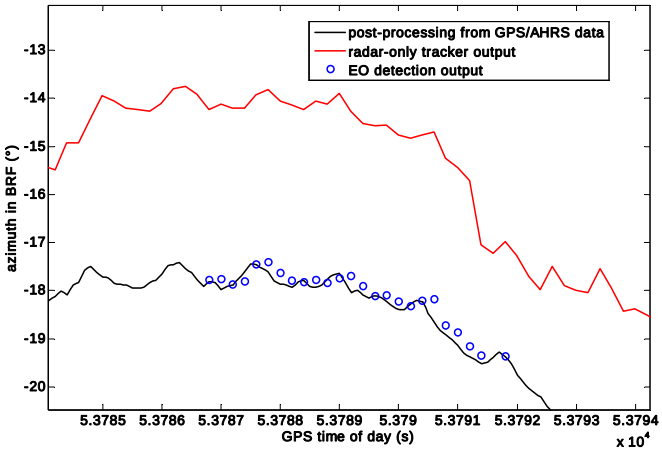


Figure 8 Intruder Azimuth in the BRF

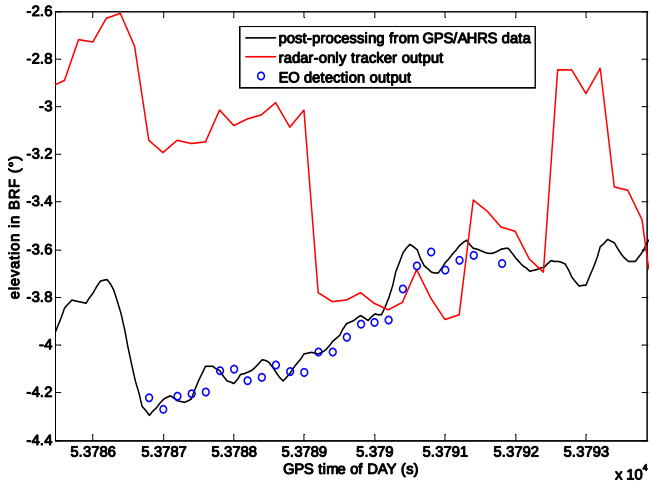


Figure9 Intruder Elevation in the BRF

Table 3 Sensors Angular Measurements Performance

	Accuracy (°)	Data Rate (Hz)
Radar	1.7	1
Panchro Camera	10^{-1}	5

6 Results

All algorithm controls added to reduce the risk of false alarms, i.e. the ones explained in the previous sections, have been applied. They allowed for increasing the EO object detection performance as it is synthesized in the Table 4.

Table 4 Edge detection–labeling performance optimized in the considered frontal encounters

False Alarms	Missed Detections	Correct Detections	Detection Range
1.6 %	37.6 %	60.8 %	2400 m

It is worth noting that these results are compliant with tracking requirements expressed in terms of false alarms rate, computation time, and detection range. In fact, a conservative assumption was made in the realization of the algorithm because the main camera does not need to output data with a constant data rate. Its main scope is to improve the accuracy of the tracker that is based on the radar as primary sensor. For this reason, it is preferable that the IP-CPU outputs to RTT-CPU fewer, but more reliable EO measurements in order to enhance the integrated tracker performance. In fact, the main effect of a missed detection is to reduce EO measurement rate and thus the algorithm capability to filter sensors noise and improve estimation accuracy, especially on angular derivatives. Instead, if a false alarm is input to the Kalman filter [5], an erroneous association may occur and, due to the small EO estimated measurement covariance. Consequently, it can have negative impact on tracker reliability such as biases generation and then it can determine a track loss. In future, this problem could be overcome by adding a real time estimate of EO covariance that is based on a guess on the quality of the performed image processing.

7 Conclusions

This paper is focused on the optimization of an image processing algorithm for a visible camera, that is integrated in a hierarchical sensor architecture for Sense and Avoid applications. The system is installed onboard a Very Light Aircraft and it is characterized by a multi-sensor solution: the radar, the main sensor, and four EO cameras, i.e. aiding sensors to radar, two visible and two IR. In particular, the panchromatic camera aims at increasing the tracking accuracy and data rate after the initial detection by the radar. Therefore, it is mandatory that the EO detection is reliable, fast, adaptable at different weather and illumination conditions (sunny, cloudy or rainy day), and it has a large detection range. The most reliable image processing technique for the presented application resulted the edge detection–labeling. It has been demonstrated that its computational load is adequate to real time systems, it works well on several image backgrounds and it has a detection range of about 2400 meters in clean air conditions. Moreover, the algorithm was improved in order to obtain fewer false alarms, thus increasing the overall reliability of the integrated system outputs.

Finally, the overall Sense and Avoid tracking system performance is shown to increase significantly with respect to the standalone radar tracking system performance. Indeed, radar can work in all-weather all-time conditions and it performs the direct range-to-obstacle measure, while EO system allows a better angular accuracy, about 0.5° with respect to 1.7° of radar measurements, and an increased data rate, up to 5 Hz instead of 1 Hz provided by the radar sensor.

Acknowledgements

The PhD grant of Ms. Lidia Forlenza was funded by CIRA.

References

[1]. EUROCONTROL, July 27th, 2007, "EUROCONTROL Specifications for The Use of Military Unmanned Aerial Vehicles as Operational Air Traffic Outside Segregated Airspace", doc no. EUROCONTROL-SPEC-0102, v.1.0, Bruxelles, Belgium.

[2]. US Federal Aviation Administration (FAA), March 27th, 2008, "Airworthiness Certification of Unmanned Aircraft Systems", Order 8130.34, Washington DC, USA.

[3]. US Department of Defence (DoD), August 4th, 2005, "Unmanned Aircraft Systems Roadmap 2005-2030," Washington DC, USA.

[4]. Cox, T.H., Somers, I., and Fratello, D.J., NASA DFRC, August, 2006, "Earth Observations and the Role of UAVs: A Capabilities Assessment,", v.1.1, Edwards CA, USA

[5]. Fasano, G., Accardo, D., Moccia, A., Carbone, C., Ciniglio, U., Corraro, F., and Luongo, S. , 2008 , "Multi-Sensor-Based Fully Autonomous Non-Cooperative Collision Avoidance System for Unmanned Air Vehicles," AIAA Journal of Aerospace Computing, Information, and Communication, vol. 5, issue 10.

[6]. Maroney, D. R., Bolling R. H., Athale, R., and Christiansen, A. D., 2007, "Experimentally Scoping the Range of UAS Sense and Avoid Capability", 2nd AIAA Infotech@Aerospace Conference and Exhibit, Paper 2007-2850, Rohnert Park CA, May, pp 1-16.

[7]. Forlenza, L., Fasano, G., Accardo, D., Rispoli, A., "Image Processing Algorithm for Integrated Sense And Avoid Systems", Proceeding of SPIE Defence + Security Conference, Volume 7833, 20 - 23 September 2010, Toulouse, France.

[8]. Forlenza, L., Fasano, G., Accardo, D., Moccia, A., "LABORATORY TEST FACILITY FOR SIMULATING A SENSE AND AVOID FLIGHT SYSTEM", Proceeding of Infotech@Aerospace AIAA Conference, 20 – 22 April 2010, Atlanta, Georgia.

[9]. Pratt, W.K., "Image Analysis", Chap. 5 in Digital Image Processing, Wiley Interscience, Ed., pp. 447-625, Mountain View, CA, USA (1991).

[10]. Nadernajad, E., Sharifzadeh, S., "Edge Detection Techniques: Evaluations and Comparisons", Applied Mathematical Sciences, Vol. 2, 2008, no. 31, 1507 – 1520

[11]. Trucco, E., and Verri, A., Introductory Techniques for 3-D Computer Vision, Prentice Hall, 1998

[12]. Jain, R., Kasturi, R., and Schunk, B., Machine Vision, Computer Science Series, 1995 McGraw-Hill International Editions

[13]. Downey, G., A., "Electro-optical Tracking Considerations II", Proc. of SPIE Acquisition, Tracking and Pointing XVIII, Vol. 5082

[14]. Forlenza, L., Fasano, G., Accardo, D., Moccia, A., Rispoli, A.,"A hardware in the loop facility for testing multisensor sense and avoid systems", 28th Digital Avionics Systems Conference, 25-29 October 2009, Orlando FL, USA, Print ISBN: 978-1-4244-4078-8, Digital Object Identifier : 10.1109/DASC.2009.5347464

[15]. Fasano G., Accardo D., Moccia A., Rispoli A., "An Innovative Procedure for Calibration of Strapdown Electro-Optical Sensors Onboard Unmanned Air Vehicles", Sensors. 2010; 10(1):639-654

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Sense & Avoid Algorithm based on Aircraft Performances Estimation

Marco Melega, Samuel Lazarus, and Al Savvaris
FMCS Department, Cranfield University, Bedfordshire, MK43 0AL, United Kingdom

Keywords: sense and avoid, autonomous collision avoidance, path following, unmanned aircraft vehicle

Abstract

In this paper an Autonomous Navigation System (ANS) integrating both the Path Following (PF) and the autonomous Sense And Avoid (S&A) functions is proposed. The S&A algorithm is developed in order to guarantee an avoidance manoeuvre that ensure a minimum separation between the ownship and all the other agents during its overall execution in multiple flying threats scenarios. The proposed S&A system is modeled in the Matlab/Simulink® environment for a Piper J3 Cub 40 model aircraft. The threats considered are aircrafts that communicate their state to the system through their Automatic Dependent Surveillance-Broadcast (ADS-B) mode S transponders.

1 General Introduction

The unavoidable separation between the *Unmanned Aerial Vehicles (UAVs)* and their pilots leads to a significant reduction of the human piloting performance, due to the loss of sensory cues valuable for flight control and challenges in providing the pilot with complete situational awareness of the surrounding environment around the air vehicle [1]. This results in worsening the knowledge of the pilot regarding the current estimation and the prediction of the position of the *UAV*, other flying agents and terrain hazards. Moreover it involves also the lack of perception of the distance from the local ground level and the weather conditions in the operating region of the *UAV* [2].

Therefore it is necessary to introduce the *S&A* system which can be defined as an artificial system equivalent to the human pilot for detecting and avoiding hazard situations such as midair collisions, controlled flight into terrain, flight path obstacles and clouds. This imply that it must provide a self-separation of the aircraft in case if other methods fail to prevent the conflict [3]. The development of this system is still an open research and involves different issues [4]. One of the most difficult issues that need to be solved is the definition of the configuration of the sensors to detect and to identify the conflict risks. The most common solutions involve transponders mode S or emerging technologies like *ADS-B* [6], radars [7–9] or cameras [10–14] or mixed configuration of sensors of both types [15] [5]. Another important topic to which still no accepted solution has been defined is the algorithm that is used for the conflict detection and the resolution manoeuvre evaluation [5]. Further work is also needed in developing and defining the type and the layout of *Human-Machine Interface (HMI)* that guarantee an easy and effective interaction between the system and the pilot [16, 17]. The interactions and the trade-offs between all the aforementioned issues should be defined by considering the certification requirements, the limita-

tions and the performances of the technology. What also needs to be considered are the effects of the human pilot involvement in the system management when it is not completely autonomous.

A prototype of *S&A* algorithm for single threats scenarios was developed in [18]. It evaluate the avoidance manoeuvre as a simple step finite variation of the command imposed to the heading angle autopilot of the aircraft. The heading angle change is optimized in order to keep the distance between the ownship and the flying threat superior or equal to a minimum value of the horizontal separation during the overall manoeuvre. In this paper this algorithm is further developed in order to estimate vertical resolutions that guarantees a minimum vertical separation during the overall manoeuvre. This approach is then extended to cover the multiple threat scenarios. The developed algorithm is integrated in an *ANS* together with a *PF* algorithm where the flying threats considered are aircraft equipped with mode S *ADS-B* transponders. The overall system was modeled in *Matlab/Simulink*® environment, including the *HMI*.

The implementation described here gives the possibility to the pilot to use the system with the *PF* algorithm in switched on (autonomous navigation mode) or off (manual navigation mode). The paper concludes with the results of two simulation scenarios for the two operative modes in which two simultaneous head-on conflict risks are considered.

2 Preliminary definitions

The Cartesian reference frame used for the development of the algorithm is shown in Fig. 1, where an arbitrary vector $\vec{v}$ can be defined through its Cartesian components according to the relationship:

$$\vec{v} \triangleq (v_x, v_y, v_z) \tag{1}$$

but also through spherical coordinates in terms of modulus $|\vec{v}|$, azimuth and elevation angles $\alpha(\vec{v})$ and $\beta(\vec{v})$ defined from the following relationships:

$$|\vec{v}| = \triangleq \sqrt{v_x^2 + v_y^2 + v_z^2} \tag{2}$$

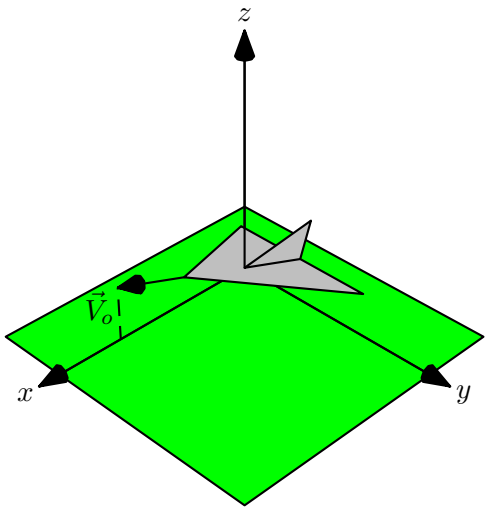


Fig. 1 Cartesian reference frame considered in the developed algorithm: the origin is in the current position of the ownship, the x and y axis are in the local horizontal plane pointing towards the projection of the aircraft direction of motion (defined from the ownship airspeed $\vec{V}_o$) and the left wing respectively and the z axis is directed upwards.

$$\alpha(\vec{v}) \triangleq \arctan \frac{v_y}{v_x} \tag{3}$$

$$\beta(\vec{v}) \triangleq \arctan \frac{v_z}{\sqrt{v_x^2 + v_y^2}} \tag{4}$$

Another useful parameter is the horizontal projection modulus $|\vec{v}|_h$ defined as:

$$|\vec{v}|_h \triangleq \sqrt{v_x^2 + v_y^2} \tag{5}$$

Supposing that the ownship position estimation comes from a *Global Positioning System (GPS)* receiver, it is necessary to take into account the uncertainty related with the position estimation accuracy of both ownship and threats. This necessitates to define a *position uncertainty zone* as the locus in which the true aircraft position lies. According to the conventional format used to express the accuracy of the GPS positions estimation, a cylindrical region is considered which is centered on the aircraft position with a radius of r_p and a half-height of h_p respectively that are equal to the horizontal and the vertical position estimation accuracy as

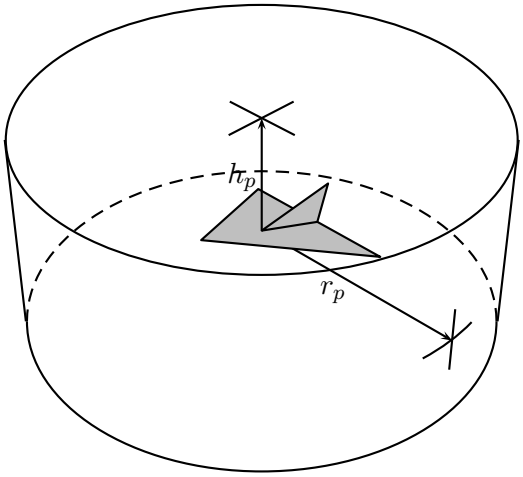


Fig. 2 Extension and geometric parameters of the position uncertainty zone

shown in Fig. 2. In addition to this a cylindrical safety zone region is defined as a position uncertainty zone the dimensions of which are increased by a safety factor in order to take into account the approximations introduced in the resolution manoeuvre estimation.

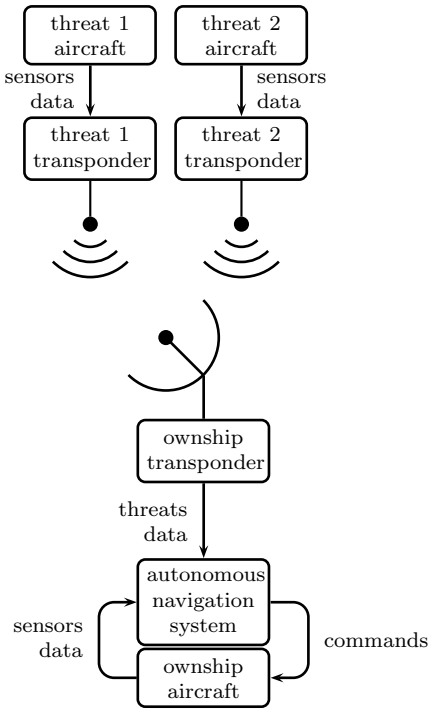


Fig. 3 Structure of the simulation environment used to test the *ANS*

3 The system structure

The simulation environment used to develop and test the *ANS* system is sketched in Fig. 3. Both the ownship and the threats aircraft blocks are obtained from the simulator of a *Piper J3 Cub 40* model aircraft [18]. The on-board *ANS* receives the sensors data regarding the ownship aircraft current state and sends the commands to the autopilots in order to be executed. The sensors data include the threats transponders broadcast, the ownship aircraft *GPS* position, altitude, true and ground airspeed, magnetic heading angle and vertical speed. In order to simulate the real communication between threats, the data exchanges between the transponder blocks are enabled only when the horizontal distance between the aircrafts is inferior to the minimum among the reception range of the ownship transponder and the transmission range of the threats transponders. This distance is calculated by transforming the *GPS* position of the threat from the geocentric reference frame to the *North-East-Down Reference Frame (NED)* which is centered in the current ownship position and estimating the length of the resulting vector.

The structure of the *ANS* is shown in Fig. 4. The switch with the label **user selection** selects the modes between the autonomous and the manual navigation mode of the system. The collision risks and safety alerts are processed by the **conflict detection** algorithm. If one or more potential conflicts are detected, the resolution manoeuvre is computed in the **resolution evaluation** algorithm. The informations about the alerts and the recommended resolution manoeuvre are then displayed via the *HMI* to the pilot. The **autonomous collision avoidance system** takes over and executes a collision avoidance manoeuvre "if" there is no command input given from the pilot.

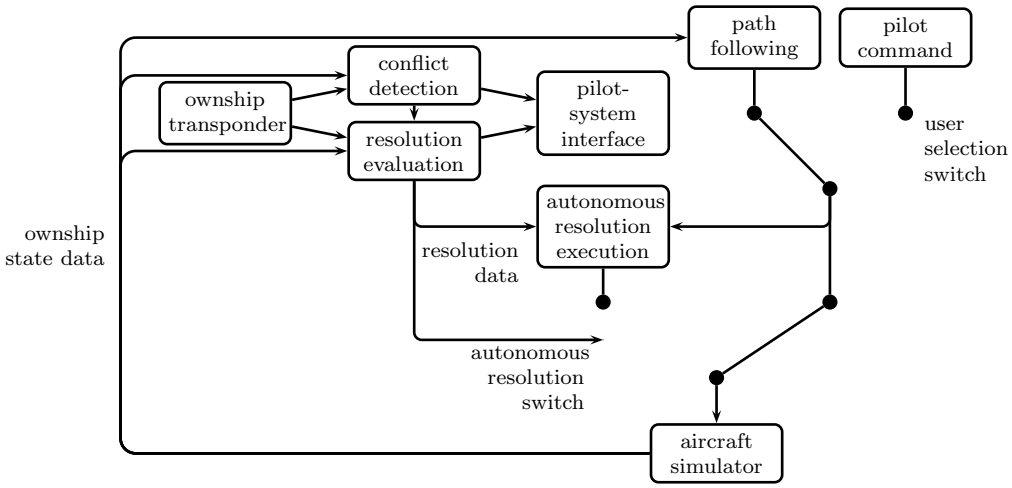


Fig. 4 Block diagram of the *ANS* system

4 Path Following Algorithm

The *PF* algorithm navigates the aircraft through a given set of predefined waypoints. These are defined in geocentric coordinates as:

$$\vec{P}_{W/P} \triangleq \left(\phi_{W/P}, \lambda_{W/P}, h_{W/P}\right) \tag{6}$$

where $\phi_{W/P}$, $\lambda_{W/P}$ and $h_{W/P}$ are the latitude, longitude and altitude, respectively. The airspeed $V_{W/P}$ is controlled between each of the waypoints and it is maintained to be a constant one.

If it is assumed that the ownship position is:

$$\vec{P}_{O/S}(t) \triangleq \left(\phi_{O/S}(t), \lambda O/S(t), h_{O/S}(t)\right) \tag{7}$$

Then the waypoint can be represented in the *NED* reference frame which is centered in $\vec{P}_{O/S}(t)$. Therefore, the heading angle command $\psi_d(t)$ and the airspeed command $V_d(t)$ are:

$$\psi_d(t) = \alpha\left(\vec{x}_{W/P,h}(t)\right) \tag{8}$$

$$V_d(t) = V_{W/P} \tag{9}$$

The altitude command $h_d(t)$ can be defined as a linear variation in the altitude during the flight between each of the waypoints. In order to accomplish this, the vertical speed $V_z(t)$ has to be maintained constant and it can be estimated as:

$$V_z(t) = \frac{h_d(t) - h_{W/P}}{t_{W/P}} \tag{10}$$

where, $t_{W/P}$ is the elapsed time to reach the waypoint and this can be calculated as:

$$t_{W/P} = \frac{\left|\vec{x}_{W|P}\right|}{V_{O/S}} \tag{11}$$

therefore, the mathematical expression of the altitude command $h_d(t)$ is:

$$h_d(t) = \int_0^t V_z(t)\,\mathrm{d}\,t \tag{12}$$

Due to the uncertainties in the aircraft position, the aircraft may not exactly reach the desired waypoint. Therefore the aircraft is assumed to reach the desired waypoints if the following condition is satisfied:

$$\left|\vec{x}_{W/P}\right| < \epsilon \tag{13}$$

where $\epsilon_{W/P}$ is a threshold value defined based on the aircraft position accuracy. If the above defined condition is fulfilled, then the next waypoint has to be considered for the next iteration.

5 Conflict detection and resolution manoeuvre evaluation

According to the aforementioned definition in section 2, the ownship maintains the constant position during the whole mission (i.e. in the origin of the reference frame). Based on

the straight projection hypothesis introduced in [18], every other aircraft in the close airspace move on a straight line. This line passes from the current position of the threat and is parallel to its relative airspeed $\vec{V}_r$ defined as the sum of the ownship's airspeed vector $\vec{V}_o$ and the threat's airspeed vector $\vec{V}_i$, as shown in Fig. 5. If the threat trajectory intersect the position uncertainty zone, therefore a collision risk is being detected. In this case, a resolution manoeuvre has to be generated in order to keep every threat outside the safety zone during the overall manoeuvre. Moreover, the control input variation has to be minimized in order to reduce the deviation distance between the resolution manoeuvre and the aircraft original trajectory.

If the ownship is assumed to be equipped and controlled with *Flight Path Control System (FPCS)*, the interactions between the longitudinal and latero-directional dynamics during the manoeuvres are negligible. This implies that it is possible to define the resolution manoeuvres in the horizontal and vertical plane independently.

Considering a single threat, the resolution manoeuvre in horizontal plane is a heading angle step command that falls within the safe heading angle range $\Delta\Psi res$ [18].

The resolution manoeuvres in vertical plane could be simply defined as the altitude variation Δh_{res} necessary to put the ownship in a final stationary condition in which the vertical distance from the threat is equal or superior to the half-height h_s of the safety zone (Fig. 6). Therefore the range ΔH_{res} of the safe altitude changes that can be used for the resolution manoeuvre is:

$$\Delta H_{res} = [-\infty, h_o(t) - (h_i(t) + h_s)] \cup \\ \cup [h_o(t) + (h_i(t) + h_s), +\infty] \quad (14)$$

In the manual navigation mode, the switching between the resolution execution in the horizontal and vertical planes is automatically controlled by introducing two binary variables $\chi_{h,AR}$ and $\chi_{v,AR}$.

The switch $\chi_{h,AR}$ for the autonomous horizontal resolution manoeuvre and its switching principle are described in [18].

The switch $\chi_{v,AR}$ for the autonomous vertical resolution manoeuvre execution is defined

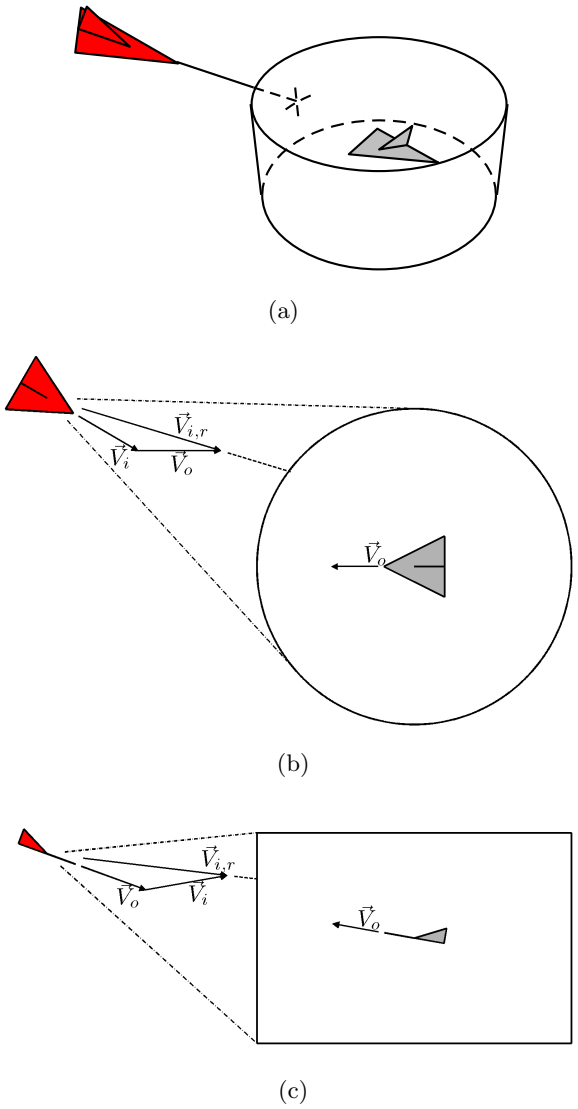


Fig. 5 Conflict detection approach: (a)three dimensional space, (b)horizontal plane and (c)vertical plane. The red aircraft is the threat and the gray one is the ownship.

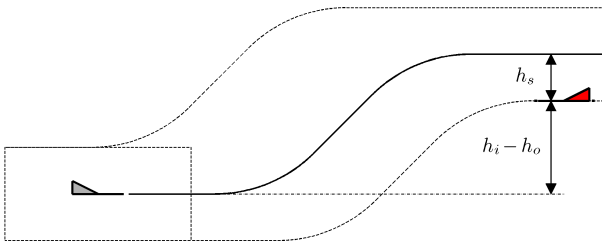


Fig. 6 Approach used for the vertical resolution evaluation: the red aircraft is the threat and the gray one id the ownship.

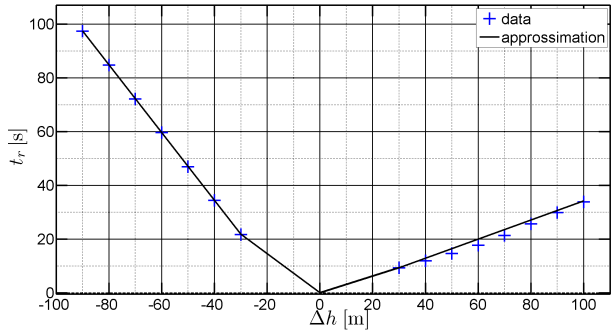


Fig. 7 Estimation of the t_r values for a discrete grid of Δh value and comparison with the approximation (vertical view)

by considering the horizontal distance necessary for the aircraft to reach the new altitude commanded by the resolution manoeuvre. This requires the estimation of the **rise time** t_r, which is the time taken by the aircraft to move from the current altitude to the new commanded altitude. The rise time can be estimated from the response of the aircraft based on a discrete grid of altitude variations Δh, starting from the same initial cruise condition. These values can be used to get a linear segment regression as given in the equation below and shown in Fig. 7:

$$t_r = \begin{cases} 21.82 - 1.26(\psi_j + 30) & \psi_j < -30 \\ -0.73\psi_j & -30 < \psi_j < 0 \\ 0.31\psi_j & 0 < \psi_j < 30 \\ 6.25 + 0.36(\psi_j - 30) & \psi_j > 30 \end{cases} \tag{15}$$

Once t_r is estimated, the minimum distance for a safe manoeuvre execution can be calculated as:

$$d_{\min,res} = V_{GAS,c} t_r \tag{16}$$

where $V_{GAS,c}$ is the estimated ground airspeed during the manoeuvre. This is assumed to be equal to the aircraft true airspeed to get conservative results for the head-on encounters. Therefore $\chi_{v,AR}$ is true when the following condition holds:

$$\left|\vec{P}_{i,r}(0)\right|_h = d_{\min,res} \tag{17}$$

In the manual navigation mode, the algorithm is defined in order to operate both the horizontal and the vertical resolution manoeuvre simultaneously as soon as any one of the autonomous resolution execution switches $\chi_{h,AR}$ or $\chi_{v,AR}$ is switched on. This means that it is possible to define an overall autonomous resolution execution switch χ_{AR} as:

$$\chi_{AR} = \chi_{h,AR} \vee \chi_{v,AR} \tag{18}$$

where $\vee$ indicates the logical operator *or*. In the autonomous navigation mode, an optimal navigation is obtained where the resolution manoeuvre is initiated as soon as the collision risk is detected and therefore the switch χ_{AR} is not used.

In this paper, the above defined approach is extended to a multiple threats scenario. Considering n threats, the safe heading angle, altitude ranges and the autonomous resolution switches are considered as $\Delta\Psi_1, \Delta\Psi_2, ..., \Delta\Psi_n$, $\Delta H_1, \Delta H_2, ...\Delta H_n$ and $\chi_{1,AR}, \chi_{2,AR}, ..., t_n\chi_{n,AR}$ respectively. Then the resulting resolution heading angle and the altitude ranges $\Delta\bar{\Psi}$ and $\Delta\bar{H}_{res}$ can be defined as the intersection between the ranges calculated for every threat:

$$\Delta\bar{\Psi}_{res} = \bigcap_{i=1}^{n} \Delta\Psi_i \tag{19}$$

$$\Delta\bar{H}_{res} = \bigcap_{i=1}^{n} \Delta H_i \tag{20}$$

The general autonomous resolution switch $\bar{\chi}_{AR}$ is switched on as soon as any particular switch which relates to any one of the threats is switched on.

In order to minimize the gap with the optimal navigation trajectory, the effective heading angle value used for the heading angle change $\Delta\psi_{res}$ has to be chosen as the minimum absolute value in the range $\Delta\bar{\psi}_{res}$. Similarly, the resolution altitude change Δh_{res} has to be defined as the value of the set $\Delta\bar{H}_{res}$ that has the minimum absolute value.

6 System-Pilot Interface

The developed *HMI* display for the system is shown in Fig. 8. It includes a panel with the simulation controls, a panel with the switches of the system functionalities, the *PF* panel and the *S&A* panel.

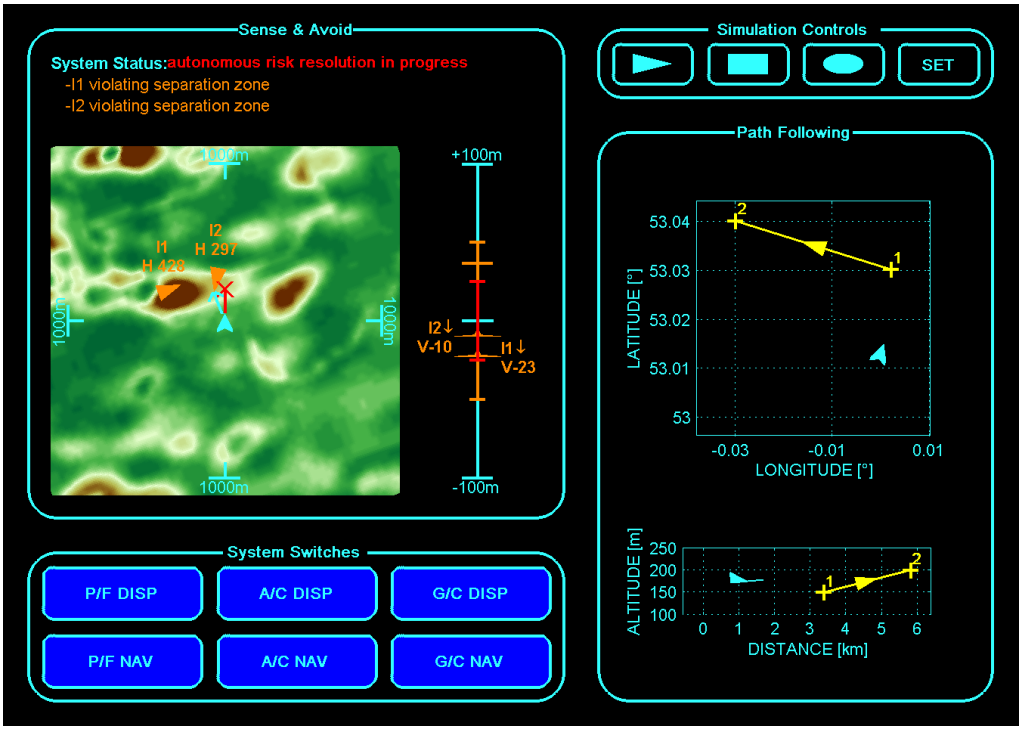


Fig. 8 Snapshot of the *HMI* display of the *ANS* system

The *PF panel* shows the position of the waypoints with respect to the current position of the aircraft and the route that the aircraft has to follow for an optimal path. It includes a horizontal display of the positions expressed in latitude and longitude and a vertical display of the distances in kilometers between the current position of the aircraft and the waypoints.

The *S&A panel* is mainly used to display informations about the threats, system alerts, resolution manoeuvres and the ground terrain topology.

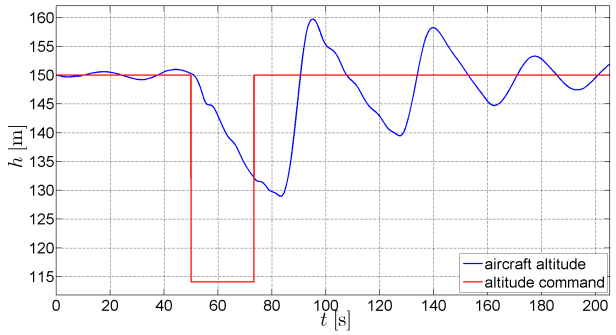
7 Results and discussion

The overall architecture of the proposed approach was modeled and validated through simulations in *Matlab/Simulink*® environment. In particular, the results of two different simulations are presented for a two threats scenario considering head-on encounters. The simulations are based on the manual and the autonomous navigation modes of the *ANS* (i.e. the *PF* algorithm is switched off and on modes), respectively. The position uncertainty zone and the safety zone are set respectively with a radius of 50 m and 100 m, with a half-height of 25 m and 35 m. The ownship starts from its initial position with the latitude of $53.030° N$ and longitude $0.002° W$ in a cruise condition at an altitude of 150 m, airspeed of 15 m/s and the heading direction towards the North. One of the two threats is assumed to start from the latitude of $53.016° N$ and the longitude of $0.016° W$ cruising at the same altitude and airspeed of 14 m/s with the heading direction of 90° from North. The second one is assumed to start from $53.022° N$ (latitude) and $0.0065° W$ (longitude) with the heading direction of 210° from the North and descending from an altitude of 200 m to 100 m; airspeed of 14 m/s. The waypoints used in the autonomous navigation modes are tabulated in Table 1.

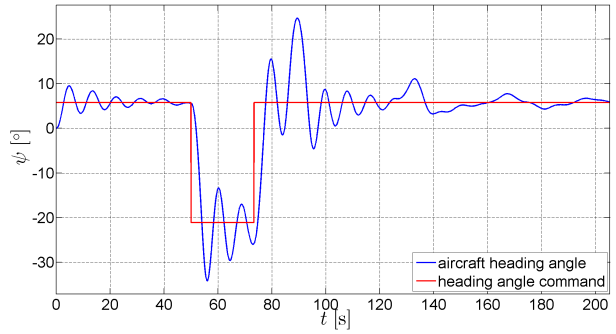
The results of the simulations for the *ANS* in manual and autonomous navigation mode are shown in Fig. 9 and Fig. 10, respectively. The altitude profiles and the altitude commands of the ownship are shown in Fig. 9a and Fig. 10a. The aircraft heading angle profile and the heading angle commands are shown in Fig. 9b and

waypoint name	**latitude**	**longitude**	**altitude** [m]	**airspeed** [m/s]
WP 1	$53.030°N$	$0.002°E$	150	15
WP 2	$53.040°N$	$0.030°W$	200	15

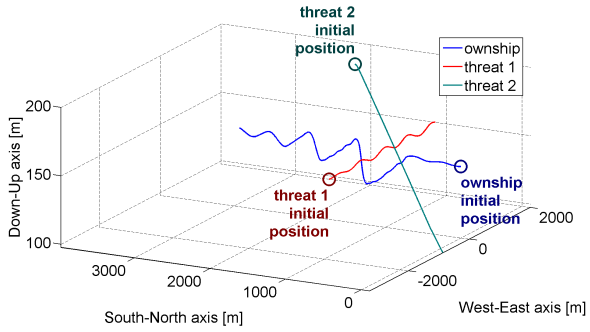
Table 1 Data of the waypoints considered in the autonomous navigation mode simulation



(a)

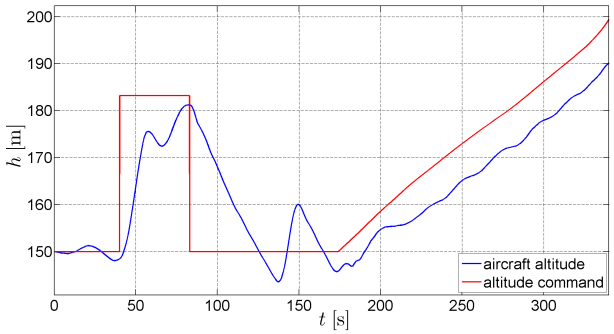


(b)

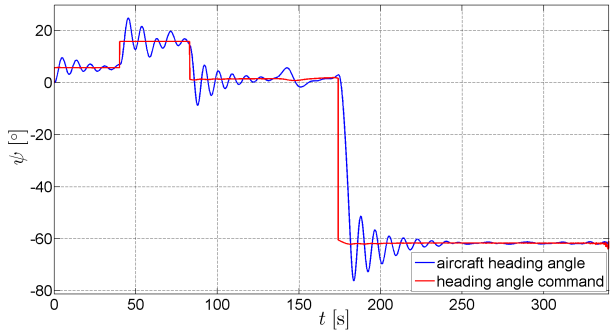


(c)

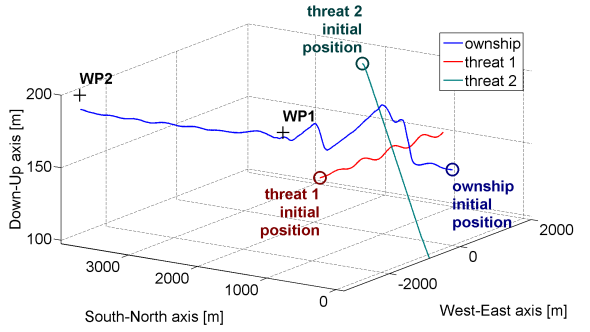
Fig. 9 Results of the *ANS* system simulation for the manual navigation mode: (a)altitude profile and command, (b)heading angle profile and command, (c)three dimensional trajectory of ownship and threats



(a)



(b)



(c)

Fig. 10 Results of the *ANS* system simulation for the autonomous navigation mode: (a)altitude profile and command, (b)heading angle profile and command, (c)three dimensional trajectory of ownship and threats

Fig. 10b. Finally the three dimensional trajectory paths of the ownship and the threats for both manual and autonomous navigation modes are shown in Fig. 9c and Fig. 10c.

Regarding the *PF* algorithm performances, it is worth mentioning there is about 10 meters undershot in reaching the desired altitude of the second waypoint in Fig. 10a. The minimum values of separation between the ownship and each threat in the first simulation are 94 m and 89 m, respectively. This means that the threats do not enter in the position uncertainty zone, despite their trajectories pass inside the safety zone. The difference between the minimum distances and the maximum dimension of the safety zone is however limited to the 6% and the 11% of the last value.

More results are obtained in the manual navigation mode, in which the minimum distances between the ownship and the threats are 120 m and 172 m, respectively. These results imply that the threats are always outside the safety zone during all its manoeuvres.

This difference in the performances can be related in the two cases to the resolution manoeuvre starting point. In fact in the manual navigation mode the resolution is executed when the distance between the ownship and closest threat is the minimum value that guarantee a safe avoidance manoeuvre. In the autonomous navigation mode, it is initiated as soon as the threat is detected, (i.e. the distance between ownship and threats is generally greater than the previous case). Despite this difference, the approach used for the manoeuvre evaluation is the same and involves the estimation of the future trajectory of the ownship during the manoeuvre. This prediction is more reliable in the short term than in the long term due to the hypothesis used. Therefore, when the resolution maneuver is executed in the manual navigation mode, the deviation between the predicted and the true trajectory is inferior and the performances are more conservative.

8 Conclusions and future works

The results presented in the paper provide a proof of acceptable performances of the *S&A* algorithm for the *ANS* in both autonomous and manual navigation modes. The proposed algorithm is particularly suitable for *Remotely Piloted Vehicle (RPV)* as a safety pilot support tool. However it is also reliable for an autonomous navigation system similar to the one described in this investigation. The *PF* algorithm works well apart from a small delay in reaching the desired altitude. Therefore the results obtained from the integration of the two systems guarantees an acceptable *PF* algorithm performance for the *ANS*.

Further developments can be proposed for the described *S&A* system with the detection and avoidance of the ground obstacles. Additionally the *PF* algorithm can be improved by integrating an optimal waypoint navigation system with filtering techniques. Finally it is also necessary to have pilot evaluations of the *HMI*.

References

[1] Gawron V.J., "Human factors issues in the development, evaluation, and operation of uninhabited aerial vehicles", *AUVSI '98: Proceedings of the Association for Unmanned Vehicle Systems International*,Huntsville, Alabama, 1998, pp. 431-438.

[2] Drury J.L., Riek L. and Rackliffe N., "A decomposition of UAV-related situation awareness", in: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (HRI '06) ACM*, New York, New York, 2006, pp. 88-94

[3] Australian Transport Safety Bureau (*ATSB*), *Limitations of the See-and-Avoid Principle*, 1991

[4] Zeitlin A. D., "Developing Requirements for the Unmanned Aircraft Sense & Avoid Function", *AIAA Infotech@Aerospace Conference*, Seattle, Washington, 2009, report number 2009-1924

[5] Albaker B. M. and Rahim N. A., "A Survey of Collision Avoidance Approaches for

Unmanned Aerial Vehicles", *2009 International Conference for Technical Postgraduates (TECHPOS)*, Kuala Lumpur, Malaysia, 2009, pp. 1-7

[6] Martel F., Schultz R. R., Semke W. H., Wang Z. and Czarnomski M., "Unmanned Aircraft Systems Sense and Avoid Avionics Utilizing ADS-B Transceiver", in: *AIAA Infotech@Aerospace Conference*, Seattle, Washington, 2009, report number 2009-1923

[7] Kwag Y. K. and Hwang K. Y. (2005), "Performance Simulation of Radar Sensor Based Obstacle Detection and Collision Avoidance for Smart *UAV*", *Digital Avionics Systems Conference, 2005. DASC 2005. The 24th*, Washington D. C., D. C., 2005, p.1-10, Vol. 2

[8] Kemkemian S., Nouvel-Fiani M., Cornic P. and Garrec P., "MIMO Radar for Sense and Avoid for UAV", *Phased Array Systems and Technology (ARRAY), 2010 IEEE International Symposium on*, Waltham, Massachusetts, 2010, p.573-580

[9] Freeman P. and Moosbrugger P., "A Low Cost Phased Array Solution for UAV Collision Avoidance", *PAIAA Infotech@Aerospace 2010*, Atlanta, Georgia, 2010, p.573-580

[10] Shah S.I.A. and Johnson E.N., "3D Obstacle Detection Using a Single Camera", *AIAA Guidance, Navigation, and Control Conference*, Chicago, Illinois, 2009, p.573-580

[11] Call B., Beard R. and Taylor C., "Obstacle Avoidance For Unmanned Air Vehicles Using Image Feature Tracking", *AIAA Guidance, Navigation, and Control Conference and Exhibit*, Keystone, Colorado, 2006, p.573-580

[12] Carnie R., Walker R. and Corke P., "Image Processing Algorithms for UAV *Sense and Avoid*", in: *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, Orlando, Florida, 2006, p.573-580

[13] Dennis A.,Archibald J., Edwards B. and Lee D.J., "On-Board Vision-Based Sense-and-Avoid for Small UAVs", *AIAA Guidance, Navigation and Control Conference and Exhibit*, Honolulu, Hawaii, 2008, p.573-580

[14] Watanabe Y., Calisey A. J. and Johnson E.N., "Vision-Based Obstacle Avoidance for UAVs", *AIAA Guidance, Navigation and Control Conference and Exhibit*, Hilton Head, South Carolina, 2007, p.573-580

[15] Fasano G., Forlenza L., Accardo D. and Moccia A., "Integrated Obstacle Detection System based on Radar and Optical Sensors", *AIAA Infotech@Aerospace 2010*, Atlanta, Georgia, 2010, p.573-580

[16] Tadema J., Theunissen E. and Kirk K. M., "Self Separation Support for UAS", *AIAA Infotech@Aerospace 2010*, Atlanta, Georgia, 2010, p.573-580

[17] McCarley J.S. and Wickens C.D., *Human Factors Implications of UAVs in the National Airspace*, Technical Report AHFD-05-05/FAA-05-1, University of Illinois, Aviation Human Factors Division, Savoy, Illinois, 2005

[18] Melega M., Lazarus S. and Savvaris A., "Conflict Resolution Manoeuvre Definition based on Aircraft Performances", *30th Digital Avionics Systems Conference*, Seattle, Washington, 2011

[19] Carr J.J. and Brown J.M., *Introduction to Biomedical Equipment Technology* (3th ed.), Prentice Hall, 2010

[20] European Organization for Civil Aviation Equipment (EUROCAE), *WG-73: Unmanned aircraft systems , A concept for UAS Airworthiness Certification and Operational Approval - Volume 2: UAS operations*, 2010

[21] *UK* Secretary of State for Transport, *The Rules of the Air Regulations 2007*, The Stationery Office Limited, *UK*, 2007

Route Planning Under Uncertainty For a Lighter Than Air Robot in Bridge Monitoring

Yasmina Bestaoui, Associate Professor
Nabil Ait Taleb, Sadek Oukid, Master Students
Laboratoire IBISC, Université d'Evry, France

Keywords: *UAV, Planning, Uncertainty, Markov Process*

Abstract

A Lighter Than Air Robot (LTAR) is an airship capable of sustained flight with no direct human control and able to perform a specific task. An important potential application is bridge monitoring. The disaster caused by the collapse of one of Minneapolis (Minnesota) highway bridges points to the need for better technologies to inspect bridges.

1 Introduction

Lighter than air vehicles behave and function differently to heavier than air craft: airplanes and helicopters. With sufficient autonomy, an unmanned lighter than air vehicle can be considered as a Lighter Than Air Robot (LTAR) [1,2]. It does not require any motor action to maintain a certain altitude and position in the space as it relies on low density gas inside the envelope to balance its own weight. It uses lift to float in the air in ways similar to submarines buoyancy in the water. One way to deflect some of the concerns associated with high fuel consumption of heavier than air craft is to rely on lighter than air vehicles. They offer an unmatched capability to fly for long periods of time and to do it silently. Airships float in the atmosphere due to their buoyancy so are airborne when filled with gas, before being launched and remain airborne until the gas eventually is released. With this feature, a lighter than air robot could conduct a continuous aerial operation with very low energy consumption. Non rigid airships with a cigar shaped profile are the most common type. These airships do not have any rigid internal framework.

The missions foreseen are environmental monitoring, inspection, surveillance and other applications ... Another application founds its roots in a growing worldwide interest in a new concept consisting of using autonomous atmospheric flight vehicles as platforms operating for extended periods of time at very high altitudes (between 20,000 and 100,000 m) to accomplish missions actually accomplished using spacecraft. As airships need less energy to stay aloft, they are considered as potential near-space platforms [3,4,5].

Lighter than air robots have essentially two flight regimes: hover and cruising flight:

- Aerostatics hovering (Hover): The speed of the LTAR relative to the surrounding air is small such that few or no forces act on the vehicle, except those resulting from the propulsion itself. The ability to hover is extremely useful for any operations that require close proximity to terrain or obstacles.}

- Aerodynamics flight (Cruising Flight): There is a significant relative speed between the vehicle and its surrounding environment and aerodynamics forces act on the vehicle. These aerodynamic forces then largely dominate those generated by the power system. Cruising flight usually results in the aerial robot constantly meeting upcoming air, which makes the range of adverse events to flight quite narrower.

From the earliest days of aeronautical experimentations, the natural wind proved itself to be a major parameter to successful flights [6,7]. The wind mostly affects a trajectory through its speed. In general, the wind speed can be modeled as a sum of two components: a nominal deterministic component (available through meteorological forecasts or measured with a Doppler radar) and a stochastic component, representing deviations from the nominal one. The closed loop controller takes care of the stochastic part considered as perturbations, while the deterministic component is introduced into the motion planner.

This paper is organized as follows. Section 2 introduces Flight planning while section 3 presents the bridge monitoring flight simulator. Section 4 is a brief review of various approaches in route planning with uncertainty and Section 5 shows the application of

2 Flight Planning

2.1 Introduction

Flight planning routines attempt to create paths that are fully consistent with the physical constraints of the airship, the obstacle avoidance, optimum flight path and weighed regions. Weighed regions are regions with abnormally low or high pressure, wind speeds or any other factor affecting flight. Vehicle autonomy is a discipline fertilized by the robotics and computer science fields.

The purpose of 3D mission is to create a pathfinder. This pathfinder generates the path from the initial point to the mission goal and navigates the vehicle. The position, orientation and speed of the vehicle are known from the sensors data and information about the meteorological conditions and probably obstacles to avoid is available. These elements are connected and close the cycle of the vehicle command and control.

Aerial robots require fuel, power and time resources to move about their environment. To fully define a trajectory, a path must be augmented with velocities and angular motion parameters. Resources costs in the form of forces/torques and traversal times can then be computed from the governing equations of motion. To incorporate quantities such as fuel and time into a traditional path planner's cost function state system state must be augmented with velocities...

2.2 Problem Formulation

The idea of uncertainty in planning has traditionally been considered in the domain of compliant robot, where the robot is allowed, or even required, to touch obstacles. The type of situations considered are relevant to problems involving manipulators and gripping, but are not of much use to vehicle motion planning where contact with an obstacle is to be avoided at all costs (with the exception of landing or perching). In most of the real world UAV planning problems, the issue of uncertainty in sensing and control is unavoidable.

The general problem of planning with uncertainty can be given as follows: Given a vehicle with uncertain position information, uncertain environment knowledge (e.g. obstacle locations) and having limited precision in tracking commands, find the best path to the goal.

3 Bridge Monitoring Flight Simulator

The objective of this section is to present a software simulation of missions to demonstrate innovative concepts for airborne systems. This flight simulator software is used to accurately represent the system behavior and to generate multiple numerical data useful for analysis of

results and algorithms improvement. Thus it is interesting to see a graphical representation of the displacement of the object simulated. Furthermore, the development of a standardized interface between software modeling and visualization environments can more easily use a variety of tools. The proposed simulator is based on Matlab/ Simulink® developed by MATHWORKS (www.mathworks.com). This flight simulator implements the numerical models: the model of the airship, the simulated model of the environment, the atmosphere and its evolution, the databases needed for scenarios, and finally tools for recording simulation data. This flight simulator software enables to present an easy and understandable way of monitoring mission specialist. The environment of virtual reality can be connected directly with all possible commands and all types of translation. To demonstrate the possibilities of such a tool, a specific example of bridge was chosen: the bridge of Cheviré (Nantes, France) which an overview is shown in Figure 1. It is a typical example where the monitoring work of a bridge can be automated. This bridge has three distinct parts to watch for:

- On deck,
- The bearings that support the deck: abutments at both ends and a dozen piles,
- The foundation that allows the transmission of forces of structure and land.

Monitoring of the deck, abutments and piers lends itself well to automation by monitoring with an airborne platform.



Figure 1 : Cheviré Brigde (Nantes, France)

In the simulated mission, the airship should fly from its point of rest, watch the first bridge pier in translation, monitor the span and continue monitoring the second cell, eventually returning to another point of rest. Several views are possible: on the road, on the airship, the whole trajectory and an observer far away. It is possible to move around the stage with arrows separated by 45 °: North, Northeast, East, South East, South, Southwest, West, North-West. Thus, it is possible to follow the airship in his motion.

To perform the simulation, the position in space (x, y, z) and orientation (Euler angles: roll ϕ, pitch θ, yaw ψ) of the vehicle are sent to the interface located in a Simulink model with an "s-function" Subsequently, the interface converts the data into the coordinate system of the viewer selected. This is a model which is used to calculate the data needed to display graphics. This model consists of several blocks from library sources. These blocks represent the mission planning, trajectory generation, the airship model and 3D visualization.

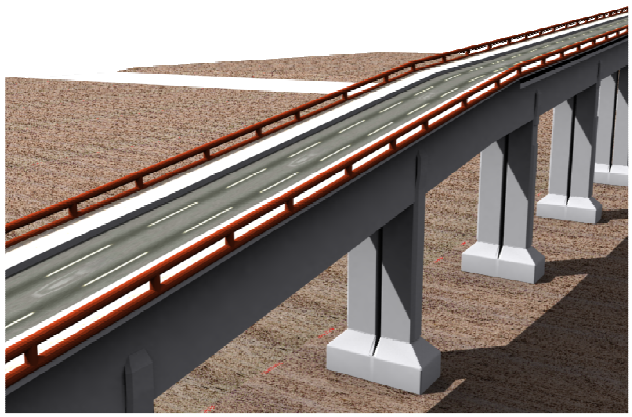


Figure 2 shows the path followed by the autonomous airship in nominal conditions.

4 Different Approaches in Route planning with uncertainty

Flight management of a vehicle in a dynamic partially known environment is a complex real time problem. The flight route can be preprogrammed at launch based on current weather conditions and forecast information, but once en route, changes to the flight plan have to be made in accordance to changes in the

weather. Even though forecast models have improved, weather and wind patterns continuously change and are still difficult to predict. Thus, the use of frequently updated weather information during the flight will allow adaptation and replanning based on current and near-term forecasts. The Weather forecast must become an important part of the flight management system, to detect signs of a slight change in wind velocity, monitor development and movement of clouds and recognize conditions suitable for the development of temperature inversions or of local wind phenomena.

4.1 Receding Horizon Control

The Receding Horizon Control (RHC) framework can also be employed to deal with uncertainties; the off-line pre-planned trajectory or approximate cost-to-go function accounts for global convergence based on known knowledge and the on-line RHC can used to negotiate with mid-flight uncertainties. It can be used to generate trajectories for a vehicle operating in an environment with atmospheric turbulence [8]. As uncertainty always exists due to numerous reasons, a way of taking care of it is to update the reference paths. The immediate measurements of the position and orientation are taken as initial conditions for the next reference trajectories generation.

The basic problem is to move the vehicle from the current configuration to the next desired configuration (waypoint or final destination). The current measured wind velocity is assumed to be constant between two updates of the reference trajectories. Because the environment is dynamic and trajectories are likely to change, a method is sought to generate paths quickly. A directed path segment is sought to join the current measured configuration $(x_i, y_i, z_i, \chi_i, \gamma_i)$ and the next way point configuration $(x_f, y_f, z_f, \chi_f, \gamma_f)$. At this point, the path smoothness is defined by the continuity of the path curvature and torsion. This method is presented in detail in [8].

4.2 Reactive Planning

The term 'reactive planning' refers in general to a broad class of algorithms that use only local knowledge of the obstacle field to plan the trajectory. Reactive algorithms are important in dealing with uncertainty, and run very quickly since no elaborate couplings are involved. In the case where a global obstacle map is not available and obstacle positions are known within a small radius, a reactive algorithm prevents last-minute collisions by stopping the vehicle when an obstacle is known to be in the trajectory, which has been planned by a different algorithm. This type of approach is important in many existing practical implementations in order to 'patch' an unsound algorithm to ensure that it is sound, as well as to deal with obstacle fields that may change suddenly. However, reactive planners, due to their inability to take the global planning problem into consideration, are seldom used as the sole trajectory generation process. In other words, if only the reactive planner is used, the vehicle may never find a trajectory that will lead to the goal, let alone an optimal one.

The Motion Description Language (MDL) and its extension (MDLe) can be used to define and design the reactive algorithms. In this framework, a sensor based interrupt (i.e. obstacle detected) will cause the vehicle to switch to another behavior. Fuzzy logic could also be used as a basis for a reactive algorithm, as well as visual flow based reactive planning.

There are numerous ongoing works on this topic, especially bio-inspired reactive planning algorithms. This can be explained due to the deficiencies of existing planning algorithms to involve sensory information in a principled way so that a complete planning framework results.

4.3 Decision Tree Formulation

A decision basis is a set of probable trajectories a moving object can take given a set of initial conditions. A decision tree is formed by concatenating a decision basis at the end of each trajectory of a previous decision basis. Decision tiers are sets of decision basis ordered along each path. Decision tier 0 is the root

decision basis, tier 1 is the set of decision bases attached to the ends of each trajectory emanating from the root basis, tier 2 is the next set and so on.

4.4 Markov Decision Process

Markov Decisional Process works as an entity that takes decisions through time. At any instant (τ) in time, an agent collects information w (τ) on the state 's (τ)' system. With the history of pairs of states and actions 'h(τ) = (s0, a0, a1,....., w0 (τ-1), w (τ))', The decision maker chooses the action a(τ). The choice of this action has two effects,

- the decision maker receives an immediate reward 'r(τ)',
- and the system transits to a new state 's(τ+1)' in the light of information received w(τ) collected on this state. This problem is looped as the agent desires.

These decisions (choice of actions) occur at different moments in time. There are two categories: When it is necessary to decide in continuous time, it is called continuous sequential decision making, and vice versa, sequential decision making is called discrete moments when the decision is discrete. Let N be the number of decisions to make. When N is bounded, this is a problem with finite horizon. Conversely, when N goes to infinity, this is an infinite horizon problem

Markov Decision Processes (MDP) seem to be the most adapted to our problem, because it incorporates a formal uncertainty and the mathematical resolution provides satisfactory results in the context of aerial robotics.

5 Application of the MDP to LTAR route planning

5.1 Introduction

When the actual wind vector field is not known exactly and may deviate significantly from the wind velocities estimated by the model, a technique to address this issue is to explicitly incorporate wind uncertainties into the planning algorithm.

To prepare the motion planning problem, the uncertainty in the wind field is modeled. The problem of reaching a particular goal location as a Markov Decision Process (MDP) is formulated using a discretized space approach. Solving the MDP provides a policy of what actuation option should be expected at any given location.

Because the wind velocity $W(r_i,t)$ is uncertain, the next state 's_j' may be considered as a random variable and a probability distribution can be constructed over all adjacent cells. Given these transition probabilities from all states, the motion planning problem is to select the actions that minimize time-to-goal. This problem is thus naturally posed as a Markov Decision Process *S*, *A*, *P*, *R* where *S* represents the set of possible states s_i, *A* is the set of actions available from each state, *P* gives the transition probabilities $P_a(s_i,s_j)$ from current state s_i to possible next states s_j' under action a, *R* defines the expected immediate reward for each transition and each transition a. In this problem, reward is negative travel time. .

5.2 Wind Uncertainties

Consider the problem of an aerial robot on the mission of bridge inspection explained earlier in this document. The challenge is to define a policy that is to say a series of decision rules to guide the robot to accomplish the mission assigned to it. This assumes a partial knowledge or overall environment in which the robot will evolve. So, moving from one place to another, avoiding obstacles, in this case, air turbulence requires the use of sensors of the robot. When these sensors are completely reliable, the robot perceives the complete state of the system. Unfortunately, in most cases, the sensors of the robots are only partially reliable, and the perception of the state of the system is said to be partial. So as a solution we can propose Markov decision processes. The sampling of the 3D environment (cell sampling of the 3D space) is needed to solve the problem. However, a compromise must be struck between

accuracy (related to the sampling frequency) and the risk of a combinatorial explosion that could be caused by selecting too many variables comprising the state vector. Available actions the system will be materialized by the various possibilities of the airship actuators: thrust angle α for example. The probability of transition between states T(s', a, s) will be chosen based on uncertainties related to the wind. Assuming that this law is Gaussian then it will have to choose the parameters (mean and standard deviation) as appropriate (For example, if it is determined that the winds are very low and the probability of deviation of the position of Desired arrival following the execution of an order, so we can assign a lower standard deviation, and vice versa). The reward function is unique for each problem or situation, and then it should be defined so that the process converges to the solution.

5.3 Proposed Approach

The objective of any mobile robot is to reach a target position from a certain position, also referred as pose when including orientations Cartesian positions. So, at any time, the configuration of a mobile robot is defined by six variables x, y, z, θ, φ, ψ representing the Cartesian position relative to a fixed axis and its orientation relative to a moving repair attached to robot. To reach a given goal, a robot calculates plans of motion (ie a sequence of actions that will lead him to this purpose) and executes them. Markov decision processes are very well suited to solve such problems. Their main advantage is that they offer a powerful mathematical framework for dealing with uncertainties in the information or disturbances. The problem of trajectory generation for a robot responsible for bridge monitoring is a three-dimensional problem. To carry out this work, we chose first to study the trajectory generation by Markov Decision Processes for a mobile robot moving in a 2D environment, and then generalize the problem of higher dimension. In both cases the problem formulation remains the same. As the robot is still at the design stage (Laboratoire Central des Ponts et Chaussées), so we do not model dynamics. Then we will assume for simplicity that the robot can be represented by a geometric point. To distinguish the states, we chose to sample the space environment of the robot in order to minimize the number of states considered. And assume that the control system is characterized by a very good tracking error between two instants of decision of the planning system. The air disturbances will be considered as obstacles to be avoided, they will be represented by spherical geometries during the simulation. For the simulation of the algorithm we have chosen the Matlab M-file and V-Realm Builder to design the 3D simulation environment.

5.4 Mathematical definitions

Our approach uses a simplified version of the decomposition into cells to determine the different states of Markov Decision Processes. The decomposition is also used to determine the actions that integrate the best kinematic constraints of a mobile robot.

5.4.1 Definition of states S

Markov Decision Process is defined by a quadruplet <S, A, T, R>. The set of states S is the first step in any formulation of a process. The decomposition of the environment in cells produces a finite number of rectangles in the case of 2D and 3D cubes in the case of 3D space. However, it is necessary to take certain considerations as the size of an elementary cell. It must be large enough to enclose the robot and not too small to avoid a large number of states for obvious reasons of simplicity of the problem. To fully define the possible states orientation is also considered. All theoretical angles possible [-π, π] for each axis is sampled to obtain a global state of the form <c,o>:

c: is the position of the center of the rectangle or cube. The dimension of c is either three or two.

o: orientation is obtained by sampling the set [-π, π] for each axis of rotation considered. In our case, we felt it would be reasonable to subdivide the whole so

as to have eight directions around each axis.

5.4.2 Definition of actions

It is often customary to take action in such classic problems as "initial rotation of x degrees, go straight, final rotation of y degrees". In the 2D case, it often limits the number of possible actions in each state: they represent the angles through which one could reach an adjacent state (neighbor cells after space sampling). Indeed, as shown in the following figure: at a given position, a robot operating in a 2D environment can only reach eight adjacent cells. This has the effect of considerably reducing the complexity of the problem and to assimilate the action to be taken at an angle of direction.

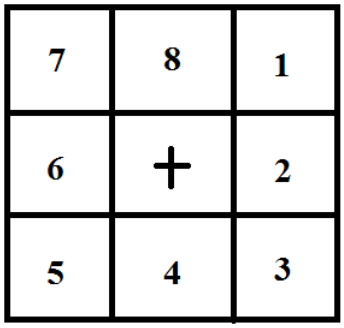


Figure 3: Examples of actions taken for the transition between two states.

The same justification can be extended to define all the actions in the 3D case. Except that the number of states reachable from any position is 26 without considering the orientations. At first, we can ignore the orientations and take into account only the Cartesian coordinates of the robot.

5.4.3 Transition function T

In the Markov decision processes, the transition function defines the effect of uncertainty when performing a given action. The goal is to model a function of uncertainty *I* configuration after execution of an action. We can choose *I* as a Gaussian whose parameters provide information about the intended effect of applying a policy decision. It can also be an external disturbance such as wind modeled by a probability function similar.

5.4.4 Reward function R(**s)**

The reward function is an application that associates to each state or action causing the system in a given state to a real number encrypting the utility's system to be in a situation (state). We can arbitrarily choose to assign to states of positive real to highlight the advantage of having the state in question. Thus, a state award by a negative number means that the system is in dangerous situations. Intuitively, we can assign to the state goal (desired final position of the robot) the value + Max. For the states located in disturbed areas, we chose to assign a value opposite that is to say, - Max. For other states, it is to say those who do not belong to the set of dangerous states and do not represent the state goal; we felt it important to give them a reward inversely proportional to the Euclidean distance which separates them from the state goal. This has the effect of preferred states whose distance from the target is lower. Finally, to avoid deadlock situations: when the robot is blocked between two neighboring states to infinitely. We decided to assign a value for penalizing states already visited (-val). In short, the reward function is defined as follows:

$$R(s) = \begin{cases} -Max, & If\ s\ belongs\ to\ a\ disturbed\ area \\ +Max, & If\ s\ is\ the\ goal\ state \\ 1/dist(s,sf) & \\ -val & If\ s\ is\ already\ visited \end{cases}$$

It is also important to note that disadvantaging states visited should preferably be chosen appropriately like *Max* and *val* (we propose to fix them empirically). We suggest to assign a negative number greater in absolute value to the largest Euclidean distance between the starting point and the end point to prevent the robot returns repeatedly to the points already visited.

5.4.5 Optimization criterion

The most appropriate for this kind of control is the finite criterion. It is assumed that

the agent should monitor the system in a finite number of steps N. The final criterion leads us to construct a value function that assigns to each state the expectation of the sum of N future rewards. The sum of the next N rewards is the criterion to be optimized.

$$E[r_0, r_1, \ldots r_{N-1} | s_0]$$

5.5 Modeling wind uncertainty

The wind is decomposed into direction and magnitude components, denoted respectively by θ_i and W_i. Independent probability distributions are assigned to each. The expected value of each component is equal to the field model value given by $\overline{W}_i(r,t)$. To model the uncertainty in θ_i, a Von Mises distribution is employed

$$f_{vm}\left(\theta_i, \overline{\theta}_i\right) = \frac{\exp\left(k\cos\left(\theta_i - \overline{\theta}_i\right)\right)}{2\pi I_0(k)}$$

where $\overline{\theta}_i$ = \angle $\overline{W}_i(r,t)$ is the mean, k is a concentration parameter and $I_0(k)$ is the modified Bessel function of the first kind of order 0. The wind magnitude W_i uncertainty is modeled as Gaussian with $\overline{W}_i(r,t) = \left|W_i(r,t)\right|$. The standard deviation is set as proportional to the magnitude, $\sigma_i = \rho W_i$ (ρ is chosen by the user).

5.5.1 Transition Probabilities}

Next, the transition probabilities $P_a(s_i,s'_j)$ are defined that govern state s'_j is entered after executing each action a from each state s_i. If there is no vehicle actuation, the wind direction wholly determines the next state's position. In that case, the probability of the wind forcing the vehicle from state s_i to an adjacent state s_j is where θ_{ij} is the smaller bordering angle of the 1/8 circular sector pointing from s_i to s_j.

A Monte Carlo method is used to determine the probabilities P and expected rewards R from resultant vector for each action, described as follows. First a set θ_0 of N sample points are drawn from the Von Mises distribution $f_{vm}\left(\theta_i, \overline{\theta}_i\right)$ and a set W_0 of N samples from the standard normal distribution (this step is required only once, whereas the following steps must be done by iterating for each s_i). Second these samples are adjusted for the state s_i.

6 Test and Results

6.1 2D Simulation Results

Simulations of the algorithm are performed for configurations ranging from simple cases (only one obstacle to overcome) to configurations in the presence of many obstacles of different sizes (see figure 4).

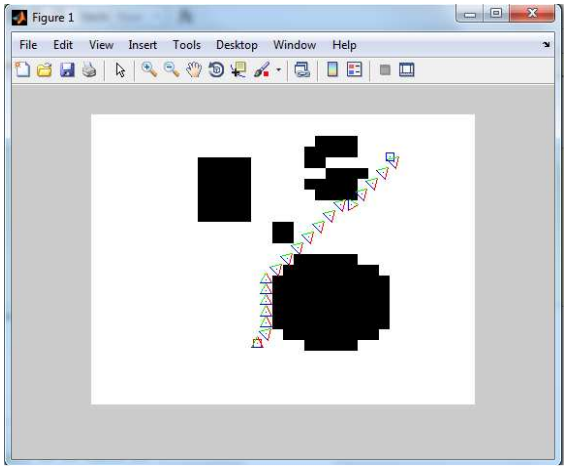


Figure 4: 2D path of the mobile robot

6.2 3D Simulation Results

The simulation of the algorithm in 3D is necessary because it brings us closer to the real case application of bridge monitoring. All components of the virtual world are built to simulate the program using V-Realm Builder. Thus, we built a bridge, a platform to illustrate soil. The robot and the point reached will be represented by squares. One is blue (robot), the second in red (end point). Areas of air disturbances that can be considered as obstacles will be illustrated by spheres of different colors. All these information are stored in a VRML standard file. Here is an overview of the virtual world on which the program will be applied to generate the 3D trajectory.

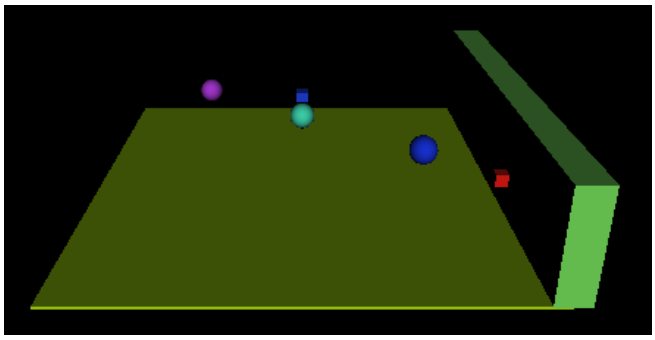


Figure 5: Overview of elements used for the 3D simulation program

The square on the far right position materializes the target point; the robot, on the far left of the screen is represented also by a blue cube.

On its way, it must avoid the spheres.

Initialization data of the algorithm are initially in the VRML file. Translation properties of each node (element of the virtual world) should be incorporated into the Matlab program. This approach has the advantage of separating the visual appearance of manipulation by Matlab code as the VRML file is written separately. The program starts with building the quadruplet <S,A,T,R>, except that the set of states includes all possible configurations for three-dimensional spaces. As for the set A, it consists of 26 elementary actions corresponding to the 26 neighboring cells for each state considered. Indeed, if any cell is selected, it is then close to 26 of his fellows. Actions A available for each state (position) then determine the next waypoint that will rally the robot that satisfies the most of the optimization criterion. When the program runs, it first opens the VRML file to load all elements and their properties and to manipulate through Matlab functions. Thus, the position of each element can be calculated by a block of code and then modified using an assignment of property translation of the node corresponding to the selected item. It is the same with the other properties (rotation, radius, ...).

Simulation results for the 3D case

To simulate the algorithm we chose to design a virtual world in which the robot represented by a blue cube has to reach a position represented by a red cube. Three obstacles on the path symbolize turbulence areas that the robot must avoid. Here is a sequence of images taken at regular intervals of the simulation.

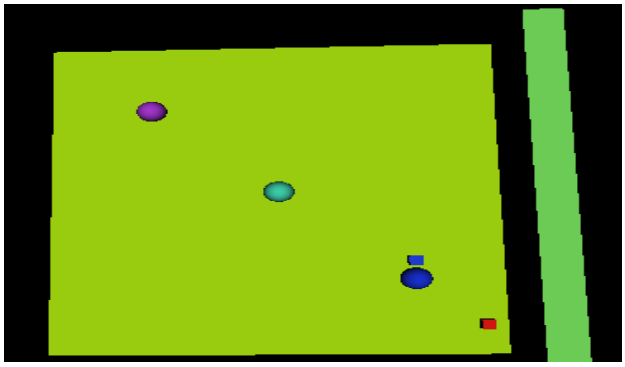


Figure 6 : Final configuration

Discussion

The proposed algorithm for trajectory generation in 3D has been tested successfully in situations where the robot was around obstacles (spheres). This allows us to validate the proposed method and generalize the 2D method. Note that for the 3D case, it is imperative to choose a horizon limited decision-making because of the exponential accumulation of the decisions. For a horizon $\tau = 1$, the number of possible decisions is already 26. Thus for a horizon $\tau = 2$, this number is 26^2 or 676 sequences of possible decisions. In our approach we have introduced a new way to construct the reward function that incorporates dangerous states, advantaged states and the concept of Euclidian convergence to the goal.

7 Conclusion

Aircraft guidance in the presence of winds and uncertainty presents several challenges. The proposed solution to uncertainty is a Markov Decision process considering the uncertainty in wind and turbulence.

In flight, the planned paths are used as guidance trajectories with a feedback controller being used to drive the state of the vehicle to the planned trajectory in the face of disturbances model uncertainties and approximation error.

References

[1] Bestaoui, Y. *Modeling and Trajectory Generation of Lighter Than Air Robot* , Robot Motion Control 2007, K. Kozlowski editor, Springer LNCIS 360, pp. 3-28

[2] Bestaoui, Y. *Bridge monitoring by a lighter than Air Robot*. AIAA Aerospace Sciences Meeting, Orlando, Floride, USA, 2011.

[3] Lee, S., Bang, H. *3D ascent trajectory optimization for stratospheric airships platforms in the jet stream* AIAA J. of Guidance, Control & Dynamics, vol. 30, 1341-1352, 2007

[4] Schmidt, D. *Modeling and near space keeping in control of a large high altitude airship* AIAA J. of Guidance, Control & Dynamics, vol. 30, 540-547, 2007

[5] Mueller, J. B., Zhao, Y. *Optimal ascent trajectories for stratospheric airships using wind energy* AIAA J. of Guidance, Control & Dynamics, vol. 32, 1341-1352 , 2009

[6] Jiang, Z., Ordonez, R. *Robust approach and landing trajectory generation for reusable launch vehicle in wind* 17th IEEE Int. Conf. on Control Applications, San Antonio, TX, pp. 930-935, 2008

[7] Nelson, R., Barber, B. *Vector field path following for miniature air vehicle* IEEE Trans. On Robotics, vol. 23, pp519-529, 2007

[8] Bestaoui, Y. *Mission plan under uncertainty for an autonomous aircraft* J. of Aerospace Engineering, vol. 24, pp. 1297-1307, 2010.

[9] Blackmore, L., Ono, M., Bektassov, A., Williams, B. *A probabilistic particle control approximation of chance constrained stochastic predictive control* IEEE Transactions on Robotics, vol. 26, pp. 502-517, 2010

[10] Calafiore, G., Campi, G. *The scenario approach to robust control design* IEEE Transactions on Automatic Control, vol. 51, pp. 742-753, 2006

[11] Teichteil-Königsbuch, F., Fabiani. P. *Influence de la structure de modélisation en planification probabiliste*. ONERA-CERT rep. 31055, France.

[12] Theodorakopoulos P. *On autonomous target tracking for UAVs.* Laboratoire d'Analyse et d'Architecture des Systèmes CNRS, rep. 31077 Toulouse, France. 2009.

[13] Burlet, J., Aycard, O., Fraichard, T. *Robust motion planning using Markov Decision Processes and Quadtree Decomposition*. Inria Rhônes-Alpes & Gravirv 38334 St Ismier, France. 2004.

[14] Igarashi,H, Loi, K. *Path-planning and navigation of a mobile robot as a discrete optimisation problems*. Art Life Robotics. 2001.

Acknowledgments: The authors gratefully acknowledge the financial support of the 'Direction de la Recherche et de l'Innovation, Ministère de l'Ecologie, du Développement Durable, du Transport et du Logement', FRANCE, Grant 10 MGC S 003

CEAS 2011

Design Methods and Aerodynamics Configurations

Aerodynamic guidelines in the design and optimization of new regional turboprop aircraft

Fabrizio Nicolosi and Pierluigi Della Vecchia
Department of Aerospace Engineering (DIAS) – University of Naples "Federico II"
Via Claudio 21, 80125 Naples, Italy
E-mail: fabrnico@unina.it, pierluigi.dellavecchia@unina.it
Ph: +39 0817683583 Fax : +39 081 624609

Keywords: *Aircraft design, aerodynamic analysis and optimization, regional turboprop.*

Abstract

This paper aims to provide some guidelines in the aerodynamic design and optimization of future regional turboprop aircraft with about 90 passengers. Currently there are no configurations on the market of this type, thus a typical 70 passengers turboprop aircraft is taken as reference starting point. The most critical areas in terms of aerodynamic behaviour are highlighted and an automatic procedure manageable through MATLab® is described: this interface allows to import and modify geometries using interpolating curves and surfaces via NURBS. Within the optimization loop, each new geometry is analyzed trough a panel code solver until optimized shapes are found.
Wing-fuselage junction, undercarriage pod, cockpit and wing-tip device are investigated. Design of the winglet is presented highlighting performance improvements during the entire mission profile.
Finally two different fuselage configurations are shown: the first with a 4-abreast fuselage arrangement and the second with 5-abreast, highlighting pros and cons of each configuration.

1 Introduction

Nowadays the increase in oil price, the huge growth of air transport traffic and the increasing attention to the aircraft environmental footprint led to considerable interest of specialists in new configurations of regional transport aircraft. As highlighted by the ATR Senior Vice President of Operations, Luigi Lombardi, during EWADE 2011 conference, the airlines will need about 3000 new turboprops in the next 20 years[1]. The 42% of the new turboprop deliveries expected to be 70 seats. The new 90+ seat segment is a strong percentage of the total, i.e. the 39% as shown in Fig. 1. Also Bombardier Commercial Aircraft Vice President Marketing, Philippe Poutissou, has expressed optimism about the future and he sees strong demand for this size aircraft in the market in the next two decades[2]. In particular, in its latest forecast of new aircraft deliveries in the 20-149 seat market segment over the next 20 years, Bombardier forecasts that 5800 aircraft will be delivered in the 60-99 seat segment and 6300 (only 500 more) in the 100-149 seat category.

For these reasons this work aims to provide some guidelines in the aerodynamic design of future regional turboprop aircraft with about 90 or more passengers.

Currently there are no configurations on the market of this type, so a typical 70 passengers turboprop aircraft is taken as reference starting point. This configuration has the wing in the high position for propeller clearance, T-tail and under wing engines as shown in Fig. 2.

An aerodynamic analysis of the reference geometry is performed trough a panel code in order to highlight the most critical areas in terms of aerodynamic behaviour.

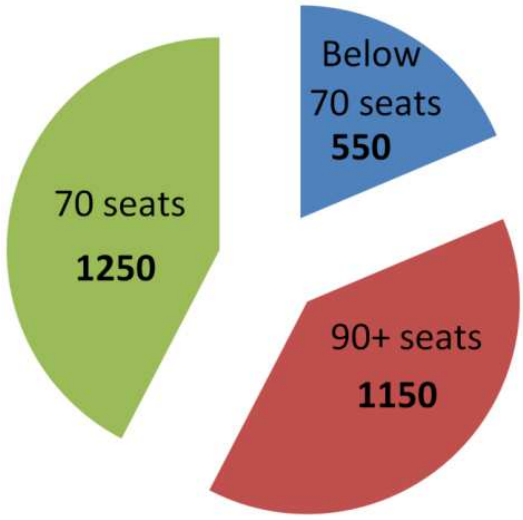


Fig. 1 Long term demand for Large Turboprop, ATR Forecast, March 2010[1]

Standard old configurations of regional turboprop can highlight some possible aerodynamic improvements that can be achieved through a careful design of some aircraft components.

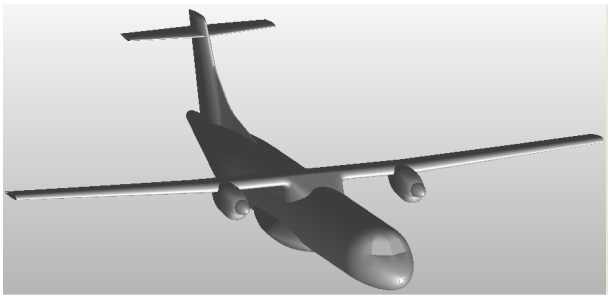


Fig. 2 Typical regional turboprop aircraft

Particular attention is placed on three components of the airplane (see Fig. 3): the wing – fuselage junction (named Karman), the landing gear pod (named Fairing) and the cockpit (named Fuselage nose).

Once identified these critical components, an automatic procedure manageable trough MATLab® allows to import and modify these geometries using interpolating curves and surfaces, NURBS[3]. The non-uniform rational B-spline (NURBS), are extensively used in CAD application. According to Piegl[3], NURBS curves are vector-valued piecewise rational polynomial functions (see [4][5] for further reference). They are used within this work to calculate an approximation of given surface grid points which represent the original shape. This is a powerful technique to change the grid points position and therefore the shape of the objects by adjusting the control points as outlined by Becker and Schäfer[6]. In contrast to the amount of grid points the amount of control points is rather small.

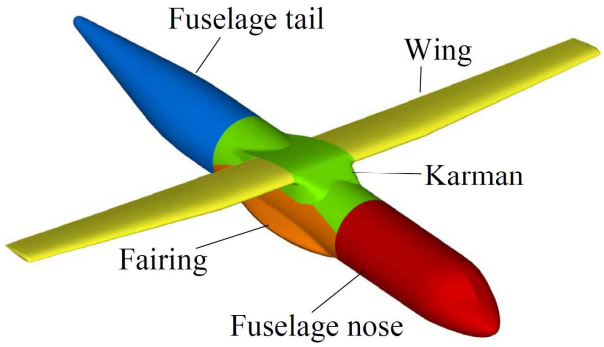


Fig. 3 Wing body components of the reference geometry

The strategy of optimization here presented is based on modifying original geometry (using the NURBS) until new shapes satisfy the objective function minimization (i.e. drag coefficient) via steepest gradient method.

Another major focus is the wing tip design, mainly addressed to the performance improvements (in particular rate of climb with One Engine Inoperative and OEI ceiling). The influence of winglet on every aircraft aerodynamic derivatives, including directional stability and lateral stability (dihedral effect) has been carefully considered.

In the analysis of possible future configurations two different solutions of a new regional turboprop aircraft with 90 passengers are proposed: the first with a 4-abreast fuselage arrangement and the second with 5-abreast. The

reason to propose two different fuselage geometries is mainly dictated by cabin comfort issue. The analysis of the two different solutions has been carried out in terms of aerodynamics (skin friction and form drag) and the effect on aircraft stability (also on tail planes have been re-designed).

2 Strategy and Optimization Process

The automated optimization process is here presented. Starting from the reference wing-body geometry shown in Fig. 3, the aerodynamic behaviour of some components is investigated to know if and how they can be optimized in terms of drag reduction. Once defined the components to optimize, the original shape is reconstructed via NURBS approximation. Then a number of control points (usually a pair), which can vary within a predetermined space (constrained optimization), are selected. The new surface geometry is then built according to the control points position and then analyzed with the panel code solver until the convergence criterion is satisfied.
At the end of the process an iges file (.igs), suitable for CAD processing is automatically generated. A flowchart of the process is shown in Fig. 4. All the phases are automatically managed in MATLab® environment.

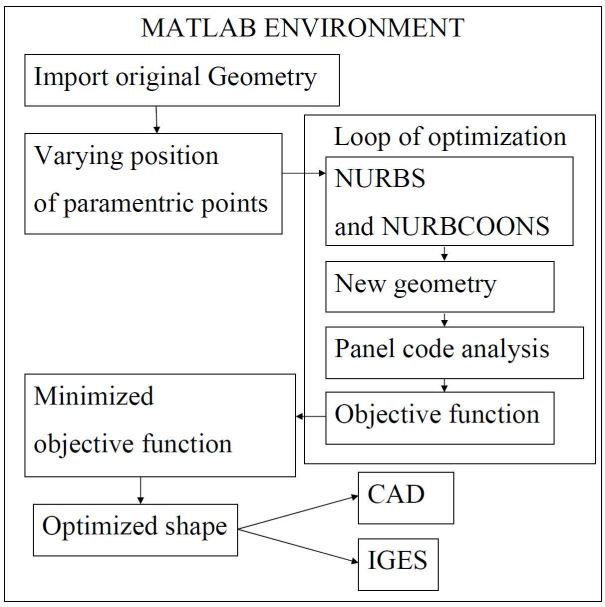


Fig. 4 Flowchart of the optimization process

2.1 Reference Geometry Analysis

A typical turboprop aircraft is considered assuming characteristics showed in Table 1, Table 2 and Table 3. A typical large turbo-propeller engine model variable with the altitude has been assumed. The parabolic drag approximation has been used. The zero lift drag coefficient has been calculated using the equivalent flat plate method plus several effects according to the methodologies proposed by Roskam[7] and Raymer[8]. These data have been used to evaluate aircraft performance.

Table 1 Geometrical characteristics

W_{TO}	b	S	L_f	$D_{MAX,f}$
23000 kg	27 m	61 m^2	27 m	2.5 m

Table 2 Characteristics to estimate drag polar

AR	e	C_{D0}
12	0.83	0.0306

Table 3 Performance characteristics

SBP	N° of engines	η_p	Cruise altitude
2750 hp	2	0.85	20000 ft

A preliminary aerodynamic analysis is necessary to define zones on the aircraft where the flow behaviour negatively affects total drag coefficient because of quasi-stagnations and separations.

The aerodynamic analysis has been performed trough a fast and reliable panel code solver available at DIAS (Department of Aerospace Engineering). This software allows the calculation of the nonlinear aerodynamic characteristics of arbitrary configurations in subsonic flow. Potential flow is analyzed with a subsonic panel method; the program is a surface singularity distribution based on Green's identity. Nonlinear effects of wake shape are treated in an iterative wake relaxation procedure; the effects of viscosity are treated in an iterative loop coupling potential flow and integral boundary layer calculations. The

compressibility correction is based on Prandtl-Glauert rule.

Table 4 Reference aerodynamic analysis conditions

Condition	α	β	M	Re
Cruise	0°	0°	0.43	$11.5 \cdot 10^6$
Climb	6°	0°	0.3	$8.0 \cdot 10^6$

In order to reduce computational time the analysis has been carried out on the wing body configuration of the aircraft. The geometry has been properly meshed and covered with about 12000 body panels subdivided into 5 components such as Fuselage nose, Fuselage tail, Karman, Fairing and Wing (Fig. 3) with 3200 wake panels. The aerodynamic analysis has been performed at two reference flight condition, which are typical cruise and climb conditions(see Table 4).

The main aerodynamic results in terms of pressure coefficient and skin friction coefficient are here discussed, emphasizing more critical issues.

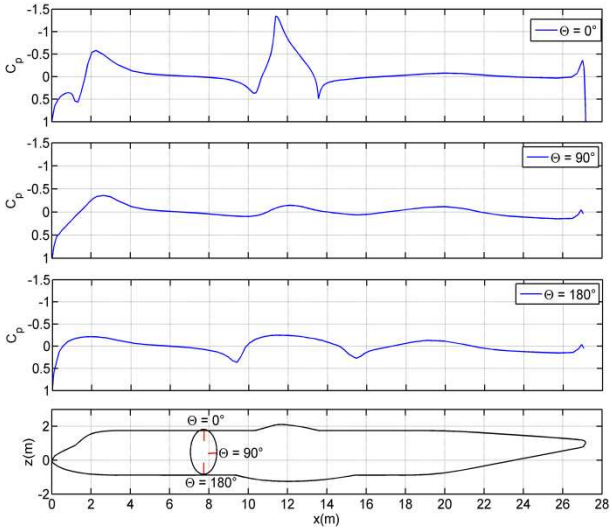


Fig. 5 Fuselage streamlines pressure distribution, Cruise

The pressure distribution over the aircraft fuselage shows that a locally convex curvature causes the pressure to decrease and a locally concave curvature will lead to an increase in pressure especially along the upper (Θ = 0°) and lower (Θ = 180°) streamlines as shown in Fig. 5. A large suction peak over the crest of the Karman component (C_p = -1.4) occurred (Fig. 5 and Fig. 6). Moreover a locally strong increase in the pressure is found where the local radius of concave curvature is very small, such as on the nose, on the front/rear Karman zone and on the Fairing component (see Fig. 5 and Fig. 6). As outlined by Schlichting[9] and Obert[10] these rapid flow variation should be controlled in the design phase. Geometry discontinuity (as those between Karman-fuselage and Fairing-fuselage) should be avoided or reduced. A more regular pressure distribution will result in lower drag than one with large variation[10].

This behaviour not too suitable in terms of pressure distribution is also reflected in the skin friction coefficient distribution, as shown in Fig. 7 and Fig. 8. The viscous analysis highlights a critical turbulence separation in the rear zone of the Karman component as shown in Fig. 7. This behaviour is clearly due to the entire boundary layer development, from the nose until the separation zone along the upper fuselage streamlines (see Θ = 0° in Fig. 8). The middle and lower streamlines (Θ = 90° and Θ = 180° in Fig. 8) show a normal development until a separation on the rear fuselage cone, mainly due to the upsweep angle.

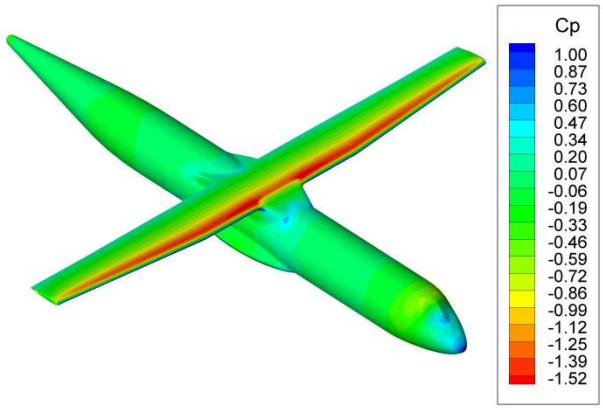


Fig. 6 Wing-body pressure distribution, Cruise

As it can be seen in Fig. 5 and Fig. 8, the flow on the cockpit after the stagnation point, suddenly accelerates (fuselage transition point has been fixed at x = 0.5 m from the aircraft nose). The flat windshield with protruding frame causes several vortexes formation and a smoother transition between the cockpit and the windshield should be suitable, as also

highlighted by Roskam[7], Obert[10] and de Mattos et Alii[11].

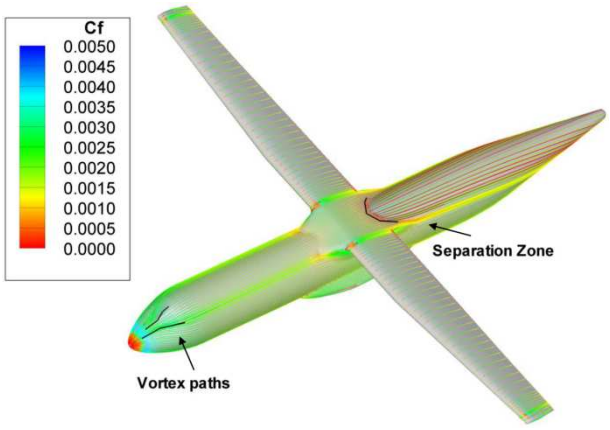


Fig. 7 Fuselage skin friction distribution, Cruise

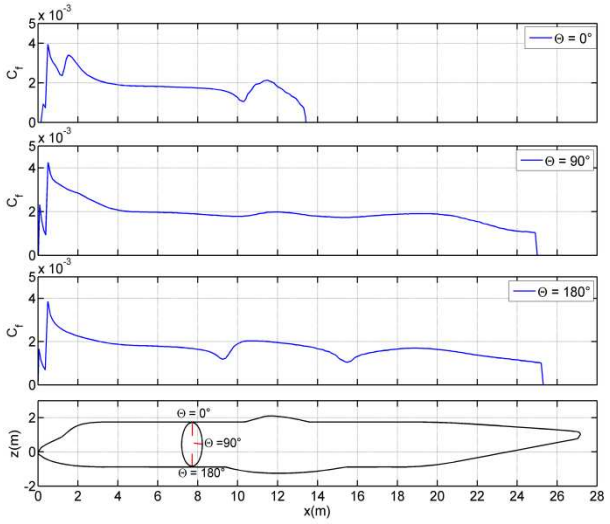


Fig. 8 Fuselage streamlines skin friction distribution, Cruise

2.2 NURBS reconstruction and objective function

The shape optimization is performed using a special set of curves and surfaces known as NURBS (Non Uniform Rational B-Splines). These are modeled with a number of control points, which are given proper weights and moved in space during the optimization loops.

The strategy of optimization consists in defining a component of the aircraft to be optimized (such as Karman). By an appropriate input file the user can define what portion should be optimized by defining the boundary curves shown in red in Fig. 9. Along these boundary curves the tangential direction is preserved thanks to several control points C, as represented in Fig. 9. The NURBS control points A and B allow the geometry variation in height and in spanwise respectively. The user has also to set the variation range of the A and B control points (Fig. 9).

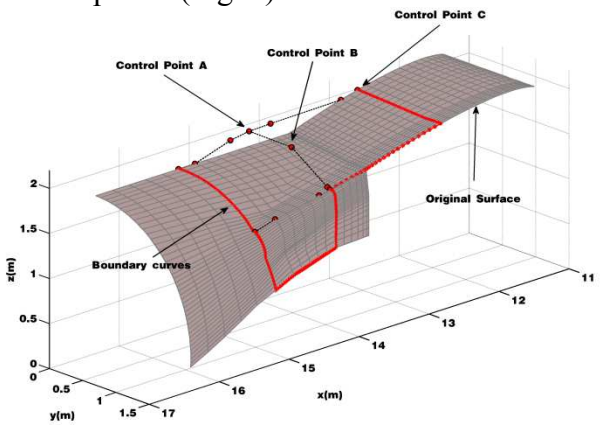


Fig. 9 Nurbs curves and control points

According to the control points position and boundary curves, the NURBS curves are built[3][4]. Subsequently the NURBS modified surface can be carried out through these curves, as represented in Fig. 10. The new surfaces are put on the reference wing-body geometry; finally the new configuration is analyzed trough the panel code solver. This process is iterated until the optimized geometry is found.

The drag coefficient has been assumed as objective function; it has been computed with the Young formula[12] which is based on integral quantities of the boundary layer, evaluated at the body's trailing edge or separation zone (see Eq. 1).

$$C_D(s) = \frac{2\pi}{S_{ref}}\left(r_T(s)\cdot\theta(s)\right)U_{e,T}(s)^{\left(\frac{H_T(s)+5}{2}\right)} \quad (1)$$

Where C_D is the drag coefficient, S_{ref} is a reference area (usually the maximum frontal area, in this case the wing surface), r is the radius, θ is the b.l. momentum thickness, U_e is the inviscid external velocity and H is the b.l. shape factor. The subscript T indicates that all

quantities are to be evaluated at the body tail (x/L=1.0), or at the separation point.

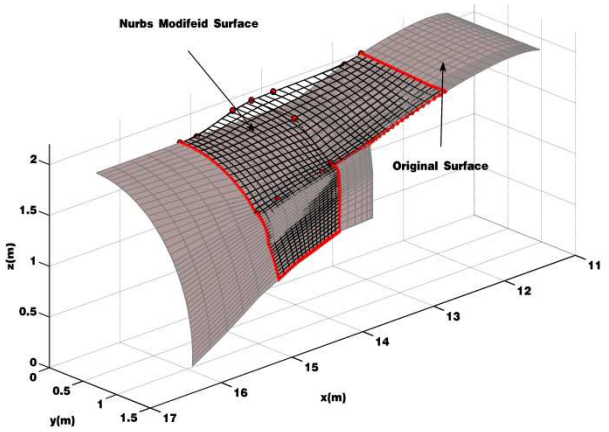


Fig. 10 Nurbs surfaces

This is considered an approximate method to perform viscous calculations for general fuselages (non axisymmetric bodies). As reported in [13] and [14], comparisons with experiments have shown reasonably good agreement with experimental drag coefficients not only in case of axisymmetric bodies, but also for some fuselages tested at Delft[15]. Moreover the analysis performed in this work aims to provide an improvement of a reference geometry and it does not want to give the real drag coefficient of a geometry.

The viscous calculations along some streamlines of the body has been performed. In particular six streamlines (3 in the upper and 3 in the lower zone) have been taken as reference for the calculation of the drag coefficient. The radius is considered as the mean value of the fuselage height, at each x-station, for both upper and lower streamline. A drag coefficient can be obtained for both streamlines (Eq. (1)) leading to six values. The fuselage drag coefficient will be taken as the average of average upper and average lower values. It is clear that this represents an approximate method because the boundary layer on the fuselage is three-dimensional, and axisymmetric viscous calculations are not an accurate representation of a non-axisymmetric flow condition on a non-axisymmetric body. It is evident that the results are more and more inaccurate as the fuselage shape is far from an axisymmetric body.

A steepest gradient method is used to minimize the objective function.

3 Design and Optimization Results

In this section the main results obtained by applying the procedure illustrated in 2.2 to three aircraft components Karman, Fairing and Nose are applied. For brevity, only the Karman and Nose design will be illustrated. At the end, the design of the winglet will be introduced. Drag coefficient reduction and its effect on the aircraft performance will be put in evidence.

3.1 Karman Design

The optimization process applied to the Karman components (see Fig. 3) aims to avoid the separation zone shown in Fig. 7 and Fig. 8 through better airflow behaviour. It has been imposed as geometrical constraints that the minimum upper height should be equal to the maximum height coordinate of the reference geometry, and the tangential direction with the fuselage must be ensured. Comparisons between optimized geometry and baseline geometry are represented in Fig. 11, Fig. 12 and Fig. 13 . The abrupt geometrical discontinuities are avoided both on the front and the rear zone, thanks to a smother and slightly longer component. Moreover the pressure gradients are reduced and the separation in the rear part of the Karman is avoided as it can be seen in Fig. 15 and in Fig. 16. Total drag coefficient is reduced of about 7 counts, giving a maximum true airspeed improvement of 3.5 knots (see Table 5). The drag coefficient reduction grows with the aircraft angle of attack because of the separation zone reduction (see Table 5).

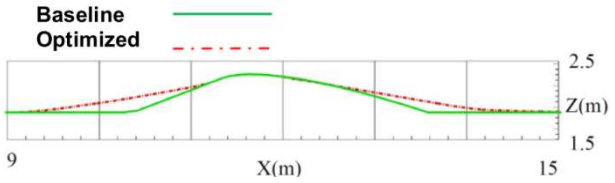


Fig. 11 Baseline and Optimized Karman geometry, symmetry plane y = 0 m

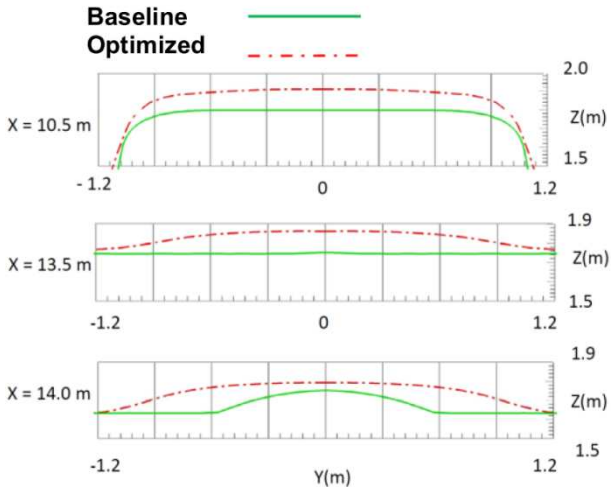


Fig. 12 Baseline and Optimized Karman geometry, slice at x = 10.5, 13.5, 14 m

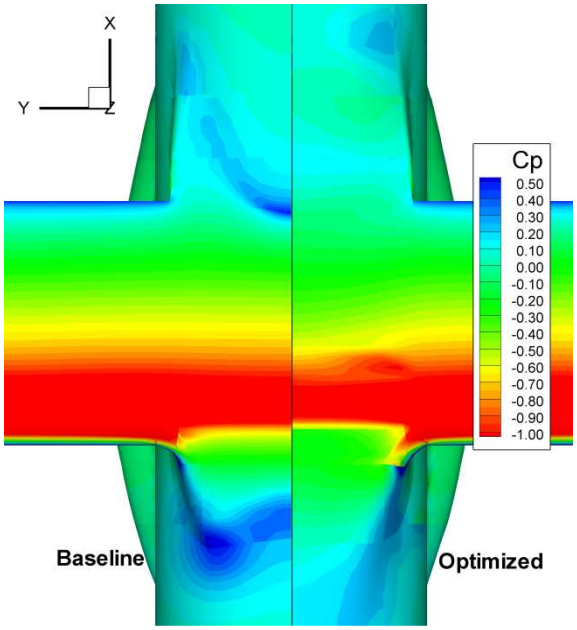


Fig. 13 Baseline and Optimized Karman pressure distribution

3.2 Nose Design

The Fuselage nose components (see Fig. 3) is modified varying its length, height and span(see Fig. 14). The fuselage constant cross section has been imposed as NURBS boundary curve (see Fig. 9). Transition has been always fixed at 0.5 m from the aircraft more forward "x" coordinate.

The best geometry is translated forward of about 0.5 m and downward of about 0.5 m respect to the reference one (see Fig. 14). This nose geometry allows to reduce forward compression on the upper streamline (Θ = 0° in Fig. 15) with a higher suction on the lower streamline (Θ = 180° in Fig. 15).

Total drag coefficient computed is reduced of 3.3 counts (see Table 5). It is evident that the optimized geometry could also lead to further improvement in laminar flow condition. Moreover, the new configuration accomplishes the regulations in term of pilot angle of visibility[16].

Since the optimized geometry of the nose is stretched the fuselage instability is 5% higher of the reference one.

Table 5 Optimized results, drag coefficient reduction, *(Reference C_D = 306 counts*, reference V_{MAX} = 261 kts)*

Results	ΔC_D Cruise	ΔC_D Climb	V_{MAX} 20kft
Components	**(counts)**	**(counts)**	**(kts)**
Karman	6.7	8.2	+3.4
Fairing	1.5	2.2	+1.1
Nose	3.3	3.5	+1.8
Total	11.5	13.9	+6.3
% difference respect to reference	-3.8%	-4.5%	+2.4%

* 1 drag counts is equal to 0.0001

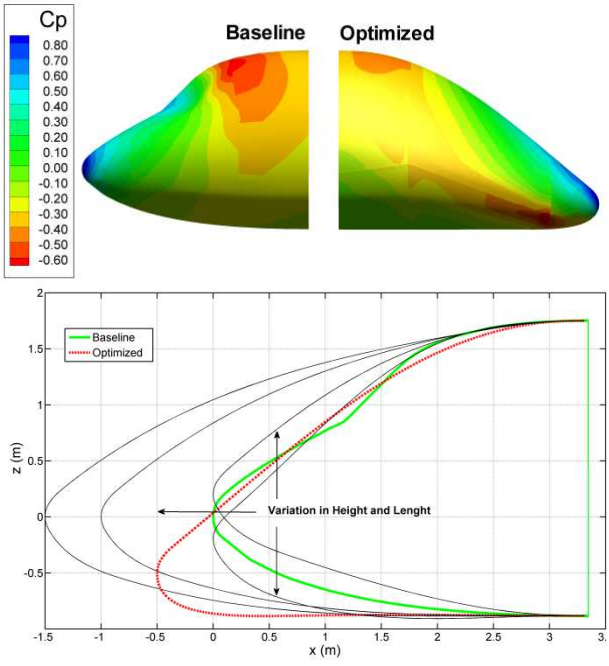


Fig. 14 Baseline and Optimized Fuselage Geometry; top: pressure distribution; bottom: midline sections

As it is shown in Table 5, the optimization process leads to a maximum true airspeed

improvement at 20 kft of about 6 knots (from 261 to 267).

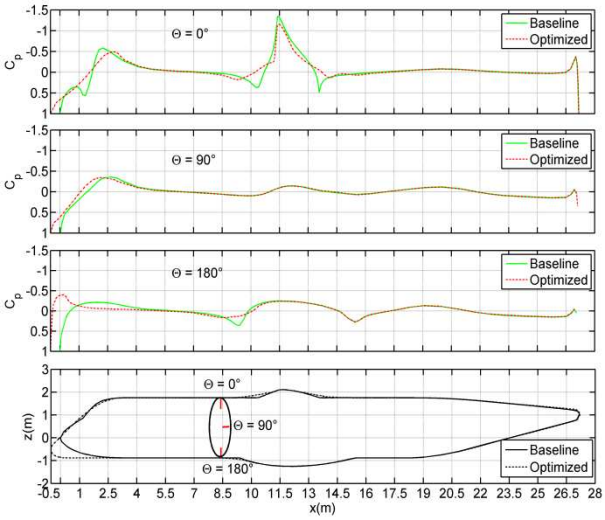


Fig. 15 Baseline and Optimized fuselage streamlines pressure distribution, Cruise

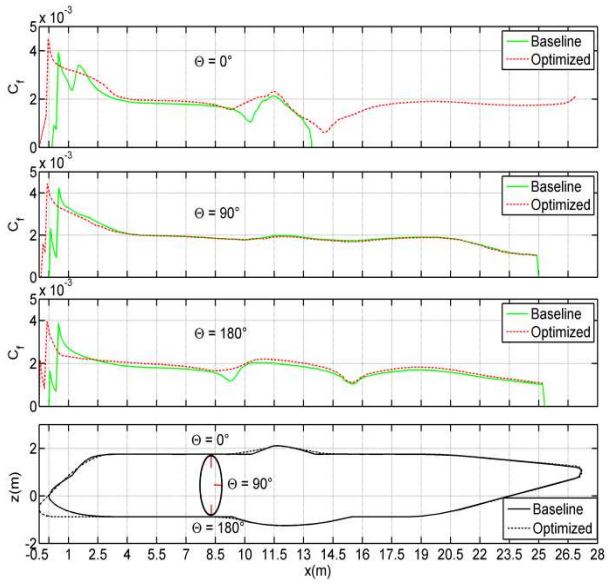


Fig. 16 Baseline and Optimized fuselage streamlines skin friction distribution, Cruise

3.3 Winglet Design

Winglet installation on transport aircraft is becoming a main aerodynamic topic design to improve performance and save fuel.

The authors have gained experience in winglet design and further analysis are under development([17][18][19]).

A MATLab® code, written at DIAS, allows to define winglet geometry by setting a reference wing geometry and several winglet design parameters (such as cant angle, toe angle, winglet height, etc.). Then an automated procedure performs aerodynamic calculation with the same panel code aforementioned (see section 2.1). Several analysis has been carried out in order to improve the wing induced drag factor, and the results on the optimized winglet are here presented, focusing the attention on the performances and on the aircraft static stability.

The optimized winglet shape could be defined as a semi-blended winglet with an height about the 10% of the wing semi span, a cant angle of about 80 degrees and a sweep angle of about 50 degrees.

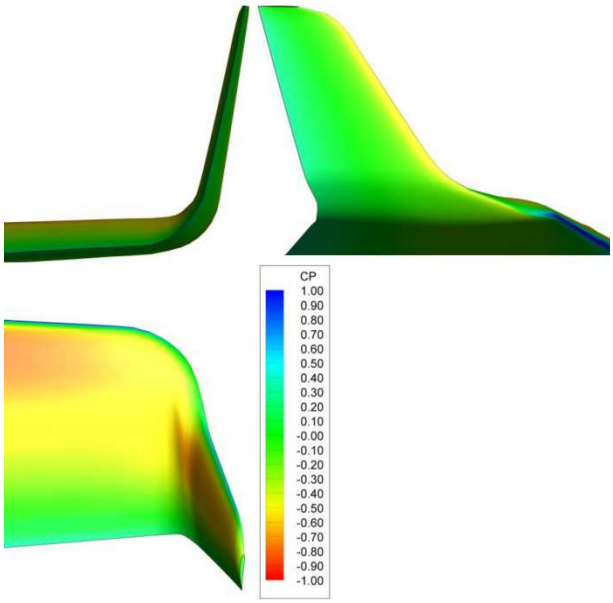


Fig. 17 Winglet pressure distribution, Cruise

The designed winglet, shown in Fig. 17, allows an improvement of the wing induced drag factor of about 11% with respect to a reference wing with the same aspect ratio. In Fig. 18, the difference between aircraft total drag coefficient without winglets and the total drag coefficient with winglets vs. lift coefficient is depicted; it is interesting to notice that the winglets give a reduction of the total aircraft drag coefficient in a wide range of aircraft lift coefficient. The major improvements due to winglets are in climb condition where the drag coefficient can be reduced until 30 drag counts. This result is mainly impacting on the climb performance of the aircraft which is improved especially in OEI condition, as it can be seen in

Table 6. Maximum true airspeed in cruise condition is improved of about 3 knots at maximum TO weight. FAR ceiling with winglet is enhanced of about 1000 ft and a fuel reduction of about 13 pounds during a typical 200 nautical miles mission has been obtained.

The influence of winglet on all aircraft aerodynamic derivatives, including directional stability and lateral stability (dihedral effect), has been carefully considered (see Table 6). In particular aircraft lateral stability is increased of 26% in accordance to [20][21].

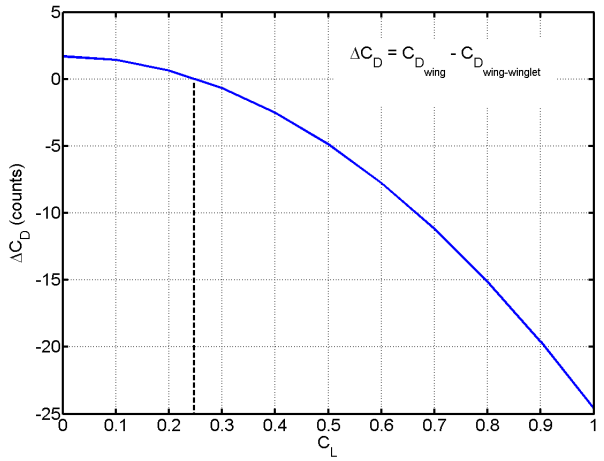


Fig. 18 Aircraft drag coefficient variation due to winglets (1 drag counts is equal to 0.0001)

Table 6 Complete Aircraft without and with winglets, main performance and aerodynamic derivatives

Aircraft	No Winglets	Winglets	%
C_{D0}	0.0306	0.0308	+0.7
e	0.83	0.92	+11.0
V_{MAX} (kts),TAS	261	264	+1.1
RC_{MAX}(AEO) s.l. (ft/min)	1425	1466	+2.9
RC_{MAX}(OEI) s.l. (ft/min)	331	382	**+15**
FAR ceiling AEO (ft)	24957	25997	+4.1
FAR ceiling OEI (ft)	11933	13181	**+10.5**
Fuel burns 200 nmi (lb)	1305.6	1292.0	-1.1
$C_{L\alpha}$ (1/deg)	0.0980	0.1038	+5.6
dε/dα	0.281	0.267	-5.2
$C_{l\beta}$(1/deg)	0.00360	0.00487	**+26**
$C_{n\beta}$(1/deg)	-0.00106	-0.00101	-4.9

4 New turboprop configurations

According to the guidelines dictated in paragraph 3, two different solutions for a new regional twin engine turboprop aircraft with 90 passengers are proposed as shown in Fig. 19: the first with a 5-abreast fuselage arrangement and the second with 4-abreast. To ensure the same aircraft stability and control, tail planes have been also designed for both configurations.

The reason to propose two different fuselage geometries is mainly dictated by cabin comfort issue. The 4-abreast configuration leads to better aerodynamics both for the fuselage geometry (longer) and tail planes, but is characterized by a relevant lower cabin comfort, especially for a 90 seats aircraft. Conversely, the 5-abreast configuration is slightly worse in terms of aerodynamic behaviour but it has more comfort in the cabin(Fig. 20), giving also the possibility to take on board standard cabin baggage in the lower fuselage section area.

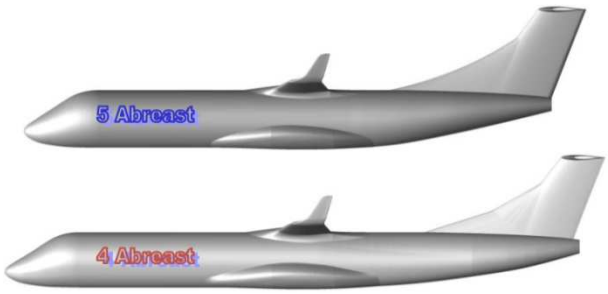


Fig. 19 Up: 5-abreast configuration; down:4-abreast configuration

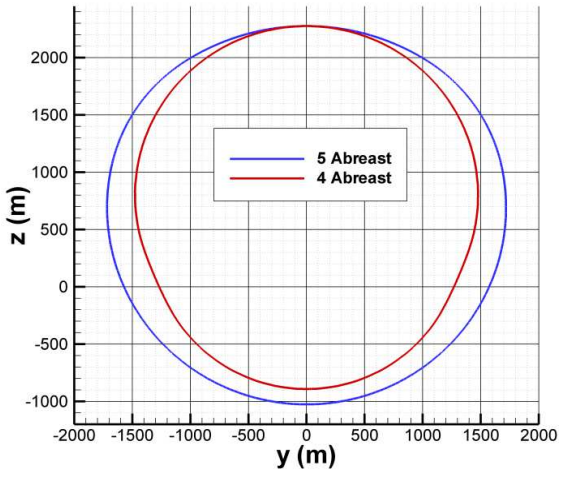


Fig. 20 4-abreast and 5-abreast fuselage cross section comparison, x = 10 meters

A performance comparison of these two configurations is summarized in Table 7. The aircraft zero lift drag coefficient (calculated according with methodologies proposed by Roskam[7] and Raymer[8]), has shown that the 4-abreast configuration has a reduction of about 7%, which correspond to a 3% higher maximum true airspeed. This difference is mainly due to the fuselage contribution (skin friction, upsweep and windshield effects), but it is also due to the tail planes (which have been designed in order to have the same stability and control requirements).

Table 7 4-abreast and 5-abreast comparison

Config.	l(m)	d(m)	CD_0	V_{MAX}	RC
4-abreast	32.6	3.2	-	-	-
5-abreast	31.2	3.3	+7%	-3.1%	-1.5%

5 Conclusion

Aerodynamic guidelines on some critical components (wing – fuselage junction, nose) in a regional turboprop has been proposed, showing a drag coefficient reduction of about 4-5%. Winglets adoption leads to performance improvements during the entire flight envelope and it gives non-negligible effects on the aerodynamic derivatives, mainly on the dihedral effect (+26%). A comparison between two new 90 passengers turboprop shows that a larger fuselage (5-abreast) can give an advantage in cabin comfort with a drag penalty.

References

[1] Lombardi, L., "ATR, Program review and market outlook", 10th EWADE Conference, Naples, 2011.

[2] Special Convention RAA News, "Regional Horizons," RAA Annual Convention, Milwaukee, Wisconsin, 26 May 2010.

[3] Piegl L., "On NURBS: A survey," Computer Graphics and Applications, IEEE, Vol. 11, No. 1, Jan 1991, pp. 55–71.

[4] Piegl, L. and Tiller, W., "Curve and surface constructions using rational B-splines," Computer-Aided Design, Vol. 19, No. 9, Nov. 1987, pp. 485–498.

[5] Rogers, D. F. and Adams, J. A., "*Mathematical elements for computer graphics,*" 1st ed., McGraw-Hill New York, 1976.

[6] Becker, G. and Schäfer, "An advanced NURBS fitting procedure for post-processing of grid-based shape optimizations," 49th AIAA Aerospace Sciences Meeting, January 4–7, 2011, Orlando, FL USA.

[7] Roskam, J., *"Airplane Design, Part VI: Preliminary Calculation of Aerodynamic, Thrust and Power Characteristics,"* 2nd ed., DARcoporation Lawrence Kansas, 2000, Chap 4.

[8] Raymer, D., P., *"Aircraft Design: A Conceptual Approach,"* 3rd ed., AIAA Education Series, Reston Virginia, 1999, Chap 12.

[9] Schlichting, H., Truckenbrodt, E., A., *"Aerodynamic of the Airplane,"* 1st ed., McGraw-Hill, USA, 1979, Chaps 5, 6.

[10] Obert, E., *"Aerodynamic Design of Trasport Aircraft",*1st ed., IOS Press TUDelft Amsterdam, 2009, Chapter 2.

[11] De Mattos, B., S., Papa, R., de Castro, L., C., "Considerations about Forward Fuselage Aerodynamic Design of a Transport Aircraft," *42nd AIAA Aerospace Sciences Meeting and Exhibit,* AIAA 2004-1241, Reno, Nevada, 5 - 8 January 2004.

[12] Young, A., D., "The Calculation of Total and Skin Friction Drag of Bodies of Revolution at Zero Incidence," ARC R&M 1874, April 1939.

[13] Boermans, L., M., M. and Nicolosi, F., "Sailplane fuselage and Wing-Fuselage junction design," XXV Ostiv Congress, S. Auban France, 3-11 July 1997.

[14] Coiro, D., P. and Nicolosi, F., "Design of Natural Laminar Flow Fuselages," *19th ICAS Conference*, Anaheim, California, 18-23 September 1994.

[15] Boerman, L., M., M. and Terleth, D., C., "Wind Tunnel tests of eight sailplane fuselage conbinations," Technical Soaring, Vol. VIII, No. 3, 1984.

[16] European Aviation Safety Agency, "Certification Specifications and Acceptable Means of Compliance for Large Aeroplanes CS-25," Amendment 11,4 July 2011.

[17] Pascale, L., Nicolosi, F., "Design and Aerodynamic analysis of a light twin-engine propeller aircraft," *26th ICAS Conference*, Anchorage, Alaska, 14-19 September 2008.

[18] Nicolosi, F., De Marco, A., Della Vecchia, P., "Flight tests of a twin-engine flight aircraft: performances, stability and parameter estimation," *XX AIDAA Congress*, Milano, Italy, June 29-July 3, 2009.

[19] Nicolosi, F., De Marco, A., Della Vecchia, P., "Flight Tests, Performances, and Flight Certification of a Twin-Engine Light Aircraft," *Journal of Aircraft,* Vol. 48, No. 1, 2011, pp 177-193.

[20] Jacobs, *P., F.,* "Effect of Winglets on a First-Generation Jet Transport Wing V, Stability

Characteristics of a Full-Span Wing With a Generalized Fuselage at High Subsonic Speeds," NASA, 1978.

[21] Jacobs, *P., F.,* Flechner, S., G., "The effect of winglets on the static aerodynamic stability characteristics of a representative second generation jet transport model," NASA, 1976.

Parametric Modelling of Aerodynamic Objects - The Key to Successful Design and Optimisation

W. Stalewski
Institute of Aviation
Al. Krakowska 110/114, 02-256
Warsaw, Poland,
Email: stal@ilot.edu.pl

Keywords: *parameterisation, parametric design and optimisation, multi-disciplinary optimisation*

Abstract

In the paper the methodology of parametric modelling of complex objects, especially within aerospace domain, has been presented. In principle, the main area of application of the methodology has been its usage as a parametric generator of geometry within the Numerical Optimization process. This assumption caused that the PARADES software, based on the presented methodology, differs significantly from conventional CAD systems. The main distinguishing features of this software are:

- *high level of compliancy with professional CAD systems and grid generators*
- *high quality of generated geometries, usually easy-readable by grid generators working in automatic (batch) mode*
- *implementation of both the classical parametric modelling based on the set of design parameters and alternative parametric modelling based on the morphing technique*
- *possibility of work both in interactive and in fully automatic mode*
- *easy adaptation to any numerical optimisation system, any operating system and any hardware*

1 Introduction

Within the aerospace engineering domain many typical problems concern an optimal designing of aerodynamic objects or their components. Usually, to solve such problems, the methodology of parametric design and optimisation is applied. General scheme of such a process is shown in Fig.1. The designer may both a human or a computer code. In the former case we are saying about Direct Design and Optimisation process, performed in interactive mode by experienced engineer. In the latter case the Numerical Optimisation process is performed in fully automatic mode by specialised software. The designer, human or software, uses parametric model of designed product to generate various variants of this product. Next, they evaluate physical properties of sequential variants usually using specialised codes. Based on this evaluation, the designer decides which variant is the best from point of view of assumed objectives and constraints.

Within the approach described above, one of the most important components of design system is the parametric model of the designed product. Generally, the parametric model is responsible for geometry generation. In fully parametric approach, the complete geometry is

unambiguously defined by certain set of variables, called Design Parameters. As a result, the design problem comes down to search of the set of Design Parameters defining the optimal geometry. Application of parametric model enables to introduce orderliness and automation into design and optimisation process, which may be clearly formulated and solved using mathematical tools.

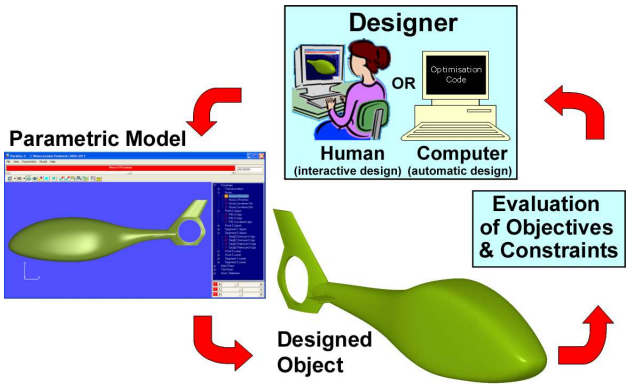


Fig.1 General scheme of process of parametric design and optimisation.

From a practical point of view, the parametric model is a computer program, intended to create a geometry based on the given set of Design Parameters. The effective and useful parametric model of aerodynamic object should fulfil some fundamental requirements:

- easy adaptation to different design and optimisation problems
- modelling of wide variety of complex and high quality geometries using relatively small number of Design Parameters
- regardless of current values of design parameters, the created geometry should always have the same topology
- possibility of work both in interactive and in fully automatic mode (batch mode)
- the software should be easy adapted to any operating system and hardware
- the software should be a relative simple code, easy to duplicate, which is very important in automatic optimisation, usually using parallel computing technique
- the exported geometry should comply with CAD standards and with professional grid generators

In the paper the methodology of parametric modelling, developed in Institute of Aviation (Warsaw, Poland), has been presented. Based on this methodology the software supporting a parametric design was worked out. The software fulfils all above requirements concerning parametric-modelling codes. The software is still developed, but it has already found application in many research and design works [1],[2],[3],[4]. Among other things, the software has been used in several European projects (HISAC, NACRE, CESAR) and commercial design works (EADS-CASA, Rolls Royce).

2 Methodology

For parametric modelling of complex objects, the methodology worked out formerly in Institute of Aviation was adapted, developed, and finally implemented as the PARADES software. Generally it was assumed, that the main area of application of the methodology would be its usage as a parametric generator of geometry within the Numerical Optimization process. This assumption was the main factor that determined the whole process of development of presented below the methodology and the software.

2.1 Fundamentals of methodology

The main idea of the methodology of parametric modelling of complex objects is shown in Fig.2. Usually the parameterised object is simply a shape of designed product (e.g. flying object or its component). In mathematical terms a shape may be treated as the function – parametric description of curve, surface or volume. In general, the presented methodology can be applied to parameterise any functions, not necessarily describing the geometry. For example it may be multi-dimensional function describing kinematics of given body or function describing distribution of certain physical quantities such as temperature or stress.

Generally, the parameterised object is modelled using a set of NURBS [5] (Non-Uniform Rational B-Splines) curves, surfaces or solids. These fundamental

components are created based on a set of Design Components: scalars, vectors and NURBS (curves, surfaces, solids). These Design Components are complementary input data necessary to create parameterised object. In particular they may define dimensions, shape, topology, internal structure, kinematics, etc. of the object.

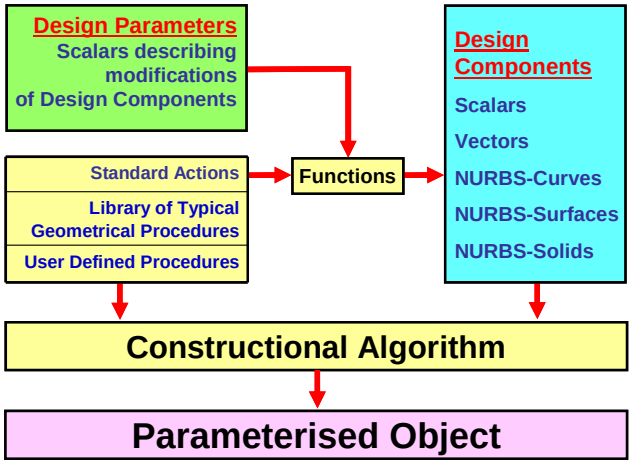


Fig.2 Functional diagram of the presented methodology of parametric modelling of complex objects.

The Constructional Algorithm describes how to create the parameterised objects using these data. The designer of given parametric model, creates the Constructional Algorithm using three types of procedures:

- Standard Actions – library of procedures modifying Design Components
- Library of Typical Geometrical Procedures
- User Defined Procedures

The Design Parameters describe modifications of Design Components. These modifications may be performed using Standard Actions and Functions defined by the designer. Generally, one Design Parameter may modify several Design Components, each according to another Function describing a speed and direction of the modification. The Standard Actions concern modifications of Design Components:

- changes of values of scalars and vectors
- affine transformations of selected subsets of control points of NURBS
- changes of weights of control points of NURBS
- global deformations of NURBS

In many cases the whole Constructional Algorithm may be defined using only the Standard Actions. In more sophisticated modelling, parts of Constructional Algorithm are attached as User Define Procedures. The designer writes them in C++ programming language. The user has full access to the Design Components and data defining final geometry of parameterised object. Additionally he may utilise Library of Typical Geometrical Procedures such as: interpolation, multi-section surface, generalised surface of revolution, trimming, etc.

In the presented methodology the Design Parameters do not directly influence the essential part of Constructional Algorithm. It means that the parameterised object may be built even without any defined Design Parameters. Such approach gives high level of freedom when given parametric model is finally defined. Depending on requirements, the number and the form of Design Parameters may be easily changed without any serious changes in the Constructional Algorithm besides changes of the Standard Actions.

2.2 Reduction of Design Parameters Set

From point of view of successful design and optimisation process the number of Design Parameters should be strongly limited. This way the dimension of search space is limited, which gives higher probability of finding optimal solution in limited period of time of the design process.

Let us assume, that the set of Design Components consists of N_s scalars, N_v vectors, N_n NURBS (curves, surfaces or solids) and that n-th NURBS is modelled by C_n control points. Then in presented methodology the number of degrees of freedom in modelling of given parameterised object is equal:

$$D_F = N_s + 3 \cdot N_v + \sum_{n=1}^{N_n} 4 \cdot C_n \qquad (1)$$

In practice, in most design problems this is a huge number. Therefore the Design Parameters

very rarely modify all Design Components separately. In properly defined parametric model either the Design Parameters modify only limited number of Design Components or the single Design Parameter modifies a large subsets of Design Components. In practise, the number of Design Parameters should be well-balanced taking into account specific requirements of given design and optimisation problem. Generally, in interactive mode it is allowed to define relatively high number of Design Parameters. In such a case, an experience of an engineer allows to control smoothness and feasibility of generated geometries. However when Numerical Optimisation is performed it is strongly recommended to limit the number of Design Parameters. It may be achieved by use of global deformation or morphing technique.

2.2.1 Global deformation

In global-deformation approach various shapes of NURBS surfaces or curves are obtained by smooth movement of their control points. Assumed types of deformations usually imitate behaviour of elastic materials. In presented methodology the following global deformations are implemented:

- bending (Fig.3)
- stretching
- skewing
- deformation of surface forced by deformation of its edge (Fig.4)

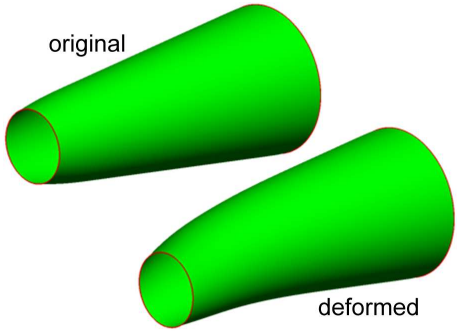


Fig.3 Bending of the NURBS-surface.

The advantage of using the global deformation is that even very complex shapes may be parameterised by a small number of Design Parameters, while maintaining smoothness and feasibility of all surfaces.

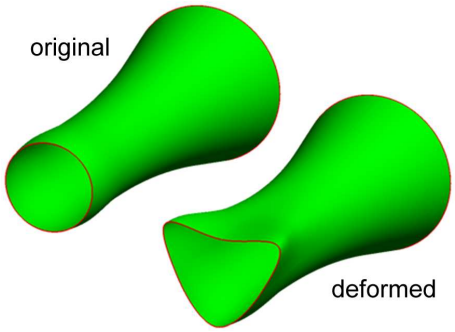


Fig.4 Deformation of the NURBS-surface forced by deformation of its edge.

2.2.2 Morphing

The parametric design based on a morphing technique consists in creating the new geometry as a weighted sum of certain basic geometries. In such a case it is assumed that set of basic geometries was prepared in earlier stage of design. This preliminary design is performed usually in interactive mode. The number of Design Parameters does not matter.

The morphing-weighted sum is realised in the space of Design Parameters, giving finally the set of morphed Design Parameters that define the morphed geometry. As a result, the weights corresponding to different basic geometries may be considered as alternative, higher-order Design Parameters.

The idea of parametric modelling based on morphing technique is illustrated in Fig.5. The basic geometries for morphing procedure are: cube, sphere and cone. All these three different objects are modelled by the same parametric model and they have the same topology. By defining appropriate morphing weights it is possible to generate a three-parametric family of solids. Examples of such morphed solids are shown in Fig.5.

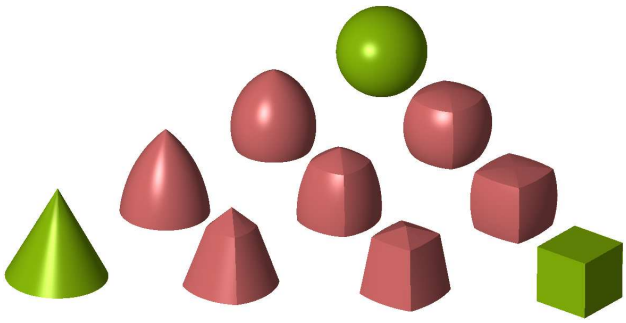


Fig.5 Application of the morphing technique in parametric modeling of three-parameter family of solids.

The application of the morphing technique in parametric modelling has several advantages. The first one is a small number of Design Parameters (dimension of morphing base), because usually the morphing is realised based on only a few basic geometries. The next advantage is that the geometries generated by the morphing technique usually inherit properties of their parents – the basic geometries. This means, that if all basic geometries are smooth and feasible, it is likely (but not certain) that the result of morphing will have similar properties.

3 PARADES software

Based on the methodology described above, the software PARADES has been worked out. The PARADES is a highly specialised tool supporting parametric modelling of complex objects, especially within the aerospace engineering domain. This software differs significantly from traditional CAD systems and the current version does not support the classical, not parametric design. By using software PARADES, the experience is only required when creating a new parametric model, while the use of properly prepared parametric model is possible even in the case of an user inexperienced in the CAD domain.

The PARADES may be utilised in two modes: interactive and automatic. The interface of the software working in interactive mode is shown in Fig. 6. The main elements of the software interface are:

- Main Window, where parameterised objects are visualised, in various display modes
- Tree of Design Parameters, where all the Design Parameters are listed and may be chosen by the user
- Design Parameter Panel – where current value of selected Design Parameter is displayed and may be modified
- Tool Buttons collected in three groups: visualisation, parameters and properties
- Main Menu

In interactive mode, a designer using a friendly interface modifies the current values of Design Parameters in order to achieve the geometry having required properties. The user establishes required values of the Design Parameters both by entering them in edit window or by sliding an appropriate slider. Additionally, instead of changing values of the Design Parameters directly, it is possible to use morphing technique to obtain designed geometry. The interactive mode is also used for checking whether a given parametric model works properly.

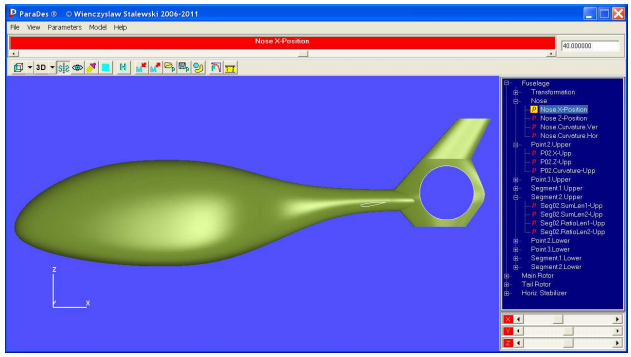


Fig. 6 The interface of PARADES software working in interactive mode.

In automatic mode the software PARADES is usually called by other software managing the process of the Numerical Optimisation. In such a case, the PARADES for given input - set of values of Design Parameters, generates automatically the geometry. In case when the parametric model utilises the morphing technique, the input data are the morphing weights.

Besides of creating the parameterised objects, the PARADES software may also perform some analyses concerning these objects. Typically, these analyses relate to geometric properties such as surface area, volume, distribution of the surface area of sections, etc., but they may be customised taking into account specifics of given problem. If needed, it is also possible to utilise external codes to do additional, specialised analyses. Results of all these analyses may be used in optimisation process to evaluation of objectives and constraints.

4 Examples of applications

The PARADES software was applied in many commercial and non-commercial projects

and research. Area of applications of the software covers a wide range of problems concerning the design and optimisation mainly in the field of aeronautical engineering. Among other things these problems concerned design and optimisation of:

- airfoils [6],[7]
- aircraft wings [2],[3],[4]
- high lift devices [2],[3] for aircraft wings
- air intake system for turboprop engine [1]
- nozzles of turbojet engines
- helicopter main rotor
- helicopter fuselage [8]
- stratospheric airplane
- propellers
- wind tunnel

Selected examples of application of PARADES software are presented below.

4.1 Multidisciplinary optimisation of aircraft wing

Within the EU Projects NACRE and CESAR the PARADES software was used to support the multidisciplinary optimisation of aircraft wings [2],[3],[4]. In these research usually the whole aircraft was modelled using the PARADES software but only wing and optionally its high lift system were parameterised.

The interdisciplinary optimisations were performed taking into account the aerodynamic and structural properties of the wing. Such formulation of optimisation problem caused that in parametric model of wing (aircraft) both the external shell of aircraft and structure of wing spar were parameterised. The parametric model of the wing shell generated input data for CFD calculations (see Fig. 7) while the parametric model of the wing structure generated data for CSM analysis assessing the weight of the wing (Fig. 8). In described research the parametric models of the wing were built usually in the same manner. The Design Parameters defined:

- wing planform (surface area, span, base chords, sweep angle, taper ratio, wing segmentation, etc.)
- spanwise distribution of relative thickness (distribution of airfoils)
- spanwise distribution of twist
- spanwise distribution of camber of airfoil mean line

This way, by using relatively small number of Design Parameters both the wing planform and wing cross-sections were parameterised. In project CESAR additionally the wing high lift systems were parameterised, as it is shown in Fig. 9. The high lift systems were optimised with regard of optimal position of Fowler flap during takeoff and landing.

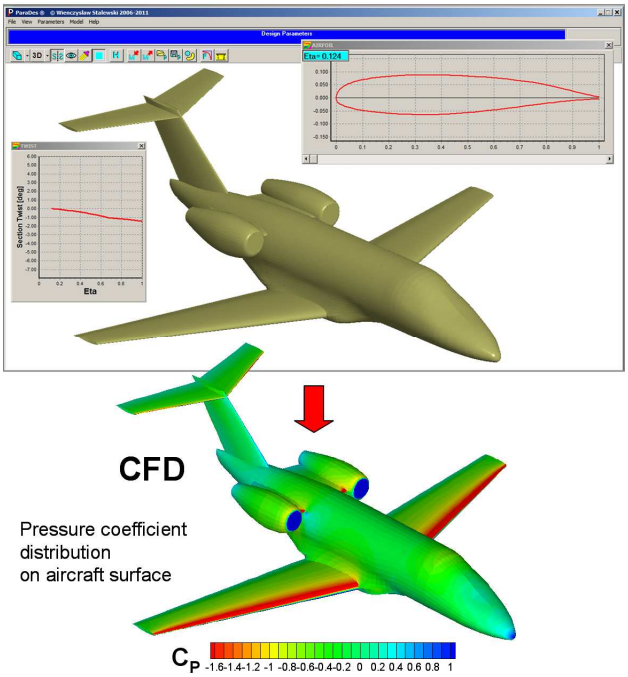


Fig. 7 Parametric model of the aircraft external shell, - the input for CFD calculations.

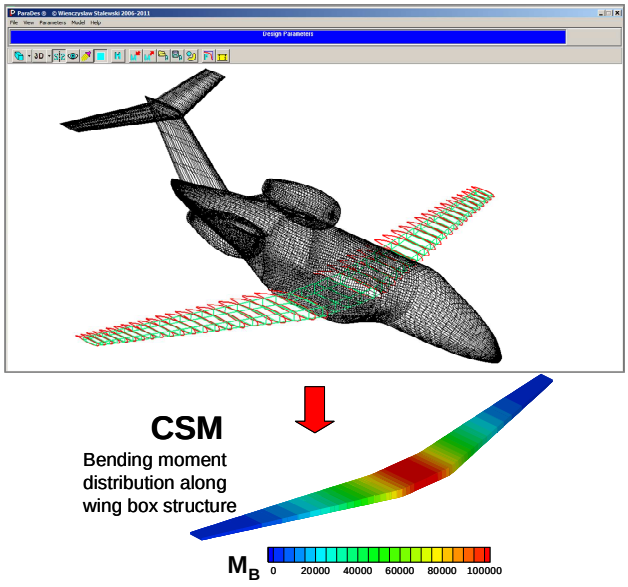


Fig. 8 Parametric model of the wing spar - the input for CSM calculations.

In Fig. 10 the results of Multidisciplinary Numerical Optimisation of low-speed, small aircraft have been presented. In this case only wing planform was parameterised and two objectives were taken into consideration: minimal drag and minimal wing weight. Fig. 10 presents obtained Pareto Set and planforms of selected, Pareto-optimal aircrafts.

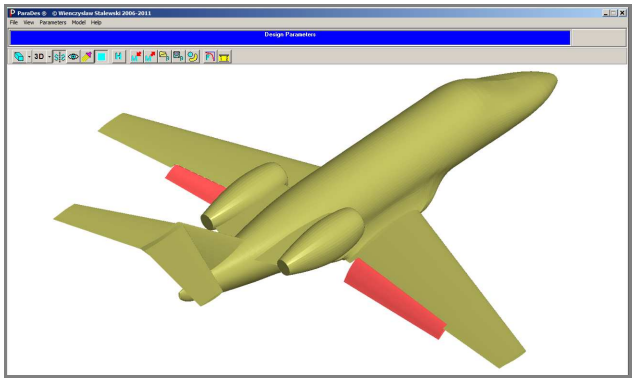


Fig. 9 Parametric model of aircraft wing and its high lift system.

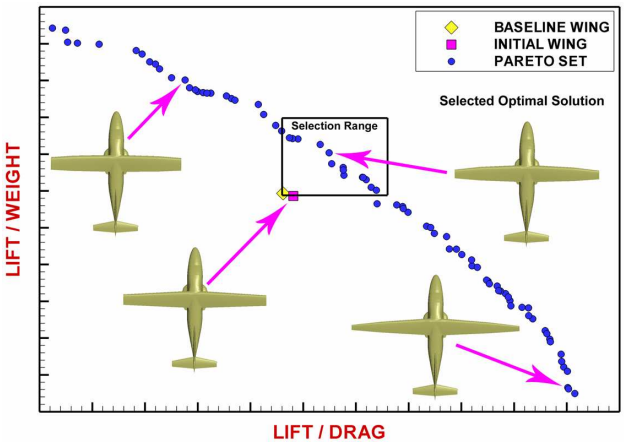


Fig. 10 Parametric optimisation of low-speed, small aircraft taking into consideration two objectives: Minimal Drag and Minimal Wing Weight

4.2 Parametric model of exhaust system of turbojet engine

Within the EU project HISAC the CFD studies of an exhaust system for Variable Cycle Turbofan of a small supersonic business jet were performed. The PARADES software was applied to create the parametric model of investigated object. This model is shown in Fig.11. By using the parametric model of exhaust system several alternative geometries of this system were generated. CFD study of flow through the various variants of exhaust system (Fig.12) let to perform sensibility analysis and select possibly the best solution.

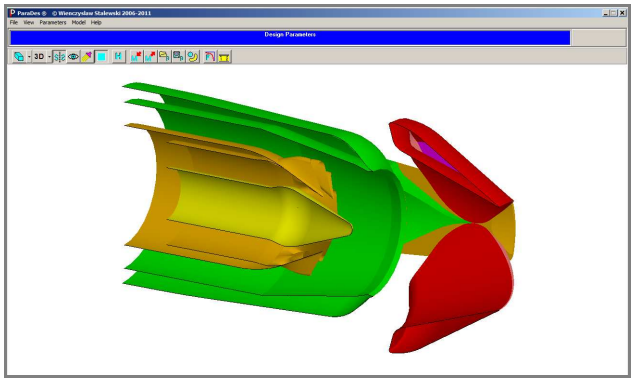


Fig.11 Parametric model of exhaust system for Variable Cycle Turbofan.

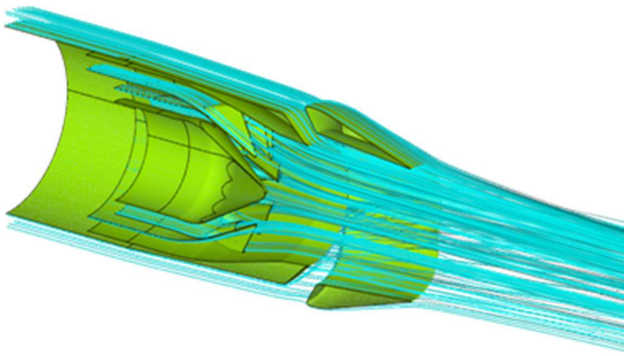


Fig.12 CFD study of flow through the exhaust system .

4.3 Optimisation of air intake system for turboprop engine

The PARADES software was applied in the Direct Optimisation process of air intake system for turboprop engine [1]. Both the air intake and the bypass were parameterised as it is shown in Fig. 13.

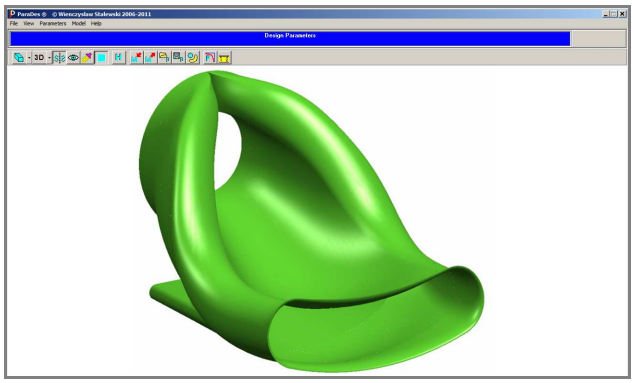


Fig. 13 Parametric model of air intake system for turboprop engine.

In preliminary stage of design process a simple, two-parameter model of isolated air intake was investigated. In the final phase the complete air intake system was optimised. During the process of design and optimisation, for every configuration created by PARADES software, the thorough CFD analyses were performed. As an example of such analysis, in Fig. 14 the total pressure distributions in selected cross-sections of air intake have been shown.

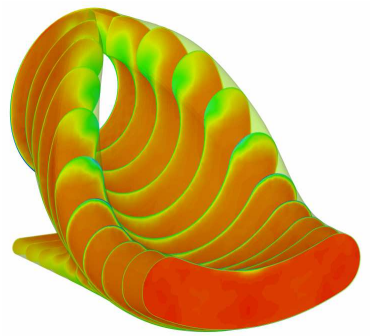


Fig. 14 CFD study of parameterised air intake system for turboprop engine. Total pressure distributions in intake cross-sections (courtesy of EADS-CASA).

4.4 Optimisation of helicopter fuselage

By using the PARADES software the parametric model of the helicopter has been worked out. The purpose of the work was to redesign the helicopter fuselage so as to improve aerodynamic properties and mass balance of the whole helicopter. It was assumed that within the optimisation process the effect of working main and tail rotor would be taken into consideration. The work was carried out in three sequential stages:

1) development of parametric model of the helicopter,
2) design of few basic modifications of helicopter using direct optimisation approach,
3) design of the final version of helicopter using numerical optimisation method based on a genetic algorithm and morphing technique.

The developed parametric model of the helicopter is shown in Fig. 15. Although the assumed modifications of geometry concerned only essential part of fuselage, other components of helicopter also were parameterised because of planned further research. In particular it concerned both the vertical and horizontal stabilizer. In presented approach, the effect of main and tail rotor was modelled in simplified way by application of Virtual Blade Model (VBM), that consists in the introduction of artificial sources of momentum in the fluid zones designated by the rotating rotor. As a result, in the parametric model the rotor blades were not physically modelled. Instead of this, the boundaries of fluid zones modelling the rotor were parameterised as well as geometric data of rotor blades (e.g. twist, thickness, planform, airfoils, etc.). All these data were the input data for VBM.

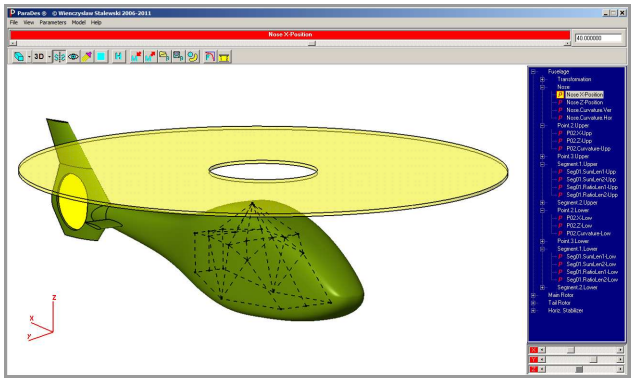


Fig. 15 Parametric model of helicopter.

Within the presented parametric model, the major part of the fuselage was modelled as a multi-section surface, defined by sections curves swept along guiding curves. A part of Design Parameters defined the transformations of control points of guiding curves (see Fig. 16). Additionally, some Design Parameters define smooth longitudinal changes of weights of control points of section curves, this way influencing the shape of fuselage cross sections. Within the optimisation process only the Design Parameters defining major part of the fuselage were taken into consideration.

The Direct Optimisation of helicopter fuselage was performed using the PARADES software in interactive mode. As a result three new variants of helicopter fuselage were designed. These variants: BAS-1, BAS-2 and BAS-3 together with initial version BAS-0 were chosen as the base for morphing procedure. The Numerical Optimisation based on Genetic Algorithm and morphing technique led to the

final solution, marked PAR-1. This helicopter has aerodynamic properties considerably better than initial version BAS-0. In Fig. 17 the geometry of helicopter PAR-1 is shown together with the basic geometries and their morphing weights corresponding the final geometry.

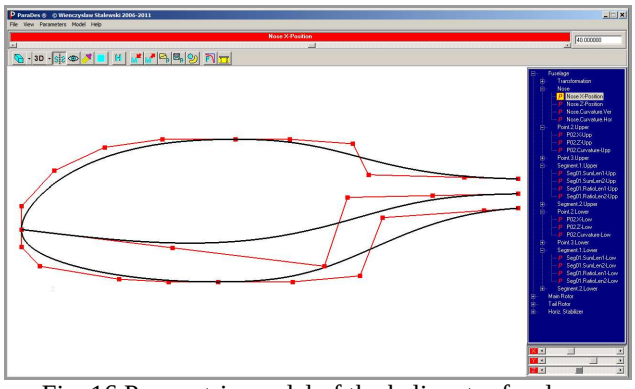


Fig. 16 Parametric model of the helicopter fuselage surface based on a system of constructional NURBS-curves.

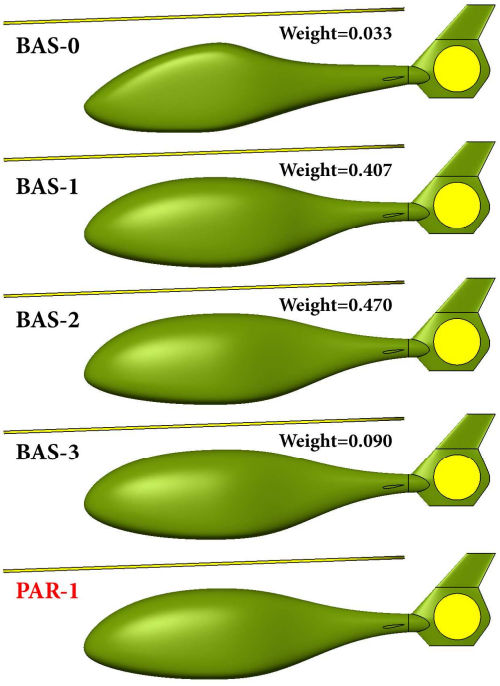


Fig. 17 The helicopter PAR-1 as a result of morphing of four basic geometries.

5 Conclusions

The PARADES is the software supporting parametric modelling of complex geometries especially within the aerospace engineering design. The software was developed based on the methodology, whose main idea was to create a tool capable to automatically generate thousands of complex geometries, smooth and feasible, properly read by the grid generators during the numerical optimisation process. The second idea was to create a tool of parametric modelling, that could be used even by less experienced CAD user to fast generation of several variants of designed product, using a friendly interface.

Previous experience in practical applications of the software seem to confirm that the PARADES copes quite well to meet above challenges. The software PARAD_{ES} is still developed and improved to become the successful tool supporting design and optimisation.

References

[1] Stalewski W., Rarata Z., “PARADES - the CAD software for Air Intake and Bypass Parameterisation”, Institute of Aviation Report TN- 16/BA/05/M, 2005.

[2] Stalewski W., Żółtak J., “Wing Design And Analysis Report AC1, Turbulent Wing, AFT, HLT”, CESAR Project (6FP), Deliverable Report D1.2.3-7.2009.

[3] Stalewski W., Żółtak J., “Wing Design And Analysis Report AC2, Turbulent Wing, AFT, HLT”, CESAR Project (6FP), Deliverable Report D1.2.3-9.2009.

[4] Rokicki J., Stalewski W., Żółtak J., “Multi-Disciplinary Optimisation Of Forward-Swept Wing”, in “Evolutionary Methods for Design, Optimization and Control”, eds.: T. Burczynski and J. Périaux, CIMNE, Barcelona, Spain, 2009.

[5] Piegl L., Tiller W., The NURBS Book, Springer. 1997.

[6] Stalewski W., Żółtak J., “Definition of Turbulent Wing for AC2 Aircraft”, CESAR Project (6FP), Deliverable Report D1.2.1-9. 2009.

[7] Kania W., Stalewski W., Zwierchanowska B., “Inverse and Direct Optimisation of Tail Rotor Airfoil Using Genetic Algorithm”, Proceedings Of The Forth International Seminar On RRDPEAE'2000, Warsaw, Poland. 2000.

[8] Stalewski W., Żółtak J., “Multicriteria Design and Optimisation of Helicopter Fuselage”, in “Evolutionary and Deterministic Methods for Design, Optimization and Control”, eds.: C. Poloni, D. Quagliarella, J. Périaux, N. Gauger and K. Giannakoglou (Eds.), CIRA, Capua, Italy. 2011. (in print)

3rd CEAS Air&Space Conference
21st AIDAA Congress

Enhancement of CEASIOM with Rapid-Meshing Tool for Aircraft Conceptual Design

M. Zhang and A. Rizzi
Royal Institute of Technology (KTH), Stockholm, Sweden

D. Raymer
Conceptual Research Corp., Playa del Rey, CA, USA

Keywords: *RDS, SUMO, CEASIOM, EDGE, rapid-CFD*

Abstract

This paper details the development and application of the RDS-SUMO-CEASIOM-EDGE rapid-CFD tool. It uses the RDS CAD model as geometry for automated meshing and CFD analysis to produce an aero-data base for control and stability analysis. It is applied to two non-conventional design proposals, an asymmetric twin-prop aircraft and an airliner with rear Open Rotor propulsion, retractable canard, and a "chin-rudder" instead of vertical tail. For both configurations, yaw control is problematic, and the stability and control analysis is used to assess control surface sizing and stability augmentation system.

1 General Introduction

Computational Fluid Dynamic analysis for the aerodynamic characteristics determining performance, stability, handling qualities, and control effectiveness is a vital part of modern aircraft design. Many CFD tools exist for this analysis but their application to the earliest phases of aircraft design is currently limited by the time it takes to perfect the conceptual geometry model and then create a computational mesh of sufficient quality to have confidence in the computed results. Historically, high-fidelity grids of a new design concept were not expected until months or even years into the design development process. Designers would neglect the finer details of surface arrangement and intersection, yet to be defined, as they focused upon their real job, the design integration of a workable and optimal vehicle concept.

The RDS Professional software[*] developed by Conceptual Research Corporation is a well-known tool for conceptual aircraft design. Tailored to allow the designer to develop a new concept or modify a previous baseline design, its proprietary Design Layout Module (DLM) permits rapid aircraft configuration layout by mouse-driven interactive computer graphics. The RDS-DLM geometry uses its own unique format DSN [1] consisting of components defined by either stored surface points or modified parametric Bezier curves (NURBS subset). Each component has intended for production-quality lofting, so its model can not be meshed without additional refinement.

[*] Dan Raymer's Aircraft Design & RDS Website, http://aircraftdesign.com

The CEASIOM [†] (Computerized Environment for Aircraft Synthesis and Integrated Optimization Methods) environment is developed to provide computational analysis of a wide class of aircraft at a very early stage of design. It is meant to support engineers in the conceptual design process of the aircraft, with emphasis on the improved prediction of stability and control properties achieved by higher-fidelity methods than found in contemporary aircraft design tools. One of the primary driving forces behind the development of the CEASIOM environment is to facilitate early computational analysis of highly novel aircraft concepts at a very early stage of its design. CEASIOM consists of a set of tools which allow the user either to generate from initial templates, or to import, a wide range of aircraft concepts and then use these to work through the design process generating the required aircraft parameters as they proceed.

CEASIOM requires an unified initial layout as input as the baseline configuration sized to the mission profile by specification of the O(10) parameters employed in the pre-design loop. Then it refines this design (in concept-design loop O(100) parameters) and outputs it as the revised layout for consideration in the down-select process (say O(1000) parameters) in Fig. 1.

SUMO[‡], integrated into CEASIOM as part of its geometry module, is being developed to handle CAD geometry from a variety of sources, including RDS Professional. It contains a fast mesh generator and generates unstructured volume grids automatically with the tetrahedral mesh generator TetGen [§]. The SUMO (with TetGen) produces the simplified CAD (O(100) parameters) in Fig. 1, used for inviscid compressible CFD.

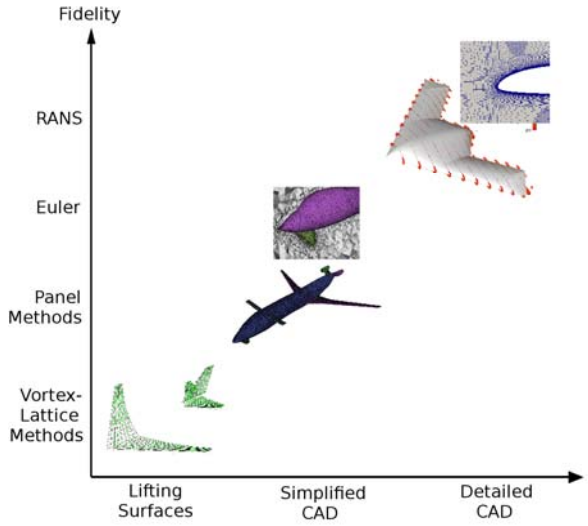


Fig. 1 Varied levels of geometry from pre-design to detailed-design, where RDS geometry corresponds to the simplified CAD.

Note that the RDS geometry is not intended for production-quality lofting, so its model can not be meshed without additional refinement. To get a meshable model from RDS geometry, a RDS-SUMO interface is developed. It reads the RDS-DLM DSN file and writes the SUMO SMX file. The volume mesh can be saved in the native format of the CFD solver EDGE [7] which computes Euler solutions in this paper and the fast and simple CAD geometry would be obtained from the interface Fig. 1. Although RDS can output geometry in the standard DXF format, the interface works directly from the RDS internal format to avoid an intermediate step which would add time and inaccuracies to the translation.

The interface links the RDS aircraft design software with CEASIOM, which is proposed for preliminary aircraft design, and provides variable fidelity CFDs and stability analysis in early design stage. Thus the ideas embedded in the RDS model become a “virtual aircraft” that can be analyzed and maneuvered by CEASIOM (Fig. 2). Once the virtual aircraft satisfies the engineer’s and the customer’s requirements, wind tunnel models could be built. The details of virtual aircraft design are spelt out in [3].

CEASIOM has its geometry construction system – AcBuilder (Aircraft Builder), to create and visualize its initial layout. It supports lifting surfaces, fuselages, tail booms, fairings, pylons,

[†] http://www.ceasiom.com
[‡] Larosterna Engieering Dynamics Lab, Aircraft Modeling & Mesh Generation, http://www.larosterna.com/sumo.html
[§] A Quality Tetrahedral Mesh Generator and a 3D Delaunay Triangulator, http://tetgen.berlios.de/

nacelles, and integration of engine intakes and nozzles and is award of control surfaces such as ailerons, elevators, rudders, flaps, slats, and all-moving surfaces. To accomplish the whole "virtual-aircraft design" loop, CEASIOM could do the job alone, referring to [4] pp. 10-11. A main advantage of coupling RDS with CEASIOM in the geometry design stage, is that the RDS-DLM can provide more geometric details than CEASIOM-AcBuider. For example, it is no way to create an n-section-wing aircraft by using AcBuilder. As Fig. 1 shows, the RDS geometry corresponds to the simplified CAD, while AcBuilder provides the very basic geometry parameters employed in the pre-design stage. Turning the RDS model into a solid model with little inaccuracies during process would save the efforts when "re-design" occurs. The rapid-meshing tool could provide qualified mesh and thus make an enhancement of CEASIOM.

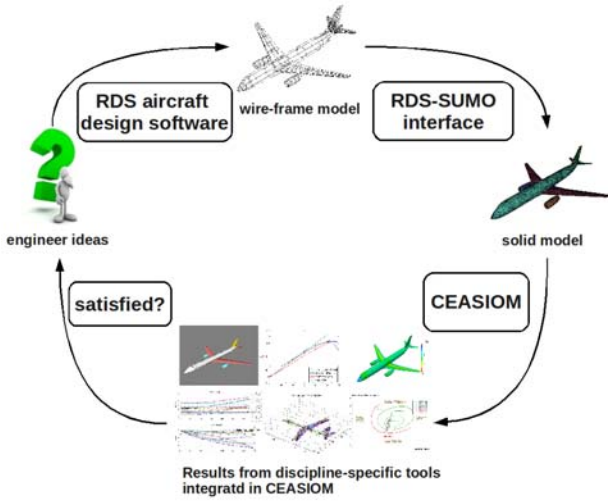


Fig. 2 Virtual aircraft design loop by RDS coupled with CEASIOM

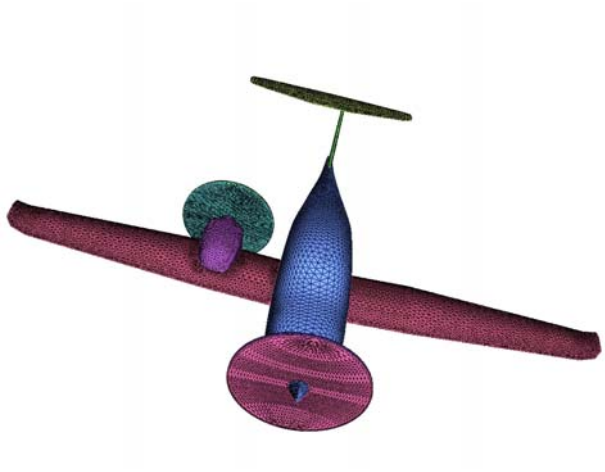
2 From Lofting Geometry to Meshes

To illustrate how the interface works, two aircraft models are studied, which are designed and developed by Dr. Dan Raymer through RDS Professional. The qualities of the mesh are acceptable for solving Euler equations and the transonic properties would be captured when needed. Yaw stability is of great interest for both configurations.

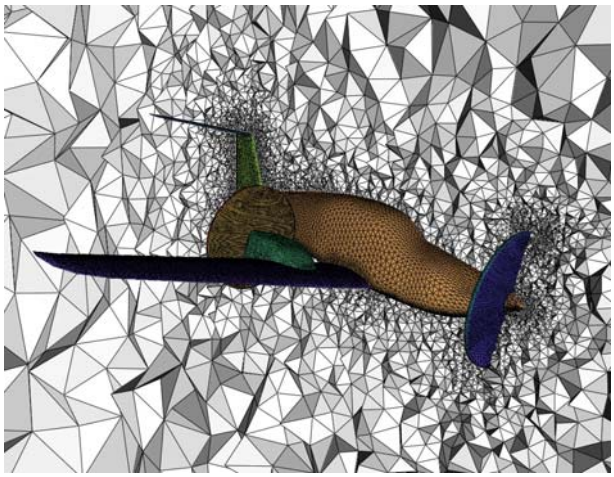
2.1 Asymmetric aircraft concept

As a test case for the RDS-SUMO-CEASIOM process, an asymmetrical aircraft concept developed several years ago was employed [6]. With both tractor and pusher propellers, an offset fuselage, and an extra nacelle on the starboard wing, the design is a suitable *stress test* for a rapid meshing program. The RDS geometry was emailed from California to Stockholm, and after about 2 hours of minor modifications, a smooth geometry was ready for automatic gridding and CFD analysis including a simple propeller disk model [8].

Figure 3 presents the SUMO meshable model of the concept, and the automatically generated tetrahedral volume mesh with 9 million cells for solving the Euler equations.



(a) Surface mesh upon the model

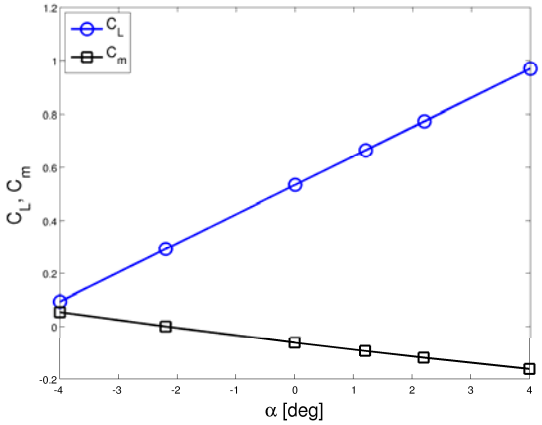


(b) Volume mesh for Euler solution

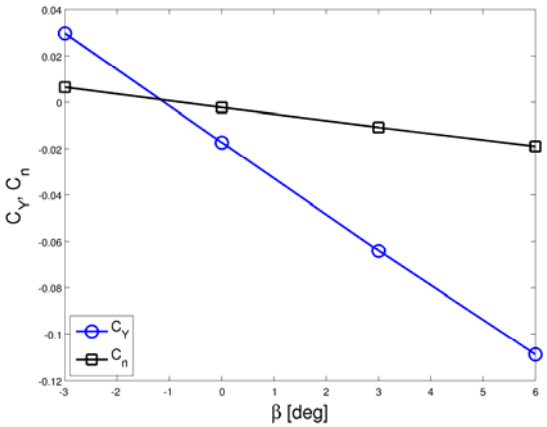
Fig. 3 SUMO mesh of the asymmetric twin-prop aircraft

Since it is essential to include the propeller-thrust effects in the aerodynamic characteristics, the aerodynamics was computed by Euler EDGE. Potential-flow methods such as vortex-lattice or panel cannot easily model the momentum source. The propeller is simulated as a "very short" nacelle in SUMO-CEASIOM-EDGE. Specification of (the same) mass flow into and out of the disk adds a momentum that provides the pressure jump, but the mass flow model cannot account for finer details such as the swirl in the propeller slipstream. The propeller thrust was chosen to balance the cruise drag. Fig. 3 visualizes the streamlines in the flow accelerated by the propellers.

Figure 4(a) shows C_L and C_m. The very cambered Natural-Laminar-Flow airfoil has extremely high lifting slope [8] which makes the lift coefficient of the whole configuration marginally exceed the "ideal" flat plate lifting slope 2π. Figure 4 shows stick-fixed stability in both pitch and yaw, $C_{m,\alpha} < 0$ and $C_{n,\beta} < 0$.



(a) Lift and pitching moment coefficients



(b) Side force and yawing moment coefficients

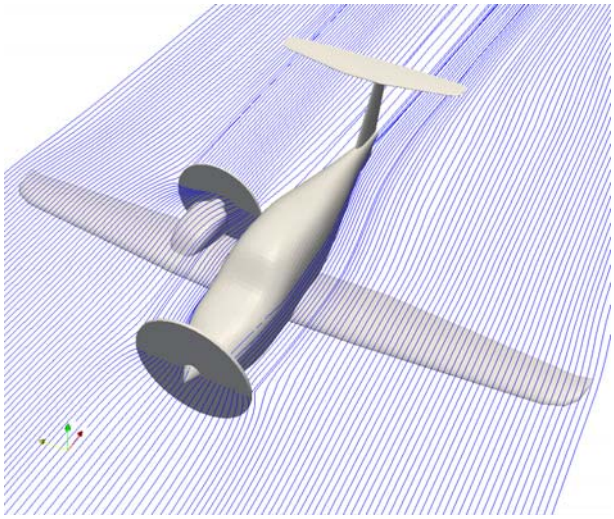
Fig. 4 Aerodynamic moments and forces computed by Euler EDGE

EDGE calculates the aerodynamics of control surface deflection by the transpiration boundary condition, with flow tangency to the deflected surface prescribed on the undeflected surface. This method avoids mesh deformation, so one mesh can be used for all the deflection cases. However, it doesn't work well for large deflections.

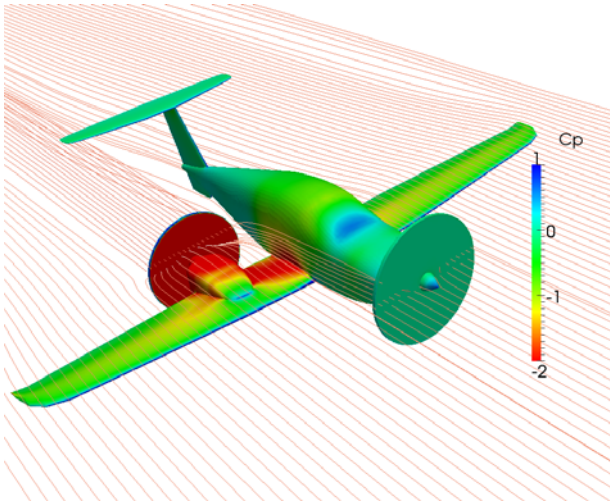
The asymmetry might cause serious control problems especially when the front engine fails and the analysis focuses on yaw trim and stability. For low speed and small side-slip angles, the aerodynamics follow a linear trend. The analysis predicts that to fly with a sideslip 3 degrees nose-left, the rudder needs to deflect around 8 degrees to left to trim.

The failure of the front engine would make yawing more serious since more thrust comes

from the single pusher propeller. Figure 5 presents the streamlines and pressure coefficient on the surface with front engine out. There are several immediate effects due to one engine failure. The initial effect is the yawing that occurs due to the off-center thrust which makes the aircraft drift to the dead engine, i.e., yaw to nose-left. The inherent directional stability tends to oppose the asymmetric yawing moment. "Yaw-induced roll" is a secondary effect which accompanies yawing motion due to the higher lift produced on the advancing wing. In this paper the asymmetric yawing moment is preliminarily considered and the balanced flight with rudder is proposed.



(a) Stream line visualization for the twin prop aircraft



(b) Pressure and stream line visualization of the one-engine-out mode (OEM), M = 0.282, $\alpha = 0^\circ$, $\beta = 3^\circ$

Fig. 5 Aerodynamic moments and forces computed by Euler EDGE

Table 1 shows the lateral coefficients for twin-prop and one-engine-out flight. The right wing-mounted pusher provides more asymmetric force, that tends to yaw the aircraft to left (i.e., the dead engine). To fly at zero side-slip, a rudder deflection around 10 degrees is required compared with 1.6 degrees rudder deflections when two propellers are both working (Fig. 6). Such deflection required for trim is probably too high. Compensation for strong lateral gusts may need more yaw moment authority than is available when 10 degrees are used for trim, and rudder stall becomes an uncomfortable plausibility. There are two ways to reduce the rudder trim deflection and enhance its authority, one is to reduce (the absolute value of) the C_n at $\beta = 0^\circ$, another is to increase the rudder efficiency $C_{n,\delta}$.

The rudder needs to be re-sized to counter the extra yawing moments produced during an engine-out event. The rudder is originally sized as 27.5% of the chord along the vertical tail. To get better or smaller rudder deflections, the rudder efficiency could be enhanced by either increasing the rudder fraction of the vertical tail, or keeping the rudder fraction while increasing the size of the vertical tail, so that $C_{n,\beta}$ becomes larger and trim beta becomes smaller. Both methods increase the force produced by the rudder. In this paper an attempt of rudder re-sizing has been made, by making the rudder 50% larger than the original size, i.e., 41.5% of the chord along the vertical tail. The new rudder shows better trimming properties, for example, to fly at zero side-slip, the rudder needs to be deflected 7° instead of 10° (Fig. 6).

The asymmetric flight would make the control more demanding and serious, whatever the working engine(s) is one or two. The best way to solve this problem probably is, to add another engine on the left wing!

mode	$C_n(\beta = 0^\circ)$	$C_{n,\beta}$	$C_{n,\delta}$
Normal	-0.00221	-0.0028	0.0014
OEM	-0.01717	-0.0036	0.0017
OEM, larger rudder	-0.01717	-0.0036	0.0025

Table 2 Lateral coefficients for twin-prop and one-engine-out mode

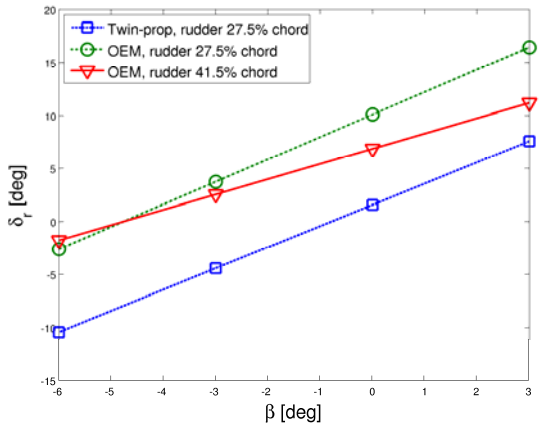


Fig. 6 Side-slip angle β versus rudder deflection δ_r when trim in yaw

2.2 Tailless Chin-Rudder Aircraft

The study of tailless aircraft with chin-rudder, in Fig. 7, was funded by NASA future airliners "N+3" project. It was modeled based on Boeing 737-800 and with a design goal that obtaining a 70% reduction in fuel consumption. The Tailless design features with a retractable canard and an all-moving chin-rudder. The Open Rotor propulsion system mounted on the rear part of the fuselage seems to offer the greatest payoff in terms of engine specific fuel consumption [10]. The canard is extended at take-off and landing to counter the huge nose-down pitching moments created by the large wing flaps by adding lift for. During cruise it retracts to reduce drag by reducing the wing area. Control is provided by a combination of wing trailing edge surfaces, the chin-rudder and nacelle-pylon ruddervators. The all-moving chin-rudder is the primary responsible for yaw control, and CFD analysis is performed in this paper to confirm the sizing of the chin-rudder for yaw control.



(a) Take-off with extractable canard

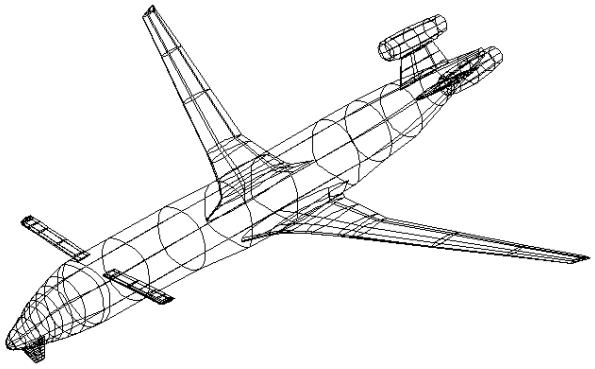


(b) Cruise with retractable canard

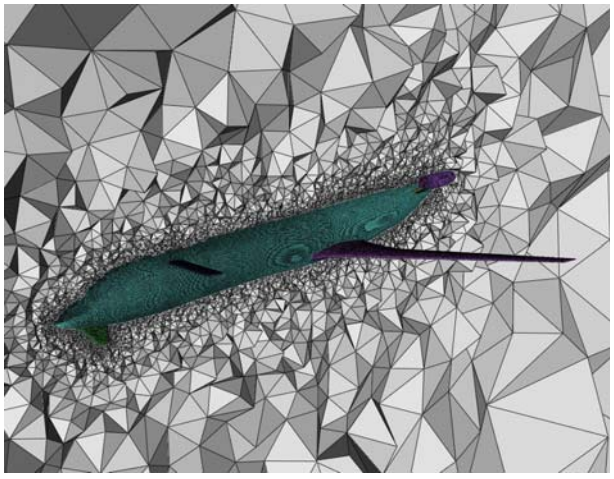
Fig. 7 Preferred concept artist's rendering: Tailless airliner

2.2.1 Mesh and CFD analysis

The RDS-DLM module provides the wire-frame model that the rapid interface could turn it as a "solid" mesh on which the Euler equation would be solved. The process is automated without any human intervention. The sizes of the control surfaces and flaps are also specified as overlapped wings in the native RDS file and modeled in SUMO accordingly.



(a) RDS wire-frame geometry



(b) SUMO-TetGen volume mesh with 5.2 million cells

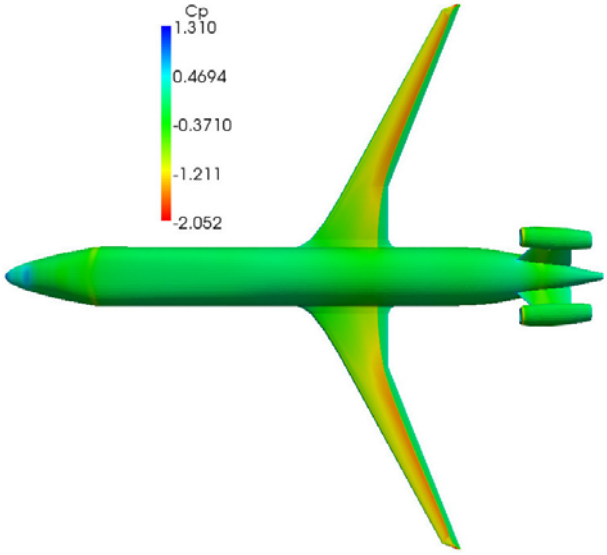
Fig. 8 Preferred concept artist's rendering: Tailless airliner

When the aircraft cruises at Mach 0.8 the transonic shock forms on the upper main wing and extends downstream until 75% of the local chord (Fig. 9(a)). To compare the results from Euler and classical DATCOM in RDS, the Euler gives $C_{L,\alpha} = 0.155$ per deg while DATCOM gives 0.160 per degree. That calibrates the Euler analysis from the RDS geometry.

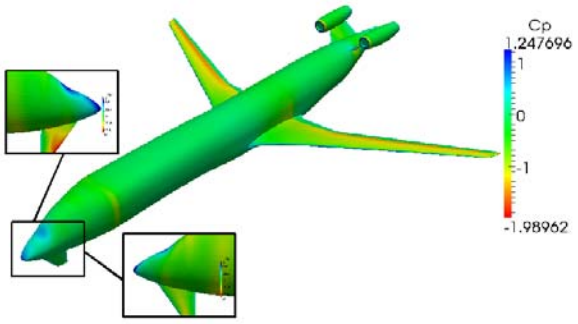
2.2.1.1 Chin rudder authority

The chin rudder is proposed to restore the moment in yaw, while the aircraft is suffering, for example, a lateral gust. It is also required to have sufficient authority to do the one-engine-out control [10]. While the nacelle pylon ruddervators could help produce yaw moment, in this study we use the chin rudder only for the yaw control during normal flight. This aircraft has inherent instability in yaw. However as long as sufficient control is attained, we could probably forget stability [10].

The chin rudder is deflected for yaw moment control. The Euler solver shows that in order to balance a 3° nose left sideslip, the chin-rudder needs to be deflected 6° relative to the free stream [10]. Figure 9 shows the transonic aerodynamics, and Fig. 9(b) shows the Euler pressure distribution with a zoom on the chin-rudder. Note that the nose of the fuselage is a bit rough, resulting in pressure oscillations. In a more detailed study it should be smoothed out.



(a) $\alpha = 1^\circ$, $\beta = 0^\circ$

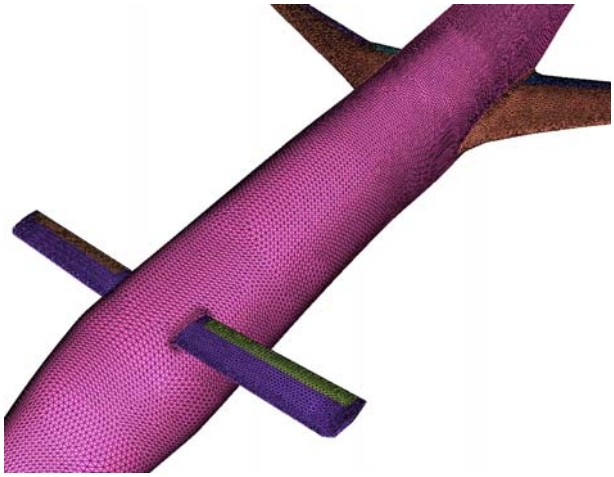


(b) $\alpha = 0^\circ$, $\beta = 3^\circ$, **and rudder deflection** $\delta_r = 9^\circ$

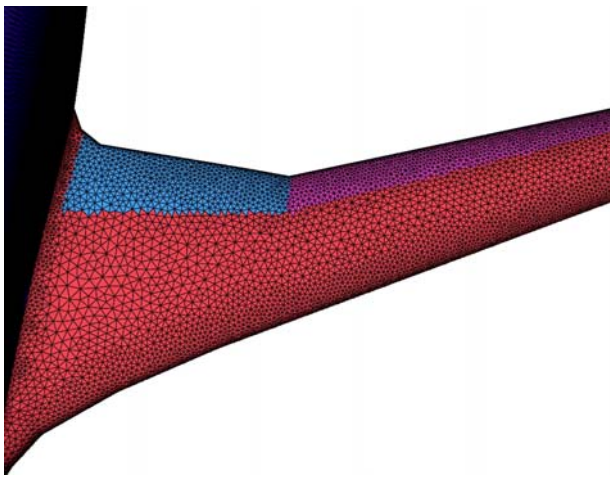
Fig. 9 Euler computed surface pressure of the Tailless aircraft at cruise point M = 0.8

2.2.1.2 Canard action at take-off

The retractable canard is used for adding nose-up pitch moment when the large wing flaps deflect and retracts completely for efficient cruising flight. The canard flap is 28% of the chord, and the wing flaps are 26% of the chord, shown in Fig. 10.



(a) Canard flap



(b) Wing flaps, inboard and outboard flaps

Fig. 10 Flaps on the surface mesh

The canard location at the nose makes the moment arm long enough to produce sufficient nose-up moment for take-off when both the inboard and outboard flaps are deflected 30 degrees producing lift and a pitch-down moment. The canard elevator is deflected 15 degrees down to counter the nose-down pitch and at Mach 0.242 gives an increment of 0.544 in C_L and 0.708 in C_m at $\alpha = 0^\circ$. Fig. 11 presents the pressure distribution. The low pressure regions on the wing and canard trailing edges due to flap deflections are obvious.

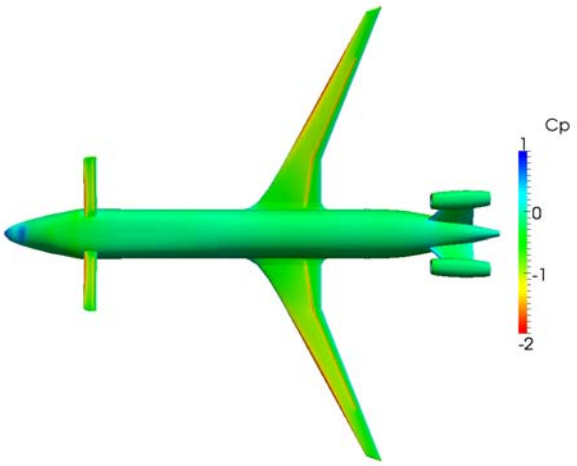


Fig. 11 Surface pressure distribution from Euler solution, M = 0.282, $\alpha = 0^\circ$, $\beta = 0^\circ$,

3 Conclusion

The CEASIOM software is to support engineers in many aspects of preliminary aircraft design. The RDS-SUMO interface rapidly produces grids for inviscid compressible CFD to enhance CEASIOM's capacities. The qualities of the mesh for the two test cases are sufficient that could provide an equivalent propeller model, and promising transonic analysis. The yaw trim for both cases are investigated for lateral stability & control. The RDS-SUMO rapid-meshing tool illustrated in this paper enables quick aerodynamic assessment of even radical new concepts in early stages of aircraft design. This opens the way to prediction of the flying qualities and 6 DOF flight simulation by e.g. the SDSA [11] module in CEASIOM.

References

[1] Raymer D., “Conceptual Design Modeling in the RDS-Professional Aircraft Design Software”, AIAA Aerospace Sciences Meeting, Orlando, FL, Paper No. 2011-161, 2011.

[2] Rizzi A., Oppelstrup J., Zhang M. and Tomac M., “Coupling Parametric Aircraft Lofting to CFD & CSM Grid Generation for Conceptual Design”, 49th AIAA Aerospace Sciences Meeting

including the New Horizons Forum and Aerospace Exposition, Orlando, Florida, Paper No. 2011-160, 2011.

[3] Rizzi A., Eliasson P., Goetzendorf-Grabowski T., Vos J., Zhang M., "Virtual-Aircraft Design & Control of TransCruiser – a Canard Configuration", preprint submitted to Progress in Aerospace Sciences, April 9, 2011.

[4] Zhang M., "Application and Development of the CEASIOM-SUMO-EDGE Suite for Rapid AeroData Assessment of Aircraft Flying Qualities", Licentiate Thesis, Department of Aeronautics, Royal Institute of Technology (KTH), Sweden, June, 2011,

[5] Raymer D., "Enhancing Aircraft Conceptual Design Using Multidisciplinary Optimization", Doctoral Thesis, Department of Aeronautics, Royal Institute of Technology (KTH), Sweden, May, 2002,

[6] Raymer D., "Simplified Aircraft Design for Homebuilders", Design Dimension Press, Los Angeles, CA, 2003

[7] Edge User Guide, FOI dnr 03-2870, ISSN-1650-1942, FOI, Swedish Defence Research Agency, Systems and Technology, Stockholm, Sweden, 2007.

[8] Zhang M. and Rizzi A., "RDS-SUMO: from Lofting to Physics-Based Grids", submitted to Aircraft Engineering and Aerospace Technology Journal, 2011.

[9] Dan M. Somers, Langley Reserach Center, Hampton, Virginia. "Design and Experimental Results for a Natural-Laminar-Flow Airfoil for General Aviation Application", NASA Technical Paper 1861, June, 1981.

[10] Raymer D., "Advanced Technology Subsonic Transport Study, N+3 Technologies and Design Concept", NASA-CR Contractor Report, Conceptual Research Corporation, Los Angeles, CA, 2011.

[11] Kaenel R., Rizzi A., Oppelstrup J., et. al., "CEASIOM: Simulating Stability & Control with CFD/CSM in Aircraft Conceptual Design", 26th International Congress of the Aeronautical Sciences, ICAS, Alatyska, USA, 2008.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Convergence Acceleration of the CFD code Edge by LU-SGS

E. Otero
Dept. Aeronautics, KTH, SWEDEN

P. Eliasson
FOI, SWEDEN

Keywords: *compressible CFD, convergence acceleration, implicit time-stepping, matrix dissipation*

Abstract

Edge is a flow solver for unstructured grids based on a dual grid and edge-based formulation. The standard dual-time stepping methods for compressible unsteady flows are inadequate for large-scale industrial problems. This has motivated the present work, in which an implicit Lower-Upper Symmetric Gauss-Seidel (LU-SGS) type of relaxation has been implemented in the code Edge with multigrid acceleration. Two different types of dissipation, a scalar and a matrix model, have been constructed which increase the diagonal dominance of the system matrix but not the numerical viscosity of the computed solution. A parametric study demonstrates convergence accelerations by a factor of three for inviscid transonic flows compared to explicit Runge-Kutta smoothing for multigrid acceleration.

1 Introduction

Convergence acceleration techniques are needed to speed up Computational fluid dynamics (CFD) solvers, which have a major role in industries and research agencies for aerodynamic design and optimization of aircraft. The flow is often characterized by high-speed and compressible flow features and, in many cases, by massive flow separation and unsteadiness. Accurate and efficient numerical solution by means of time-dependent computations is hence required. The efficiency of standard dual-time stepping methods used for unsteady flows in many CFD codes is inadequate for large-scale industrial problems.

Edge [1] is the Swedish national CFD tool for compressible flow, used at the Swedish aircraft manufacturer SAAB, and developed at FOI, lately in collaboration with external national and international partners. Edge uses an edge-based formulation on unstructured node-centered finite-volume grids based on a dual grid supplied by a preprocessor. Multigrid acceleration has been implemented in Edge as a strategy to speed up the convergence rate [1] based on the FAS (Full Approximation Scheme) [2] as a nonlinear approach. It treats the nonlinear equation system using an explicit Runge-Kutta time stepping [1] as smoother. The Line-Implicit method of acceleration in Edge combines the explicit multistage Runge–Kutta scheme with an implicit integration along lines in regions where the computational grid is highly stretched. Implicit methods can efficiently handle problems with disparate scales, since they have no stability restriction on

the time-step. See e.g. [3], for a detailed analysis. Therefore, implicit methods are widely used to accelerate convergence of steady-state problems as well as to improve the efficiency of unsteady solvers by advancing the solution with substantially larger time (or pseudo-time) steps. This has motivated the present work to speed up the computations by implicit time-stepping using iterative solvers such as LU-SGS [4][5][6][7]. The work is devoted to the definition of a proper block matrix implicit operator. Increasing the diagonal blocks of the system matrix is beneficial for convergence of the linear iteration [7]. Hence, linearizations which produce the coefficients in the linear system Left Hand Side (LHS), have been implemented, based on two different types of implicit operators, a scalar and a matrix model. Suitable combinations of LHS and Right Hand Side (RHS) dissipation models show significant convergence acceleration compared to explicit schemes. Since several parameters are important for controlling the convergence, an extensive parametric study has been conducted.

2 Governing Equations

Edge solves the compressible Reynolds-Averaged Navier-Stokes (RANS) equations expressed as in Eq. (1),

$$\frac{\partial \boldsymbol{U}}{\partial t} + \nabla . \boldsymbol{F}_I + \nabla . \boldsymbol{F}_V = \boldsymbol{Q} \qquad (1)$$

where $\boldsymbol{U} = (\rho \quad \rho u \quad \rho v \quad \rho w \quad \rho E)$ is the vector of conserved variables, $\boldsymbol{F}_I$ and $\boldsymbol{F}_V$ the inviscid and viscous flux matrices respectively, and $\boldsymbol{Q}$ the source vector. Equation (2) shows the corresponding integral form for a control volume Ω contained in the surface S.

$$\int_\Omega \frac{\partial \boldsymbol{U}}{\partial t} d\Omega + \int_S \left(\boldsymbol{F}_I + \boldsymbol{F}_V \right) \cdot \boldsymbol{n} dS = \int_\Omega \boldsymbol{Q} d\Omega \qquad (2)$$

A finite-volume discretization, Eq. (3), is obtained by applying the integral formulation to the control volume surrounding the node v_0, as illustrated in Fig. 1.

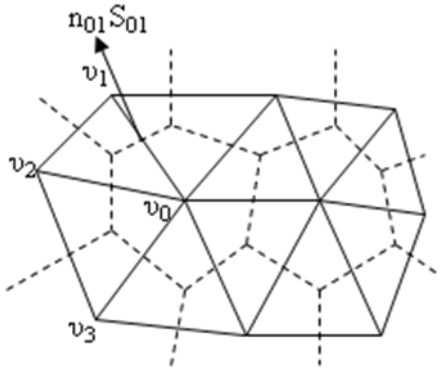


Fig. 1 Dual grid forms control volume

$$\frac{d}{dt}\left(\boldsymbol{U}_0 V_0\right) + \sum_{k=1}^{m_0} \boldsymbol{F}_{I_{0k}} \boldsymbol{n}_{0k} S_{0k} + \sum_{k=1}^{m_0} \boldsymbol{F}_{V_{0k}} \boldsymbol{n}_{0k} S_{0k} = \boldsymbol{Q}_0 V_0 \qquad (3)$$

with m_0 the number of neighbors and S_{0k} the surfaces enclosing the control volume for node v_0. $\boldsymbol{F}_{I_{0k}}$ and $\boldsymbol{F}_{V_{0k}}$ are the inviscid and viscous flux vectors respectively, computed on the edge connecting nodes v_0 and v_k.

This paper focuses on the unsteady (2D) Euler equations, Eq. (4), where the viscous flux is neglected.

$$\frac{d}{dt}\left(\boldsymbol{U}_0 V_0\right) + \sum_{k=1}^{m_0} \boldsymbol{F}_{I_{0k}} \boldsymbol{n}_{0k} S_{0k} = 0 \qquad (4)$$

3 Dissipation models

Second order schemes are non-dissipative, so dissipation must be added to stabilize the time-stepping. Edge considers two schemes [1] for the inviscid flux $\boldsymbol{F}_{I_{0k}}$, both based on a central discretization, with dissipation terms of either Jameson artificial dissipation type or upwind flux difference splitting type.

3.1 Jameson flux approximation

Equation (5) shows the inviscid flux computation across the cell face between nodes v_0 and v_1 using variable averaging,

$$\boldsymbol{F}_{I_{01}} = \boldsymbol{F}_I\left(\frac{\boldsymbol{U}_0 + \boldsymbol{U}_1}{2}\right) - \boldsymbol{d}_{01} \qquad (5)$$

The construction of the artificial dissipation term $\boldsymbol{d}_{01}$, Eq. (6), is based on the Jameson pressure switch in the ε-factors, which activate the second order difference dissipation and turn off fourth order difference dissipation in the vicinity of shocks.

$$\begin{aligned}\boldsymbol{d}_{01} = &\left(\varepsilon_{01}^{(2)}\left(\boldsymbol{U}_1 - \boldsymbol{U}_0\right)\right. \\ &\left. - \varepsilon_{01}^{(4)}\left(\nabla^2\boldsymbol{U}_1 - \nabla^2\boldsymbol{U}_0\right)\right)\varphi_{01}\lambda_{01}\end{aligned} \qquad (6)$$

The factor φ_{01} is introduced to account for the stretching in the grid. This scheme is called scalar because of the scalar coefficient of the variable differences, the local spectral radius λ_{01} (see Eq. (7)).

$$\lambda_{01} = \left(\left|\boldsymbol{u}_{01}.\boldsymbol{n}_{01}\right| + c_{01}\right)S_{01} \qquad (7)$$

where $\boldsymbol{u}_{01}$, c_{01}, S_{01}, and $\boldsymbol{n}_{01}$,denote the flow velocity, speed of sound, area, and normal of the cell face, respectively.

Since the spectral radius corresponds to the maximum eigenvalue, the Jameson model is known to be very dissipative near the shocks. On the other hand it is characterized by its simplicity, a controllable amount of artificial dissipation, and robust convergence.

3.2 Upwind scheme

The upwind scheme is of Roe flux difference splitting type. The inviscid flux formulation can be expressed as for the basic central scheme but with a flux averaging (Eq. (8)).

$$\boldsymbol{F}_{I_{01}} = \frac{1}{2}\left(\boldsymbol{F}_I\left(\boldsymbol{U}_0\right) + \boldsymbol{F}_I\left(\boldsymbol{U}_1\right)\right) - \boldsymbol{d}_{01} \qquad (8)$$

The upwind dissipation $\boldsymbol{d}_{01}$ uses the "Roe" matrix $\boldsymbol{A}_R = \boldsymbol{A}(\tilde{\boldsymbol{U}})$, $\boldsymbol{A}_R = \boldsymbol{R}\,\boldsymbol{\Lambda}\,\boldsymbol{R}^{-1}$ where $\tilde{\boldsymbol{U}}$ is the Roe average of the states $\boldsymbol{U_0}$ and $\boldsymbol{U}_1$, as shows Eq. (9).

$$\boldsymbol{d}_{01} = \frac{1}{2}\boldsymbol{R}\left|\boldsymbol{\Lambda}^*\right|\boldsymbol{R}^{-1}\left(\boldsymbol{U}_1 - \boldsymbol{U}_0\right) \qquad (9)$$

The matrix $\boldsymbol{R}$ contains the right eigenvectors of the flux Jacobian. The diagonal matrix $\boldsymbol{\Lambda}^*$ contains the eigenvalues adjusted with an entropy fix, as shown in Eq. (10), to prevent expansion shocks where an eigenvalue vanishes.

$$\left|\lambda_i\right|^* = \begin{cases} \dfrac{\lambda_i^2 + \delta^2}{2\delta}, & \left|\lambda_i\right| \le \delta, \\ \left|\lambda_i\right|, & \left|\lambda_i\right| > \delta, \end{cases} \qquad (10)$$

δ is taken as a small fraction of the spectral radius, usually around 5%. This expression is applied on the linear and non-linear eigenvalues independently, by two different δ variables.

Moreover, a second order Upwind scheme is available in Edge by applying limiters to the absolute eigenvalue matrix $\left|\boldsymbol{\Lambda}^*\right|$ from Eq. (9). Despite its slower convergence compared to the Jameson scheme, this method is preferable for its better accuracy.

3.3 Matrix dissipation

Matrix Dissipation [8] [9] is a generalization of the two previous dissipation models. This model offers a combination of high accuracy around the shocks and good convergence, coming from the Upwind and Jameson scheme respectively.

The matrix dissipation replaces the spectral radius λ_{01} by the absolute Roe matrix $\boldsymbol{R}\left|\boldsymbol{\Lambda}^*\right|\boldsymbol{R}^{-1}$

from Eq. (9). Equation (11) shows the final expression of this combination.

$$\boldsymbol{d}_{01} = \left(\varepsilon_{01}^{(2)}\left(\boldsymbol{U}_1 - \boldsymbol{U}_0\right)\boldsymbol{R}\left|\boldsymbol{\Lambda}^*\right|\boldsymbol{R}^{-1} - \varepsilon_{01}^{(4)}\left(\nabla^2\boldsymbol{U}_1 - \nabla^2\boldsymbol{U}_0\right)\lambda_{01}\right)\varphi_{01} \qquad (11)$$

This formulation offers dissipation control through the entropy fix parameters of the Roe matrix. One can notice that the strong dissipation of the spectral radius is kept in the smooth regions.

4 Implicit time-stepping

A non-linear steady state problem for inviscid flow is solved by time marching the unsteady Euler equations [1] from Eq. (4). The resulting equations in Eq. (12) are discretized in time by an implicit Backward Euler ("BE") method. Equation (13) corresponds to the final discretized system.

$$V.\frac{d\boldsymbol{U}}{dt} + \boldsymbol{R}(\boldsymbol{U}) = 0 \qquad (12)$$

$$V.\frac{\boldsymbol{U}^{n+1} - \boldsymbol{U}^n}{\Delta t} + \boldsymbol{R}(\boldsymbol{U}^{n+1}) = 0 \qquad (13)$$

After linearization of Eq. (13),

$$\boldsymbol{R}(\boldsymbol{U}^{n+1}) \approx \boldsymbol{R}(\boldsymbol{U}^n) + \frac{\partial \boldsymbol{R}(\boldsymbol{U}^n)}{\partial \boldsymbol{U}}\Delta\boldsymbol{U}^n \qquad (14)$$

the linear system Eq. (15) appears and is to be solved approximately.

$$\underbrace{\left(\frac{1}{\Delta t}V + \frac{\partial \boldsymbol{R}(\boldsymbol{U}^n)}{\partial \boldsymbol{U}}\right)}_{\boldsymbol{M}}\underbrace{\Delta\boldsymbol{U}^n}_{\boldsymbol{x}} = \underbrace{-\boldsymbol{R}\left(\boldsymbol{U}^n\right)}_{\boldsymbol{b}} \qquad (15)$$

LHS RHS

with,

$$\Delta\boldsymbol{U}^n = \boldsymbol{U}^{n+1} - \boldsymbol{U}^n$$
$$\boldsymbol{R}_i = \sum_{Edges}\boldsymbol{F}_i$$

The LHS and RHS term definitions are in the core of the paper, controlling convergence and accuracy respectively. A special effort is made to separate the parameters affecting the LHS from the ones related to the RHS. Hence the work can be focused in the increase of the diagonal dominance without affecting the accuracy of the solution.

5 LU-SGS linear solver

The linear system has a sparse coefficient matrix, but direct elimination methods require excessive computing resources for multi-million grid-point geometries.

5.1 Algorithm Description

The **L**ower-**U**pper **S**ymmetric **G**auss-**S**eidel (**LU-SGS**) linearization technique [4] solves the system very approximately but has small memory requirement. The challenge is in reducing the computational time by increasing the convergence rate through the definition of a proper approximation to the exact matrix $\boldsymbol{M}$.

The algorithm [7] is based on a number of Forward and Backward sweeps, Eq. (16), where $\boldsymbol{L}$ is the block lower triangular part, $\boldsymbol{D}$ the block diagonal, and $\boldsymbol{U}$ the block upper triangle of $\boldsymbol{M}$. In the current implementation, a block is the set of unknowns per node.

$$\begin{aligned}\boldsymbol{M}\boldsymbol{x} &= \left(\boldsymbol{L}+\boldsymbol{D}+\boldsymbol{U}\right)\boldsymbol{x} = \boldsymbol{b}\\ \left(\boldsymbol{D}+\boldsymbol{L}\right)\boldsymbol{x}^* &= \boldsymbol{b} - \boldsymbol{U}\boldsymbol{x}^n\\ \left(\boldsymbol{D}+\boldsymbol{U}\right)\boldsymbol{x}^{n+1} &= \boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^*\end{aligned} \qquad (16)$$

A fixed number of sweeps is performed in a time-step. The time marching process continues till the residual of the non-linear system vanishes as in Eq. (17).

$$\left\| \boldsymbol{R}(\boldsymbol{U}^n) \right\| \le tol \qquad (17)$$

Note that $\boldsymbol{M}$ controls the convergence rate but does not influence the converged solution.

$\boldsymbol{M}$ should be a sufficiently accurate approximation to the exact derivative to make the linearization accurate, and for convergence of the sweeps, it should be block-diagonally dominant. Since the diagonal elements come from the artificial dissipation model, the two requirements conflict. The rest of the paper focuses on the choice and construction of matrix $\boldsymbol{M}$ for convergence acceleration.

5.2 LHS dissipation models

The linearizations, i.e. the coefficients in the linear system Left Hand Side, are based on two different types of implicit operators, a scalar and a matrix model.

5.2.1 Scalar dissipation

This first implementation of the implicit operator (LHS) is based on the Yoon and Kwak [10] flux Jacobian splitting as in Eq. (18),

$$\boldsymbol{A}^{\pm} = \frac{1}{2}\left(\boldsymbol{A} \pm \boldsymbol{I}\beta\lambda_{max}\right) \qquad (18)$$

with $\boldsymbol{A} = \dfrac{\partial \boldsymbol{F}_I}{\partial \boldsymbol{U}}$ the flux Jacobian matrix, $\lambda_{\max} = \left(\left|\boldsymbol{u}.\boldsymbol{n}\right| + \boldsymbol{c}\right)$ the spectral radius of $\boldsymbol{A}$ and β a constant parameter to adapt. The flux function considered for the LHS can be expressed by Eq. (19) for node v_i and its neighbor v_n, with the corresponding splitting when deriving it, in Eq. (20).

$$\begin{aligned} \boldsymbol{F}_{i,n} = \frac{1}{2}\big(&\boldsymbol{F}(\boldsymbol{U}_n) + \boldsymbol{F}(\boldsymbol{U}_i) \\ &- \beta\lambda_{max}(\boldsymbol{U}_n - \boldsymbol{U}_i)\big) \end{aligned} \qquad (19)$$

$$\begin{aligned} \frac{\partial \boldsymbol{F}(\boldsymbol{U}_i,\boldsymbol{U}_n)}{\partial \boldsymbol{U}_n} &= \frac{1}{2}\left(\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{U}}\right)_n - \boldsymbol{I}\beta\lambda_{max}\right) \\ \frac{\partial \boldsymbol{F}(\boldsymbol{U}_i,\boldsymbol{U}_n)}{\partial \boldsymbol{U}_i} &= \frac{1}{2}\left(\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{U}}\right)_i + \boldsymbol{I}\beta\lambda_{max}\right) \end{aligned} \qquad (20)$$

5.2.2 Matrix dissipation

This extension [11] is based on the previous splitting method with the only difference that the flux derivative corresponds to the derivative of the exact RHS of Eq (8). This means that the Jacobian matrix is now expressed as in Eq. (21).

$$\boldsymbol{A}^{\pm} = \frac{1}{2}\left(\boldsymbol{A} \pm \left|\boldsymbol{A}^{''}\right|\right) \qquad (21)$$

with $\left|\boldsymbol{A}^{''}\right| = \boldsymbol{R}\left|\boldsymbol{\Lambda}^{''}\right|\boldsymbol{R}^{-1}$ the absolute Roe matrix, and that the derivatives are expressed by Eq. (22).

$$\begin{aligned} \frac{\partial \boldsymbol{F}(\boldsymbol{U}_i,\boldsymbol{U}_n)}{\partial \boldsymbol{U}_n} &= \frac{1}{2}\left(\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{U}}\right)_n - \left|\boldsymbol{A}^{''}\right|_{i,n}\right) \\ \frac{\partial \boldsymbol{F}(\boldsymbol{U}_i,\boldsymbol{U}_n)}{\partial \boldsymbol{U}_i} &= \frac{1}{2}\left(\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{U}}\right)_i + \left|\boldsymbol{A}^{''}\right|_{i,n}\right) \end{aligned} \qquad (22)$$

ignoring the derivatives of the $\boldsymbol{A}^{''}$ matrix. The new notation for the eigenvalue matrix $\boldsymbol{\Lambda}^{''}$ instead of $\boldsymbol{\Lambda}^{*}$ as seen in Eq. (9), underlines the possibility of acting on the dissipation matrices of the LHS and RHS independently.

6 Numerical results

Several simulations have been run to assess the influence of the multitude of parameters in the matrix dissipation LU-SGS scheme. The accelerated methods are compared to the present standard of FAS Multigrid with explicit Runge-Kutta smoothing. All the cases have been computed for transonic inviscid flow over a NACA 0012 airfoil, on an inviscid mesh and a

mesh for RANS simulation. The accuracy required for the computations has needed residuals of 10^{-10}.

6.1 Inviscid mesh

The first cases have been computed on an inviscid mesh for inviscid transonic Mach 0.8 flow at 1.25º angle of attack.

The results show the influence of different parameters on the convergence rate compared to explicit schemes, and gives a better comprehension of the models used. The parametric study is based on a first order upwind scheme for the RHS.

6.1.1 LHS with scalar dissipation

Figure 2 shows convergence with full multigrid for *CFL* in $(10,10^5)$ and the number of sweeps in a time-step Nswp set to 5. The number of iterations is reduced by a factor of 2.5 compared to the explicit Runge-Kutta scheme for tolerances in 10^{-4} … 10^{-10}. The rate is not very sensitive to the choice of *CFL* in (100,10000).

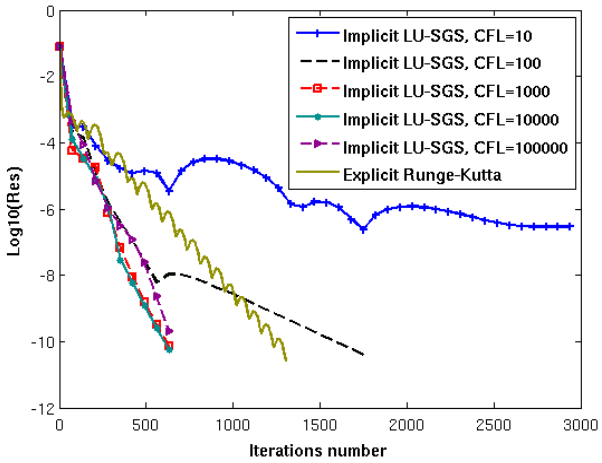


Fig. 2 Convergence dependence on CFL

Figure 3 shows convergence for different sweep numbers (Nswp) and CFL=1000. There are two distinct phases, the initial for "large" residuals > 10^{-5}, and the asymptotic phase with very small residuals. One may surmise that the non-linearities are negotiated in the first phase and in the latter phase, the residual is very close to a linear function so the scheme is actually an iterative solution of a linear system. The initial convergence seems is clearly faster for larger Nswp, but the ultimate convergence rate seems only slightly influenced by Nswp. In all of these three cases, the number of iterations is reduced by a factor of 1.5 compared to Runge-Kutta.

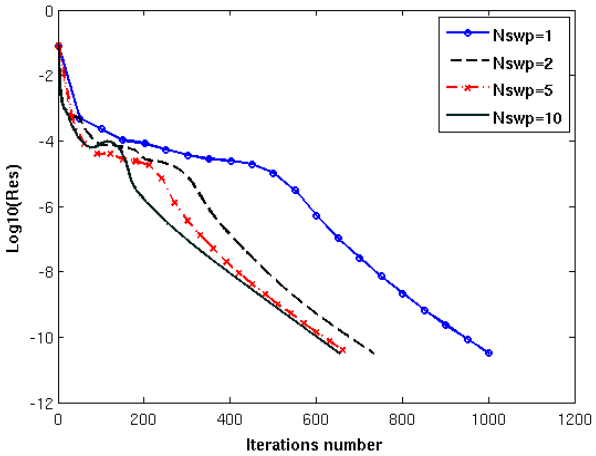


Fig. 3 Convergence dependence on number of sweeps Nswp

The coefficient β from Eq. (18) has been varied to find that the value of 1 giving full upwinding in the system matrix is optimal. $\beta > 1$ causes a convergence degradation.

For design calculations, a second order accurate scheme is necessary. In Fig. 4, LU-SGS shows more rapid convergence for a 1st order upwind residual scheme compared to a 2nd order scheme with its smaller dissipation, as expected. Moreover one observes increased convergence rate after reaching a residual of the order of 10^{-5} (see also Fig. 2) contrary to the explicit Runge-Kutta which keeps a linear convergence. Additional computations show that this LU-SGS behavior disappears when the multigrid acceleration is switched off.

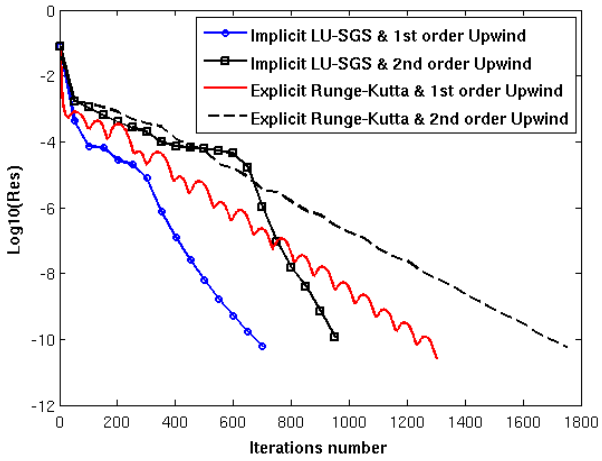


Fig. 4 Influence of space discretization scheme

6.1.2 LHS with matrix dissipation

Similar observations have been found using matrix dissipation in the LHS. The Entropy fix control described in Eq. (10) seems not to have much effect on the convergence rate, except for mitigating the *CFL* number limitation. As observed in Fig. 5, an entropy fix of 90% is needed to reach a *CFL* of 10000. Bigger values decrease the convergence as noticed for the LHS scalar dissipation.

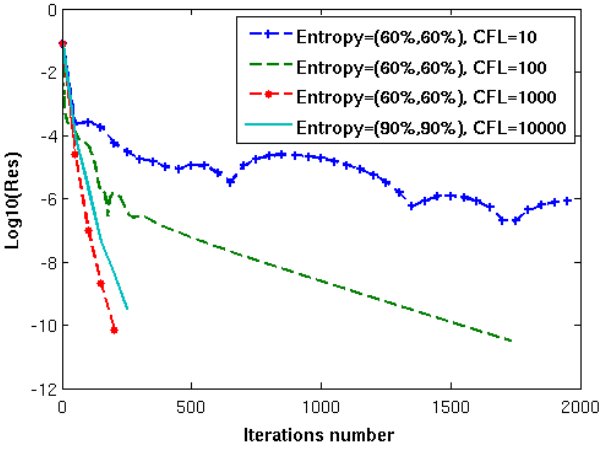


Fig. 5 Entropy limiting CFL

6.1.3 Synthesis

Figures 6 to 8 compare the convergence accelerations using the "optimal" LHS dissipation operator, for each RHS scheme, namely Jameson, Matrix, and 2nd order Upwind scheme. The Explicit Runge-Kutta shows in all the cases a convergence rate twice slower. The computations have been run for a *CFL* of 1000 and Nswp=2.

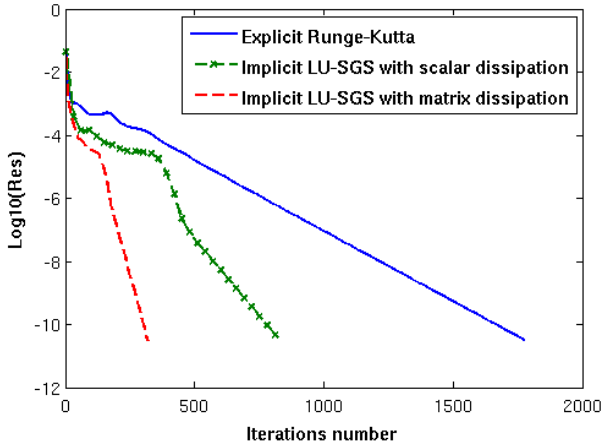


Fig. 6 RHS Jameson dissipation

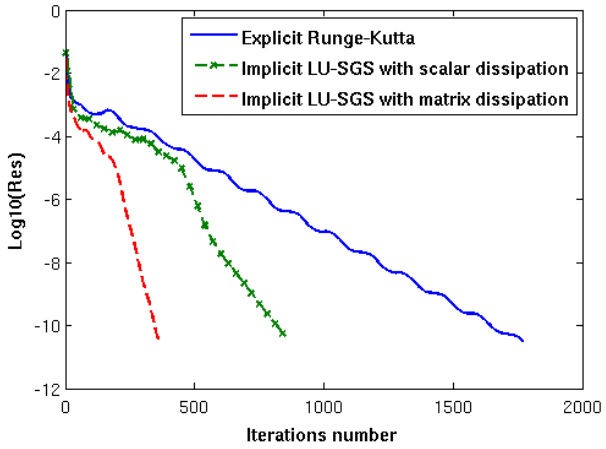


Fig. 7 RHS Matrix dissipation

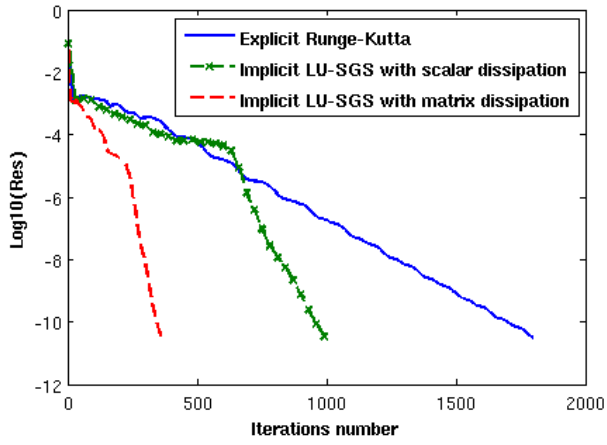


Fig. 8 RHS 2nd order Upwind scheme

6.2 Stretched mesh for RANS flow

As a transition to viscous flow simulations, inviscid transonic Mach 0.754 flow at 2.57º angle of attack, on a mesh for RANS simulation has been computed. Figure 9 gives an example of computation for a matrix dissipation LHS and 1st order Upwind scheme RHS. The comparison is done with the line-implicit method.

The grid has a normal distance from the airfoil to the first interior node of 10^{-5} chord length, about 310 nodes on the airfoil, and a varying number of quadrilateral cells in the wall normal direction with about 35 layers in average. The triangular grid outside of these layers has about 51 K nodes. The maximum aspect ratio of the stretching is 1.5×10^{3}. Four levels of full multigrid W-cycles are used in all calculations. Directional coarsening with coarsening ratios 1:4 or 1:2 along the lines are used.

Previous computations have shown a negligible gain by increasing the *CFL* number beyond 10^{4}. For an optimized configuration with CFL=10000, implicit LU-SGS converges twice faster than the line-implicit method.

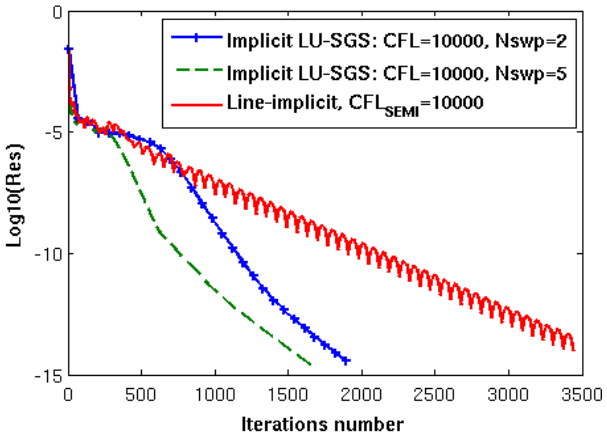


Fig. 9 Convergence on a stretched grid

7 Conclusions

An implicit LU-SGS method has been successfully implemented reducing the iteration count by a factor of 3 compared to the explicit schemes. Two LHS dissipation models of scalar and matrix types have been constructed to improve the diagonal dominance of the system matrix. First computations on inviscid mesh have revealed optimal configurations corresponding to a *CFL* of 10000 and matrix LHS dissipation. More study is necessary on stretched grids for an extension into RANS and turbulence models. There remains also efficient implementation on computational clusters.

References

[1] "Edge - Theoretical Formulation", Defence and Security, Systems and Technology, FOI, 2010.

[2] "Study Note on Multigrid Methods", Free University, Brussels and FFA Sweden, 1990.

[3] Eliasson P., Weinerfelt P., and Nordström J., "Application of a line-implicit scheme on stretched unstructured grids," *47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition*, Orlando, 2009.

[4] Kim J.S. and Kwon O.J., "An efficient implementation of implicit operator for block LU-SGS method", *Computational Fluid Dynamics*, 2005, pp.154–159.

[5] Yoon S. and Kwak D., "Implicit methods for the Navier-Stokes equations", *Computing Systems in Engineering*, Vol. 1, No. 2-4, 1990, pp. 535-547.

[6] Luo H., Sharov D., Baum J. D. and Löhner R. , "Parallel unstructured grid GMRES+LU-SGS method for turbulent flows ", *American Institute of Aeronautics and Astronautics*, 2003.

[7] Dwight R.P, "Efficiency improvements of RANS-based analysis and optimization using implicit and adjoint methods on unstructured grids", Ph.D. Dissertation, University of Manchester, 2006.

[8] Anker J.E. and Mayer J.F., "A preconditioned solution scheme for the simulation of turbomachinery flow at arbitray mach numbers", *42nd AIAA Aerospace Sciences Meeting and Exhibit*, Reno, 2004.

[9] Otero E., "Increasing the Efficiency of the CFD Tool Edge for Aircraft Design", *Aerodays 2011*, Madrid.

[10] Grismer M.J., Strang W.Z., Tomaro R.F. and Witzeman F.C, "Cobalt: A parallel, implicit, unstructured Euler/Navier–Stokes solver", *Advances in Engineering Software*, Vol. 29, No. 3-6, 1998, pp. 365-373.

[11] Nejat A. and Ollivier-Gooch C. "A High-Order Accurate Unstructured GMRES Algorithm for Inviscid Compressible Flows", *American Institute of Aeronautics and Astronautics*.

CEAS 2011

Aeroelasticity

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Structural and Aeroelastic Modeling of Curved Rotor Blades Using a Galerkin Approach

E. Piccione, G. Bernardini, M. Molica Colella and M. Gennaretti
Department of Mechanical and Industrial Engineering
University Roma Tre, Italy

Keywords: curved beams, rotary beams, aeroelasticity

Abstract

The aim of this work is the development of a numerical solver for structural and aeroelastic analysis of rotor blades with advanced geometry, like those with tip sweep and anhedral angles. The mathematical formulation is based on a beam-like model assumption, suited for numerical integration via modal approach. It is valid for slender, homogeneous, isotropic blades with curved and twisted elastic axis. Inertial loads due to rotary motion and blade deformation are included as distributed loads, as well as, the aerodynamic loads in the aeroelastic analysis. After application of a second order approximation scheme, the final nonlinear equations of motion are able to predict the dynamics of blade subject to moderate displacements. The model is conceived as a sequence of relations that yield the loads to be included directly in the equations of section moments and allows both response and stability analysis following the procedure outlined. Comparisons with results from FEM analysis will be shown for validation of the model implemented.

1 Introduction

The aim of this work is the development of a structural and aeroelastic model for rotors with advanced blade geometry characterized by curved elastic axis (for instance, with inclusion of tip sweep and anhedral angles).

Rotors with curved blades are used in many engineering applications, a large number of which are present in the aeronautical field. Indeed, nowadays a primary goal in aircraft and rotorcraft design is the reduction of the environmental impact in order to increase the community acceptance. Consequently, during last years particular attention has been focused on the reduction of emitted noise. Helicopter main rotors and tiltrotor proprotors provide lifting force and thrust, but, as a by-product, they are also sources of noise and vibrations. Blade-vortex interaction (BVI) noise generated by helicopter main rotors and proprotors during some flight conditions is one of the most severe noise problems for rotorcraft and it has been carefully studied during last years by researcher and manufacture companies (see, for instance, Refs. [1, 3]). The blade tip vortex structures, blade aeroelastic deformation, blade-vortex miss distance have been identified as the most important parameters that influence BVI (and vibrations) [4].

In order to reduce BVI noise, several blade shapes and planform configurations are currently under investigation by manufacturers and researchers and, in particular, blades with tip sweep and anhedral angles seem to be effective in reducing BVI noise. Hence, the development of theoretical and numerical models for the prediction of the structural dynamics and aeroelas-

tic behavior of advanced geometry rotor blades is a fundamental tool in the optimal design of blades oriented to noise and vibration reduction. Helicopter and tiltrotor blades are flexible, light, slender structures, and therefore prediction tools aimed at their structural/aeroelastic analysis have to be able to take into account both strong coupling between bending and torsion degrees of freedom and the nonlinearities arising from the significant deflections they usually experience. In the past years, several authors have developed structural mathematical models for curved blade, usually solved through the finite element method approach (see, for instance, Refs. [5, 6]).

Here, a structural, beam-like mathematical model suited for the analysis of curved blades is developed through application of the Galerkin approach, in that well suited for aeroelastic applications. It is valid for slender, homogeneous, isotropic rotating blades with curved and twisted elastic axis, which permits the simulation of arbitrary advanced blade shape. Specifically, differently from usual formulations based on the Galerkin approach, the test functions are the second-order spatial derivative of the trial ones and the equations projected are directly the integro-differential ones giving the sectional forces and moments.

Numerical investigation presents the validation of the presented formulation for structural dynamics and aeroelastic analysis of a simple curved blade, against results from the FEM solver presented and validated in Ref. [7].

2 Blade structural dynamics model

Here, a differential model governing the structural dynamics of rotating curved beams is outlined. As it will be shown later, it is suited for the development of a numerical solver based on the Galerkin approach, and hence is particularly useful in aeroelastic applications.

The differential model is valid for slender, homogeneous, isotropic rotating blades with curved and twisted elastic axis and includes spanwise variation of mass and stiffness properties, as well as, variable built-in pretwist, precone, sweep and anhedral angles. In order to take into account the moderate displacements usually experienced by rotor blades, nonlinear strain-displacement relations are considered with application of a second order approximation scheme.

2.1 Displacement variables and coordinate systems

Several coordinate systems are introduced to derive the equations of motion of the blade. The main ones, illustrated in Fig. 1, are the following:

- a global blade orthogonal system of unit vectors ($\vec{i}_1$, $\vec{i}_2$, $\vec{i}_3$) which is centered at the rotor hub and rotates with the blade, with $\vec{i}_1$ tangent to the elastic axis at the root section; vectors $\vec{i}_2$ and the $\vec{i}_3$ identify the principal axes of the blade root section; note that the orientation of $\vec{i}_2$ and $\vec{i}_3$ depends on both collective and cyclic blade pitch command;

- local, rotating orthogonal systems of unit vectors $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$, with $\vec{e}_1$ aligned to the undeformed blade elastic axis, and $\vec{e}_2, \vec{e}_3$ aligned with the blade section principal axes; note that the orientation of $\vec{e}_2$ and $\vec{e}_3$ depends on blade pretwist;

- local, rotating orthogonal systems related to the blade deformed configuration having base unit vectors $(\vec{\hat{e}}_1, \vec{\hat{e}}_2, \vec{\hat{e}}_3)$, with $\vec{\hat{e}}_1$ tangent to the deformed elastic axis, and $\vec{\hat{e}}_2, \vec{\hat{e}}_3$ aligned with the principal axes of the elastically twisted blade sections (this assumption means that the sections of the deformed beam remains orthogonal to the deformed elastic axis and hence, a shear undeformable beam model is considered).

Deformations are described in terms of displacements of the elastic axis and rotation of beam sections. The displacements, u, v, w, are defined in the local frames fixed with the undeformed blade, respectively along the directions identified by $\vec{e}_1, \vec{e}_2$ and $\vec{e}_3$; the blade twist, ϕ, is defined as the rotation of the blade sections about $\vec{\hat{e}}_1$-direction (*i.e.,* about the deformed elastic axis).

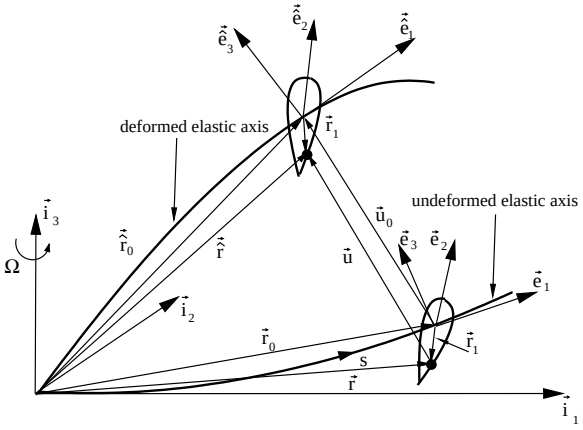


Figure 1: Schematic beam representation

2.1.1 Coordinate systems transformations

In the development of the blade dynamics formulation it is now convenient to define the transformations relating the frame of reference defined above.

Global to undeformed blade references
The unit base vectors of the global rotating frame are related to those of the local undeformed blade frames through the following expression

$$\{\vec{e}_i\} = \mathbf{A}(s) \left\{\vec{i}_i\right\}$$

where, for

$$\{\vec{e}_i\} = \left\{ \begin{array}{c} \vec{e}_1 \\ \vec{e}_2 \\ \vec{e}_3 \end{array} \right\}$$

and s denoting the curvilinear coordinate defined along the undeformed elastic axis, $\mathbf{A}(s)$ is the transformation matrix yielding the local frame unit vectors in terms of superposition of the global frame base (i.e., it collects the components of each $\vec{e}_i$ in the global frame).

Blade undeformed to blade deformed references
Local undeformed blade and deformed blade references are related by the following expression

$$\left\{\hat{\vec{e}}_i\right\} = \mathbf{T}(s) \left\{\vec{e}_i\right\}$$

where $\mathbf{T}$ is the transformation matrix depending on the deformation of the blade. The matrix $\mathbf{T}$ is obtained as a sequence of rotations about $\vec{e}_3$, rotated $\vec{e}_2$ and rotated $\vec{e}_1$ and its expression in terms of u, v, w, ϕ may be derived from that given in Ref. [8].

Global to blade deformed references
From the transformations mentioned above it is possible to determine the relation between global and deformed blade references as

$$\left\{\hat{\vec{e}}_i\right\} = \mathbf{\Lambda} \left\{\vec{i}_i\right\}$$

where $\mathbf{\Lambda} = \mathbf{TA}$.

2.2 Equilibrium relations

Considering a deformed beam element of length $\mathrm{d}s$, the equilibrium of forces and moments acting on it yields

$$\frac{\mathrm{d}\vec{v}}{\mathrm{d}s} + \vec{p} = 0 \tag{1}$$

$$\frac{\mathrm{d}\vec{m}}{\mathrm{d}s} + \hat{\vec{e}}_1 \times \vec{v} + \vec{q} = 0 \tag{2}$$

where $\vec{v}$ and $\vec{m}$ are the internal structural forces and moments at the elastic axis, while $\vec{p}$ and $\vec{q}$ are the external distributed forces and moments.

In order to integrate the above differential equilibrium equations, it is convenient to write them in terms of forces and moments components, $\mathbf{v}_l, \mathbf{m}_l, \mathbf{p}_l, \mathbf{q}_l$, in the local blade undeformed frames. Observing that for $\mathbf{v}_g, \mathbf{m}_g, \mathbf{p}_g, \mathbf{q}_g$ denoting forces and moments components in the global frame of reference one has $\mathbf{v}_g = \mathbf{A}^T \mathbf{v}_l, \mathbf{m}_g = \mathbf{A}^T \mathbf{m}_l, \mathbf{p}_g = \mathbf{A}^T \mathbf{p}_l, \mathbf{q}_g = \mathbf{A}^T \mathbf{q}_l$, the equilibrium equations projected onto the frame $\left\{\vec{i}_i\right\}$ yield

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{A}^T\mathbf{v}_l\right) + \mathbf{A}^T\mathbf{p}_l = 0 \tag{3}$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathbf{A}^T\mathbf{m}_l\right) - \mathbf{A}^T\mathbf{H}_l\mathbf{v}_l + \mathbf{A}^T\mathbf{q}_l = 0 \tag{4}$$

where $\mathbf{H}_l$ is the matrix of the components in the local undeformed frame of the axial tensor associated to vector $\hat{\vec{e}}_1$.

Next, integration of Eq. 3 yields the following distribution of the internal shear loads

$$\mathbf{v}_l(s) = \mathbf{A} \int_s^R \mathbf{A}^T \mathbf{p}_l \, \mathrm{d}\hat{s} \tag{5}$$

while the integration of Eq. 4 yields the following distribution of the internal moments

$$\mathbf{m}_l(s) = \mathbf{A}\left\{\int_s^R \mathbf{A}^T\mathbf{q}_l\,\mathrm{d}\hat{s} - \int_s^R\left(\mathbf{A}^T\mathbf{H}_l\mathbf{A}\int_{\hat{s}}^R \mathbf{A}^T\mathbf{p}_l\,\mathrm{d}\tilde{s}\right)\mathrm{d}\hat{s}\right\} \tag{6}$$

where R denotes the length of the undeformed elastic axis (under assumption of negligible second order terms related to local slope).

Equations 5 and 6 are the general solutions for the internal shear loads and moments arising in a beam, from which the equations governing the blade elastic displacement may be derived, once strain-displacement and load-displacement relations are identified (see next paragraphs). Note that, because of shear undeformable assumption, the equations governing the blade motion variables, u, v, w and ϕ, are derived from the first scalar equation in Eq. 5 and the three scalar equations in Eq. 6 (the second and third scalar equations in Eq. 5 are used to determine the shear loads components lying in the plane of the beam sections).

2.3 Strain-displacement relations

In order to express the internal loads in terms of the (shear undeformable) beam deformation variables, u, v, w and ϕ, the strain tensor has to be derived from position vectors describing undeformed and deformed beam.

Specifically, the vector position, $\vec{r}$, of a point of the undeformed beam of coordinates (s, η, ζ), with η and ζ denoting the coordinates along the principal axes of the sections, is given by

$$\vec{r}(s,\eta,\zeta) = \vec{r}_0 + \eta\vec{e}_2 + \zeta\vec{e}_3$$

while after deformation, under the assumption of negligible warping, its position is identified by the following vector

$$\vec{\hat{r}}(s,\eta,\zeta) = \vec{r}_0 + u\vec{e}_1 + v\vec{e}_2 + w\vec{e}_3 + \eta\vec{\hat{e}}_2 + \zeta\vec{\hat{e}}_3$$

where $\vec{r}_0$ denotes the position of the points along the undeformed elastic axis.

From the above equations the strain tensor, $\vec{E}$, is derived from

$$\vec{E}(s,\eta,\zeta) = \frac{1}{2}\left(\frac{\partial\vec{\hat{r}}}{\partial\xi^i}\cdot\frac{\partial\vec{\hat{r}}}{\partial\xi^j} - \frac{\partial\vec{r}}{\partial\xi^i}\cdot\frac{\partial\vec{r}}{\partial\xi^j}\right)\frac{\partial\vec{r}}{\partial\xi^i}\otimes\frac{\partial\vec{r}}{\partial\xi^j}$$

where $(\xi^1, \xi^2, \xi^3) \equiv (s, \eta, \zeta)$, observing that the local coordinate systems considered are orthogonal.

2.4 Internal load-displacement relations

From the stress-strain relations given by the theory of elasticity combined with the strain tensor expression, it is possible to derive the stress-displacement relations and in turn, after integration over the beam sections, the relations between internal structural loads and displacement variables.

Under the assumptions of rigid beam sections and shear undeformable beam, the following four loads are given in terms of the four variables describing the beam motion (the evaluation of the remaining two loads comes from balancing of the external loads):

$$\begin{aligned}
V_{\hat{x}} &= \vec{v}\cdot\vec{\hat{e}}_1 = \int_A \sigma_{xx}\mathrm{d}\eta\mathrm{d}\zeta \\
&= EA\left(\tilde{u}' + \frac{\tilde{v}'^2}{2} + \frac{\tilde{w}'^2}{2} + k_A^2 k_1^2\right) \\
M_{\hat{\eta}} &= \vec{m}\cdot\vec{\hat{e}}_2 = -\int_A \sigma_{xx}\zeta\mathrm{d}\eta\mathrm{d}\zeta = EI_\eta k_2 \\
M_{\hat{\zeta}} &= \vec{m}\cdot\vec{\hat{e}}_3 = \int_A \sigma_{xx}\eta\mathrm{d}\eta\mathrm{d}\zeta = EI_\zeta k_3 \\
M_{\hat{x}} &= \vec{m}\cdot\vec{\hat{e}}_1 = \int_A (\sigma_{x\eta}\eta - \sigma_{x\zeta}\zeta)\,d\eta d\zeta \\
&= GJ\,k_1 + V_{\hat{x}}k_A^2\,k_1
\end{aligned}$$

where A is the blade cross section area, E is the Young modulus, G is the shear modulus, while

$$I_\eta = \int_A \zeta^2\mathrm{d}\eta\mathrm{d}\zeta \qquad J = \int_A(\eta^2+\zeta^2)\mathrm{d}\eta\mathrm{d}\zeta$$

$$I_\zeta = \int_A \eta^2\mathrm{d}\eta\mathrm{d}\zeta \qquad k_A^2 = \frac{1}{A}\int_A(\eta^2+\zeta^2)\mathrm{d}\eta\mathrm{d}\zeta$$

In addition, k_1, k_2, k_3 are such that for $\vec{k} = k_1\vec{\hat{e}}_1 + k_2\vec{\hat{e}}_2 + k_3\vec{\hat{e}}_3$, the skew-symmetric matrix

$\mathbf{K} = (\mathrm{d}\mathbf{T}/\mathrm{d}s)\mathbf{T}^T$ is the matrix of the components in the local deformed frame of the axial tensor associated to vector $\vec{k}$ (specifically, k_2, k_3 are the bending curvatures of the deformed elastic axis and k_1 is the twist of the blade sections after deformation). Finally, $\tilde{u}', \tilde{v}', \tilde{w}'$ are the components of $\mathrm{d}\tilde{\mathbf{u}}/\mathrm{d}s = \mathrm{d}\mathbf{u}/\mathrm{d}s - \mathbf{K}_0\,\mathbf{u}$, for $\mathbf{u} = \{u\, v\, w\}^T$ and $\mathbf{K}_0 = (\mathrm{d}\mathbf{A}/\mathrm{d}s)\mathbf{A}^T$.

2.5 Inertial loads

Blade rotation and unsteady deformations make inertial loads arise. These contribute to the equilibrium equations as external distributed loads and, combined with the internal structural loads, yield the equations governing the blade structural dynamics.

The expressions of the inertial loads are now obtained in the local, rotating, undeformed blade reference, considering an elastic axis of arbitrary shape. The acceleration of a generic point of a rotating blade is given by

$$\vec{a} = \vec{a}_r + \vec{a}_H + \vec{\Omega} \times \vec{\Omega} \times \vec{\hat{r}} + 2\vec{\Omega} \times \vec{\nu}$$

where $\vec{a}_r$ is the acceleration of the point with respect to a frame rigidly connected to the undeformed blade, depending on the deformation variables (it is derived from the rigid motion of the beam sections expressed in terms of displacements of the elastic axis and rotations about it). In addition, $\vec{a}_H$ is the rotor hub acceleration, $\vec{\nu}$ is the velocity of the examined point with respect to the rotating frame fixed with the undeformed blade, while $\vec{\Omega}$ is the blade angular velocity.

Then, the resulting inertial distributed loads appearing in Eqs. 1 and 2 are derived as

$$\begin{aligned} \vec{p} &= -\int_A \rho\, \vec{a}\, \mathrm{d}\eta\, \mathrm{d}\zeta \\ \vec{q} &= -\int_A \rho\, \vec{r}_s \times \vec{a}\, \mathrm{d}\eta\, \mathrm{d}\zeta \end{aligned}$$

where ρ is the material density and $\vec{r}_s = \vec{\hat{r}} - (\vec{r}_0 + u\vec{e}_1 + v\vec{e}_2 + w\vec{e}_3)$.

2.6 Equations for the deformation variables

Combining the internal and inertial loads given above with the first scalar equation in Eq. 3 and the three scalar equations in Eq. 4 yields a set of four differential equations in the deformation unknowns, u, v, w, ϕ, governing the structural dynamics of the curved blade.

3 Aeroelastic modelling

Aeroelastic applications of the blade structural model outlined in the previous section require the introduction of the aerodynamic loads forcing the blade dynamics. Here, for the sake of simplicity, the aerodynamic loads are derived from a quasi-steady approximation of the Greenberg theory [9] for airfoils. Aerodynamic three-dimensional effects are taken into account by including a wake inflow model (see, for instance, [10] for details). Thus, section force, T, orthogonal to the chord, and section force, S, parallel to the chord are given by

$$T = \frac{\varrho\, C_{l_\alpha} c}{2}\left[-U_P U_T + \frac{c}{2}\,\omega\, U_T - \frac{c}{4}\dot{U}_P + \left(\frac{c}{4}\right)^2 \dot{\omega}\right]$$

$$S = \frac{\varrho\, C_{l_\alpha} c}{2}\left[U_P^2 - \frac{c}{2}\,\omega\, U_P - \frac{C_{d_0}}{C_{l_\alpha}} U_T^2\right]$$

while the section pitching moment with respect to the quarter-chord point reads

$$M_\phi = -\frac{\varrho\, C_{l_\alpha} c^3}{32}\left(\omega\, U_T - \dot{U}_P + \frac{3c}{8}\dot{\omega}\right)$$

In the above equations, U_P and U_T are, respectively, the quarter-chord velocity components normal and parallel to the chord after deformation, ω is the out-of-section component of the angular velocity of the blade section, c denotes the chord length, ϱ is the air density, C_{l_α} is the lift curve slope coefficient, while C_{d_0} is the drag coefficient.

The blade aeroelastic equations are derived by expressing U_P, U_T and ω in terms of u, v, w and ϕ and expressing the aerodynamic forces T and S in terms of components, in the local blade undeformed frame of reference.

4 Numerical implementation

In the numerical solver developed in the presented work a further simplification has been

introduced by using the elimination approach described in Ref. [10] for the axial degree of freedom, u, in order to reduce the final set of governing equations to the three scalar ones in Eq. 4, with unknown variables v, w, ϕ.

In order to explain the solution procedure, first the governing equations are formally recast as

$$\mathbf{m}_l = \mathbf{m}_{in} + \mathbf{m}_{aer} \tag{7}$$

where $\mathbf{m}_{in}$ and $\mathbf{m}_{aer}$ denote the inertial and aerodynamic contribution to the RHS of Eq. 4, respectively.

Next, these equations are integrated spatially through the Galerkin approach, starting from the following spectral description of the deformation variables

$$v(s,t) = \sum_{n=1}^{N} \alpha_n^v(t)\psi_n^v(s)$$
$$w(s,t) = \sum_{n=1}^{N} \alpha_n^w(t)\psi_n^w(s)$$
$$\phi(s,t) = \sum_{n=1}^{N} \alpha_n^\phi(t)\psi_n^\phi(s)$$

and then projecting Eq. 7 onto the second-order derivatives of the shape functions as follows

$$\int_0^R \mathbf{m}_l^T \boldsymbol{\Psi}'' \mathrm{d}s = \int_0^R \left(\mathbf{m}_{in}^T + \mathbf{m}_{aero}^T\right) \boldsymbol{\Psi}'' \mathrm{d}s \tag{8}$$

where

$$\boldsymbol{\Psi} = \begin{Bmatrix} \psi_n^\phi \\ \psi_n^w \\ \psi_n^v \end{Bmatrix}$$

For instance, considering a hingeless rotor ψ_n^v, ψ_n^w and ψ_n^ϕ might conveniently be chosen as the bending and torsion natural modes of vibration of a non-rotating cantilever beam [10]. Note that, the choice of using the second derivatives of the shape functions as test functions is mainly motivated by the equivalence between the equations obtained and those that could be derived from the application of the Rayleigh-Ritz approach.

Equation 8 may be rewritten as

$$\mathbf{f}_{el}(t,\boldsymbol{\alpha}) + \mathbf{f}_{in}(t,\boldsymbol{\alpha}) + \mathbf{f}_{aero}(t,\boldsymbol{\alpha}) = \mathbf{0} \tag{9}$$

where $\mathbf{f}_{in}$, $\mathbf{f}_{aero}$ and $\mathbf{f}_{el}$ are the inertial, aerodynamic, and elastic generalized forces, respectively, and $\boldsymbol{\alpha}$ is the vector collecting the Lagrangian variables of the problem, $\alpha_n^v, \alpha_n^w, \alpha_n^\phi$. It is a set of $3N$ nonlinear, time-dependent equations with unknowns $\boldsymbol{\alpha}$ which can be applied both for aeroelastic response and stability analyses.

Aeroelastic responses are numerically solved through the harmonic balance approach. It is a methodology suitable for the analysis of the asymptotic solution (as time goes to infinity) of differential equations forced by periodic terms, as it generally occurs (see, for instance, helicopter applications). The harmonic balance solution consists of: (i) expressing the contributions in Eq. 9 in terms of their Fourier series; (ii) equating the resulting coefficients; (iii) solving the corresponding algebraic system in terms of the unknown Fourier coefficients of the Lagrangian coordinates of the problem. Specifically, expressing $\boldsymbol{\alpha}$ and $\mathbf{f} = \mathbf{f}_{el} + \mathbf{f}_{in} + \mathbf{f}_{aero}$ in terms of the following Fourier series

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}_0 + \sum_{k=1}^{K} \left[\boldsymbol{\alpha}_k^c \cos(\Omega_k t) + \boldsymbol{\alpha}_k^s \sin(\Omega_k t)\right]$$
$$\mathbf{f}(t) = \mathbf{f}_0 + \sum_{k=1}^{K} \left[\mathbf{f}_k^c \cos(\Omega_k t) + \mathbf{f}_k^s \sin(\Omega_k t)\right]$$

(where $\boldsymbol{\alpha}_k^c, \boldsymbol{\alpha}_k^s, \mathbf{f}_k^c$ and $\mathbf{f}_k^s$ denote cosine and sine components of the k-th harmonic of $\boldsymbol{\alpha}$ and $\mathbf{f}$, whereas $\Omega_k = k\Omega$, with Ω representing the fundamental frequency, *i.e.,* the rotor angular velocity) Eq.9 becomes

$$\hat{\mathbf{f}}\left(\hat{\boldsymbol{\alpha}}\right) = \mathbf{0} \tag{10}$$

where $\hat{\boldsymbol{\alpha}}^T = \{\boldsymbol{\alpha}_0\ \boldsymbol{\alpha}_1^c\ \boldsymbol{\alpha}_1^s\ \boldsymbol{\alpha}_2^c\ \boldsymbol{\alpha}_2^s\ \cdot\ \cdot\ \cdot\}$ and $\hat{\mathbf{f}}^T = \{\mathbf{f}_0\ \mathbf{f}_1^c\ \mathbf{f}_1^s\ \mathbf{f}_2^c\ \mathbf{f}_2^s\ \cdot\ \cdot\ \cdot\}$. Because of the presence of nonlinear terms, Eq. 10 has to be solved using an iterative procedure. To this aim, the Newton-Raphson method is applied. Specifically, the harmonic aeroelastic solution is obtained from convergence of the following iterative procedure (with m indicating the iteration step number)

$$\hat{\boldsymbol{\alpha}}_m = \hat{\mathbf{f}}_{m-1} - \hat{\mathbf{J}}_{m-1}^{-1}\,\hat{\boldsymbol{\alpha}}_{m-1} \tag{11}$$

where $\hat{\mathbf{J}}$ is the Jacobian matrix of the system

numerically evaluated as

$$\hat{J}_{ij} = \frac{\partial \hat{f}_i}{\partial \hat{\alpha}_j}$$

Aeroelastic stability analysis is performed though the eigenanalysis of the following set of equations linearized about a reference (equilibrium) configuration, $\boldsymbol{\alpha}_0$, obtained as solution of Eq. 11

$$\mathbf{M}(\boldsymbol{\alpha}_0, t)\ddot{\boldsymbol{\alpha}} + \mathbf{C}(\boldsymbol{\alpha}_0, t)\dot{\boldsymbol{\alpha}} + \mathbf{K}(\boldsymbol{\alpha}_0, t)\boldsymbol{\alpha} = \mathbf{0}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are, respectively, mass, damping and stiffness matrices, numerically obtained as

$$M_{ij} = \left.\frac{\partial f_i}{\partial \ddot{\alpha}_j}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \qquad C_{ij} = \left.\frac{\partial f_i}{\partial \dot{\alpha}_j}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}$$

$$K_{ij} = \left.\frac{\partial f_i}{\partial \alpha_j}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0}$$

5 Numerical results

The numerical investigation concerns the validation of the proposed blade structural model, in terms of both free-vibration analysis and aeroelastic response of a helicopter blade with curved elastic axis. To this purpose, results from the present solver are compared with those obtained by a finite element model developed in the past by the authors and successfully applied to the aeroelastic analysis of helicopter blades with tip sweep and anhedral angles [7].

Here, the blade considered in the analysis is untwisted, without aerodynamic, mass and torque offsets, with radius $R = 4.741m$, chord $c = 0.395m$ and a parabolic elastic axis of equation $y = -0.1x^2/R$ (with x denoting the coordinate along $\vec{i}_1$ direction, and y that along $\vec{i}_2$ direction). The blade flap and lead-lag bending stiffnesses are, respectively, $EI_\eta = 51600Nm^2$, $EI_\zeta = 143441Nm^2$, while the torsion rigidity is $GJ = 13913Nm^2$. The section mass moments of inertia are $J_\eta = 0.0157kg\,m$, $J_\zeta = 0.063kg\,m$ and the blade mass per unit length is $m = 6.46kg/m$.

The validation analysis is performed in two steps: the first one (in vacuo analysis) focused on the validation of the blade structural model with and without the inclusion of inertial effects due to rotary motion, and the second one (aeroelastic analysis) focused on the inclusion of aerodynamic effects. All the computations are carried out considering nine shape functions per each dof in the Galerkin approach (present formulation), and nine straight blade elements in the FEM approach (such discretization ensures practically converged results with both solvers).

5.1 In vacuo analyses

First, the free-vibration, in-vacuo analysis of the nonrotating curved blade is considered. Figure 2 shows the comparison between the eigenfrequencies of the first five modes of vibration obtained by using the present formulation and those from the FEM solver [7]. The quality of the comparison is quite good in that only very small discrepancies (lower than 2%) appear in the evaluation of the eigenfrequencies regarding the second-flap and the first-torsion modes of vibration.

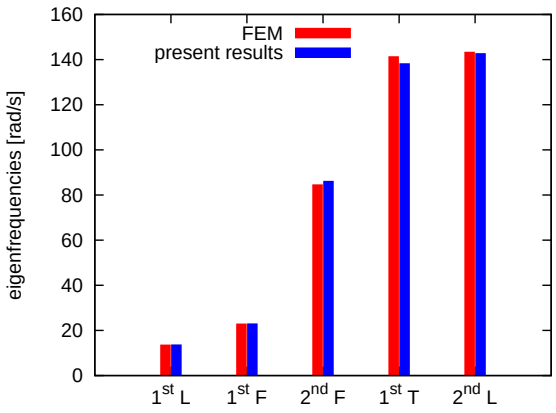


Figure 2: Nonrotating curved blade: eigenfrequencies.

Then, the free-vibration analysis of rotating blades is considered. For this analysis, the blade angular velocity is set to $\Omega = 40.123rad/s$. In the presence of rotary inertia terms, the objective is twofold: first, assessment of the effects of elastic axis curvature on the dynamic behavior of the blade and second, validation of the predicted eigenfrequencies against those evaluated from the FEM solver [7]. Figure 3 shows the steady-state lead-lag displacement due to rotation evaluated by the present formulation and

through the FEM solver. They are in satisfactory agreement, and define the reference configurations about which the free-vibrations are computed.

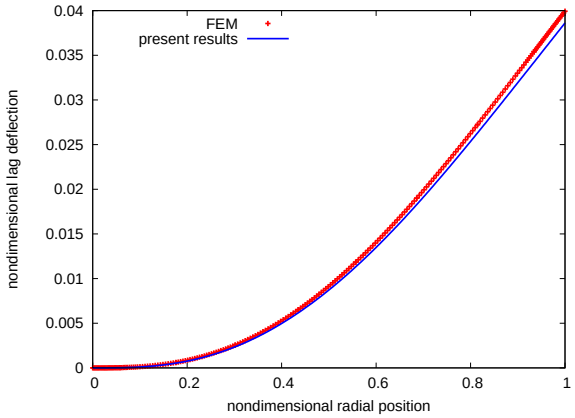


Figure 3: In vacuo rotating curved blade: equilibrium deflections.

Figures 4 and 5 present the results of the perturbation eigenanalysis in terms of eigenfrequencies, respectively concerning the blade curvature effects and the validation with respect to the FEM solver. Note that the straight blade results in Fig. 4 have been obtained considering blade having the same radius, mechanical and structural properties of the curved one.

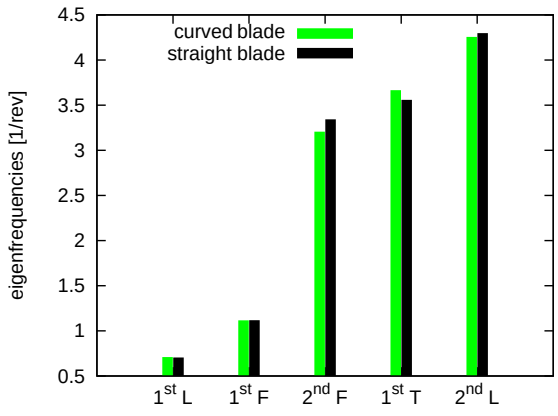


Figure 4: In vacuo rotating blades: eigenfrequencies (present approach).

The analysis of Figs. 4 and 5 shows that some small discrepancies (lower than 4%) appear in the evaluation of the second-flap and first-torsion eigenfrequencies of the curved blade.

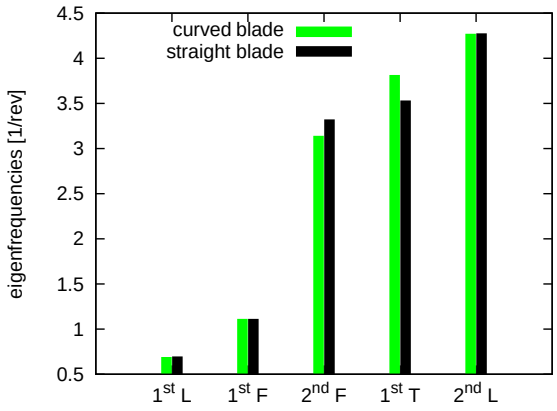


Figure 5: In vacuo rotating blades: eigenfrequencies (FEM approach).

However, the overall quality of the results might be considered satisfactory, in that the trend of eigenfrequencies variation due to elastic axis curvature effects predicted by the two approaches is similar.

5.2 Aeroelastic analyses

Finally, the aeroelastic analysis of the curved elastic axis blade in hovering condition is presented. For this analysis, the blade has been assumed to have angular velocity $\Omega = 40.123 rad/s$, with collective angle $\theta_0 = 9 deg$.

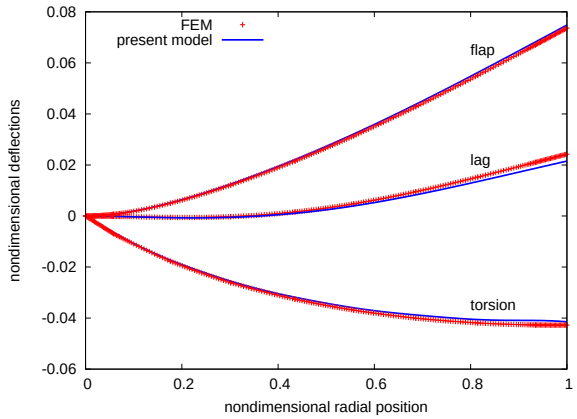


Figure 6: Rotating curved blade: aeroelastic equilibrium deflections.

Figure 6 depicts the steady-state deflections (flap-lag displacements and torsion rotation)

along the blade radius given by the present solver and those from the FEM one [7], while Fig. 7 presents the results of the perturbation

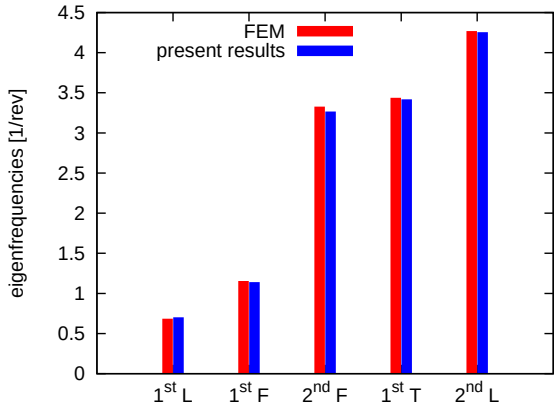


Figure 7: Rotating curved blade: aeroelastic eigenfrequencies.

aeroelastic eigenanalyses in terms of eigenfrequencies, carried out about the trim configurations of Fig. 6. Figures 6 and 7 show a good agreement between present results and those given by the FEM solver, thus validating the proposed blade model also for aeroelastic applications.

6 Conclusions

A theoretical-numerical structural dynamics and aeroelastic model for curved rotating blades has been presented. The numerical implementation has been based on a non-conventional Galerkin approach. In vacuo analysis of a parabolic-shaped blade has demonstrated that the proposed model is able to predict steady-state response and natural frequencies of vibration in very good agreement with those given by a FEM solver validated in the past against literature numerical and experimental data. Similarly, a good correlation between present results and FEM predictions has been shown in the case of aeroelastic application concerning steady-state response and natural frequencies of vibration of the rotor blade in hovering condition. For the configurations examined the proposed solver can be considered validated. However, the presented numerical investigation has to be considered as a first validation, in that more complex blade configurations and flight conditions need to be investigated.

References

[1] Munsky, B., Gandhi, F., and Tauszig, L., "An Analysis of Helicopter Blade-Vortex Interaction Noise with Flight Path or Attitude Modification," *P*roceedings of the 58th Annual Forum of the American Helicopter Society, Montréal, Canada, 2002.

[2] Patt, D., Liu, L., and Friedmann, P.P., "Rotorcraft Vibration Reduction and Noise Predictions Using a Unified Aeroelastic Response Simulation," *J*ournal of the American Helicopter Society, Vol. 50, No. 1, 2005, pp. 95-106.

[3] Beaumier, P., and Delrieux, Y., "Description and Validation of the ONERA Computational Methods for the Prediction of Blade-Vortex Interaction Noise," *A*erospace Science and Technology, Vol. 9, 2005, pp. 31-43.

[4] Tung, C., Yu, Y.H., and Low, S.L., "Aerodynamic Aspects of Blade-Vortex Interaction (BVI)," AIAA Paper 96-2010, *P*roceedings of the 27th AIAA Fluid Dynamics Conference, New Orleans, LA, 1996.

[5] Bauchau O.A., Hong C. H., "Large Displacement Analysis of Naturally Curved and Twisted Composite Beams," *AIAA Journal*, Vol. 25, No. 11, 1987, pp. 1469-1475.

[6] Yuan K.A., Friedmann P.P., "Aeroelasticity and Structural Optimization of Composite Helicopter Rotor Blades With Swept Tips", NASA CR-4665, 1995.

[7] Piccione E., Bernardini G. and Gennaretti M., "A Finite Element Model for Structural and Aeroelastic Analysis of Rotary Wings with Advanced Blade Geometry", *COMSOL European Conference* , Paris, 2010.

[8] Hodges D.H., Dowell E.H., "Nonlinear Equations of Motion for the Elastic Bending and Torsion of Twisted Nonuniform Rotor Blades", NASA TN D-7818, 1974.

[9] Greenberg J.M., "Airfoil in Sinusoidal Motion in Pulsating Stream", NACA TN-1326, 1947.

[10] Hodges D.H., Ormiston R.A. , "Stability of Elastic Bending and Torsion of Uniform Cantilever Rotor Blades in Hover with Variable Structural Couplin", NASA TN D-8192, 1976.

Influence of High Performance Computing on Aeroelasticity and Structural Dynamics

F. Sanchez-Zabala, M. Oliver and H. Climent

*AIRBUS-MILITARY, John Lennon s/n, 28906 Getafe (Madrid) – Spain.

Mercedes.Oliver@military.airbus.com

Aeroelasticity and Structural Dynamics is one of the most multidisciplinary realms in aeronautical engineering:

- In aeroelasticity, the engineer should deal with the coupling of structure, mass, unsteady aerodynamics, flight control system laws, etc.
- In ground dynamic loads, the engineer should deal with coupling of a complex linear model (complete aircraft) with a highly non-linear landing gear model.
- In flight dynamic loads the response of the structure to complex excitations like gust and continuous turbulence should be assessed.
- Environmental vibrations are one of the key elements in equipment qualification.
- Impacts simulations (ditching, crashworthiness, bird strike, hail impacts...) requires intensive knowledge of materials characterisation, rupture criteria, collapse sequences and explicit FE modelling...

All these aspects are –to mention few– an indication of the multidisciplinary environment in which the aeroelastician should dive every day. Figure 1 shows a sketch of the aeroelasticity and structural dynamics realm.

The activity of aeroelasticity and structural dynamics (A&SD) is extremely intensive in the use of computer resources. The evolution of available computer mainframes has been one of the main drivers in the development of the technology itself.

This paper will revise the effect that the evolution of High Performance Computing (HPC) has had in the methods, techniques, processes and tools of A&SD. First numerical simulation techniques will be addressed with especial focus on Finite Element Method and its transcendence for A&SD. Next, the solvers will be addressed. How these tools have evolved and how the results that these tools provide have been compared with complete aircraft ground and flight tests.

The paper will continue with the effect of the evolution of HPC on the dynamic tests:

- Effect of HPC evolution on Ground Vibration Tests (GVT). From the early days in which hardly the signal of few accelerometers could be treated to nowadays where a complete aircraft GVT may handle more than 700 accelerometers simultaneously. The paper will show examples of Airbus Military (former CASA) GVT's: the trainer C101, the medium military transport CN-235, C-295, the Airbus Tanker A330-MRTT or the very recent heavy military transport A400M.
- Another section will cover the effect of HPC on Flight Flutter Tests (FVT): how the evolution in computer technology has affected all the links of the flutter test chain: from the excitation, to the response. A dedicated section to telemetry will complement this description. Again the examples will be concentrated in CASA, Airbus Military aircraft with special attention to the EF-2000 fighter aircraft FVT.
- Impact tests will be finally addressed: the evolution of HPC has allowed jumping from only pass-fail test to sophisticatedly instrumented tests. The example shown will be the ditching tests of the CN-235 deepwater program of the US Coast Guard.

Figure 2 shows the evolution with time of HPC showing how the computer capacity increases exponentially with time, allowing to address more and more sophisticated tasks and with higher fidelity to reality.

The paper will end with conclusive remarks devoted to the increasing possibility of virtual testing and qualification by analysis only and how this possibility is being addressed between industry and the Airworthiness Authorities.

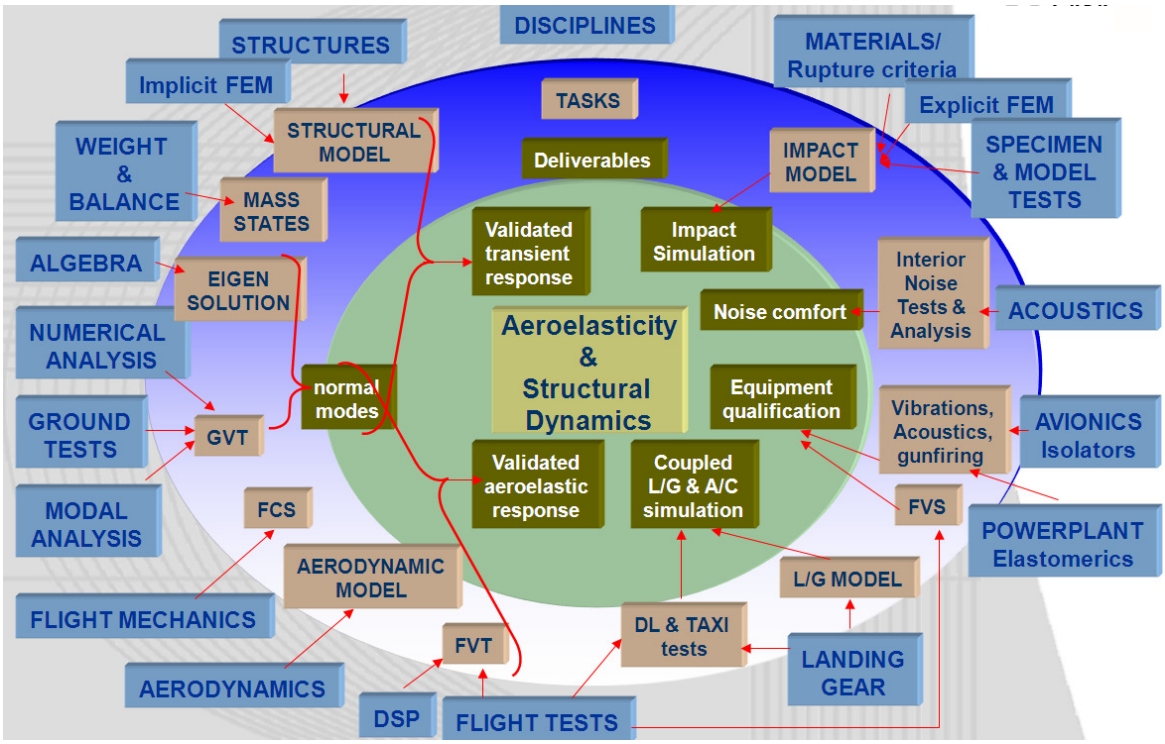


Figure 1. The Multidisciplinary realm of Aeroelasticity and Structural Dynamics

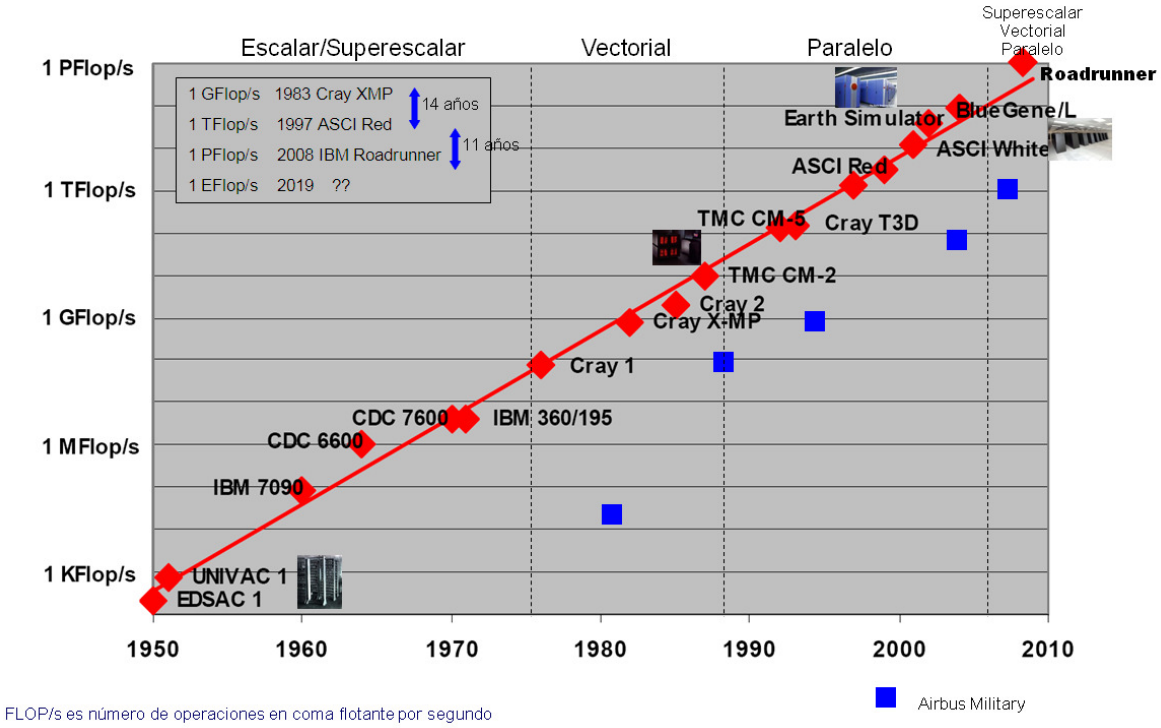


Figure 2. The Evolution of High Performance Computing (HPC)

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Experimental and numerical investigation of freeplay on a T-tail

Sebastiano Fichera*, **Giuseppe Quaranta**† and **Sergio Ricci**†
Dipartimento di Ingegneria Aerospaziale – Politecnico di Milano
Via La Masa 34, 20156, Milano, Italy

e–mail: *fichera@aero.polimi.it, †giuseppe.quaranta@polimi.it and †sergio.ricci@polimi.it

Keywords:*Nonlinear, Limit Cycle Oscillation, Freeplay, Experimental Aeroelasticity*

Abstract

Freeplay is one of the most important nonlinearities that affect the control surfaces of the aircraft; it can induce flutter phenomena and limit the performances of the same airplane. To investigate the effect of control surface freeplay, an aeroelastic wind tunnel model of a T-tail has been developed. A variable amplitude freeplay has been introduced in the control chain by a specifically designed linkage. The numerical model was designed according to the modern aeroelastic approach, describing the dynamics of the tail by a state space system. The nonlinearity has been introduced has a lumped element in the actuator feedback loop. The paper presents the experimental setup together with the correlation obtained with the numerical model.

1 Introduction

The research on nonlinear aeroelasticity and, in particular, on control surface freeplay, is motivated by the significant number of cases known in the literature of aircrafts that have experienced Limit Cycle Oscillations (LCO) caused by it (see [1]). In fact, freeplay in the control chains may arise as consequence of many factors, including wear of the parts during the aircrafts life. An interesting summary of the occurrence of aeroelastic control surface LCOs caused by freeplay on Airbus aircraft, may be found in Ref. [2]. A complete survey on nonlinear aeroelasticity can be found primarily in Ref. [3] and in the more recent work Ref. [4].

In order to perform numerical and experimental investigations a wind tunnel model of a T-Tail equipped with the rudder and the control system has been designed and manufactured. The following goals want to be achieved at the same time:

- perform experiments with or without the freeplay;
- test different freeplay amplitudes;
- investigate the effect of non-symmetric nonlinearity;
- allow to introduce an active control for LCO suppression.

The study of an active control for LCO suppression will be, however, the final target of this project and all the model has been designed around this aim. The rudder has been chosen since it can be expected to be an aerodynamic surface that operates for a large portion of its life with a null loading, a condition where the effects of freeplay may become more evident. The basic elements of the rudder control chain are shown in the block scheme of Figure 1. Here it is possible to see that two different angular

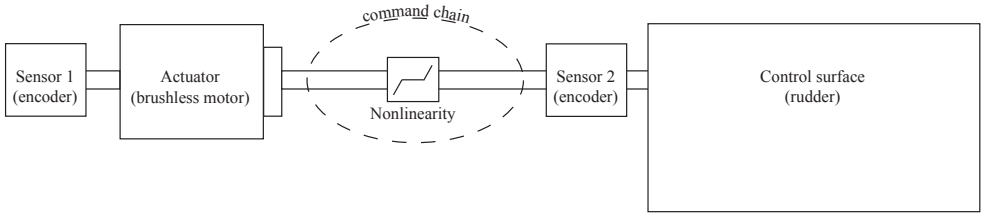
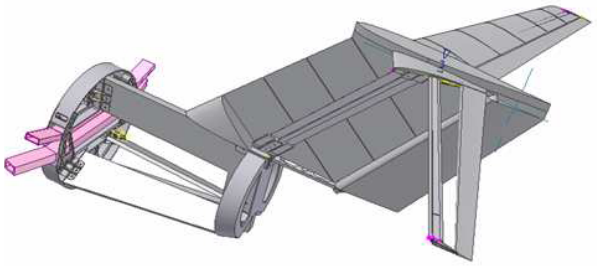


Figure 1: Control chain.

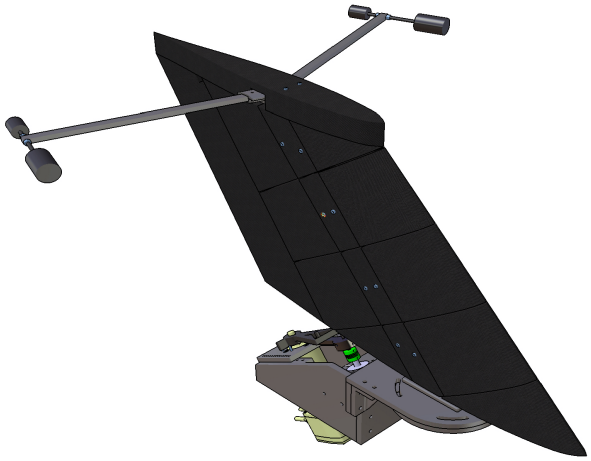
position sensors are included, one upstream and one downstream the freeplay, both to investigate the actual system response and to compute the error signals to be used by the active control for LCO suppression. The T-tail unit considered in this study is the one of the X-DIA, an aeroelastic model representative of a non-conventional three surfaces regional jet, intensively investigated in the last few years at the Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano DIA-POLIMI (see [5, 6, 7]).

2 T-tail Experimental model

The T-tail experimental model has been scaled in order to keep the same Froude number of the target aircraft. The elements that compose the wind tunnel model are shown in Figure 3: the core is composed by dynamically scaled spars, which are inserted in a series of sectors made by styrofoam covered by a carbon fiber skins. The fin structure is made by an aluminum alloy beam that is covered by five sectors that are used to get the right airfoil shape. Table 1 summarize the geometric characteristics of the T-tail. The tail structure was planned to be tested in the PoliMI Small Wind Tunnel (PSWT). The PSWT is a closed circuit low-speed wind tunnel, with a rectangular test section of $1.5 \times 1 \times 3$ m and no heat exchangers installed. The test chamber is made of square steel tubes and replaceable walls, which are sometimes substituted with glass windows for flow visualizations, or wooden panels to ease the model installation. Since the size of the test chamber is small and the contribution of the horizontal tail to the freeplay LCO is expected to be low, the horizontal tail plane considered in the preliminary design Figure 2(a) has been reduced in order to keep the same dynamic contribution but neglecting the aerodynamic one. Consequently the aluminum alloy spars were not covered by the shape sectors and ballast were added to obtain the correct inertial properties (see Figure 2(b)). This is the usual way adopted when there is not the possibility to fit the entire model, see [8, 9]. The span of the horizontal tail was set to 2/3 of the wind tunnel test section width.

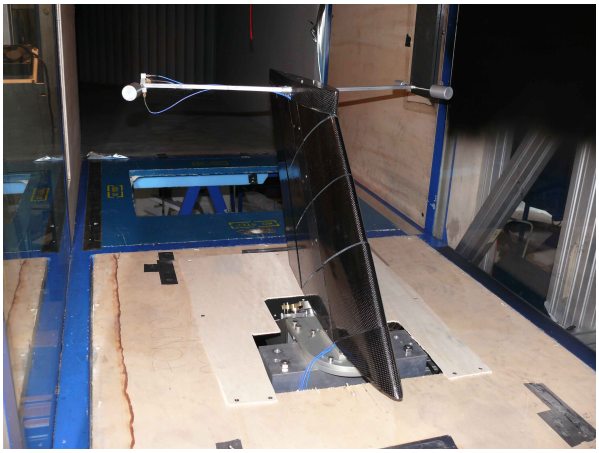


(a) Initial T-tail design

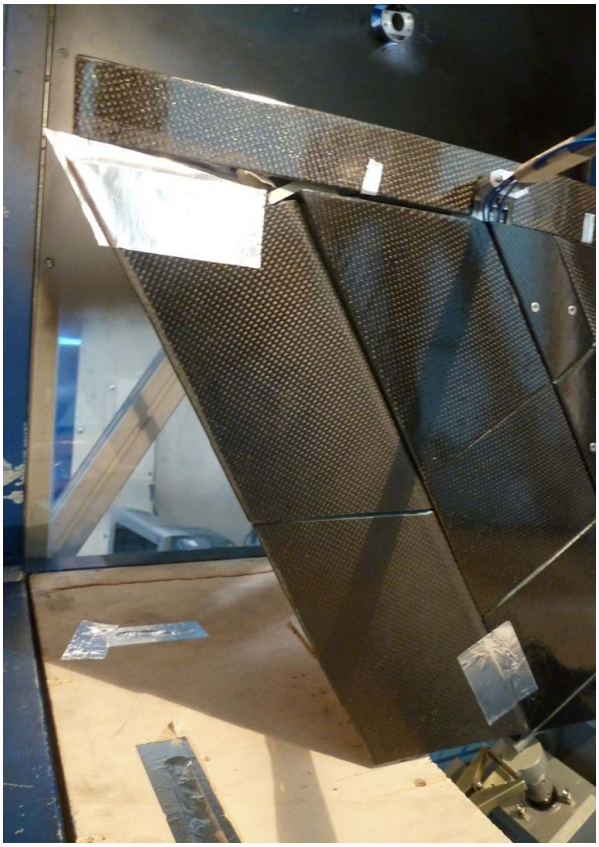


(b) Final model rendering

Figure 2: T-tail CAD model.



(a) Model in the wind tunnel



(b) Half rudder

Figure 3: T-tail aeroelastically scaled wind tunnel model.

In order to have a variable amplitude freeplay, the mechanism shown in Figure 4 has been inserted in the control chain between the actuator and the rudder. This is composed by a rigid linkage connected with the rudder and that ends with a pin that is slipped into a fork connected with the gear of the electric motor that is used to actuate the movable surface. The connecting fork allows to modify the behavior from no freeplay up to $\pm 10^{\circ}$ by changing the position of the electric motor with respect to the rudder hinge axis. The angular position of the motor-fork assembly is controlled by a PID.

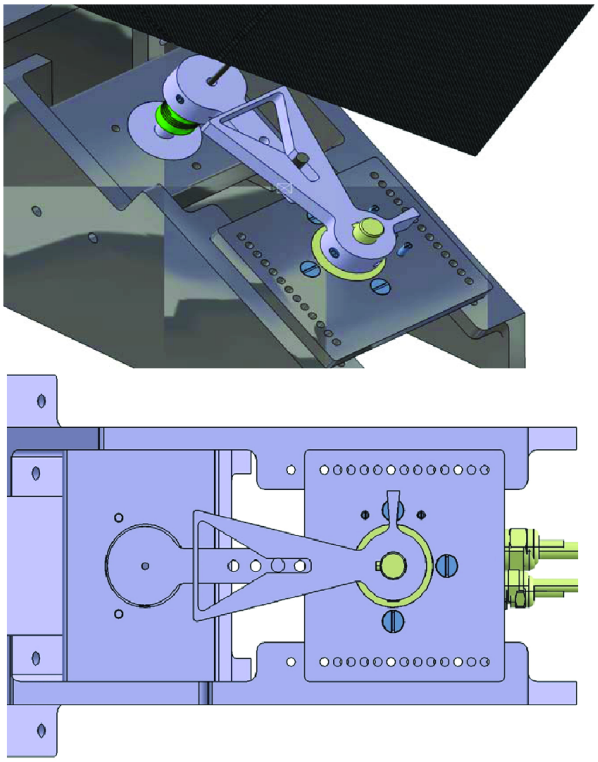


Figure 4: Kinematic chain designed to build a variable amplitude freeplay system.

The rudder surface has been cutted in the middle along the span, the portion near the root has been fixed to the main fin and the tip one has been left moveable, as shown in Figure 3(b). This design has been chosen because increases the unsteady participation of the rudder to the whole model.

2.1 Measurement, Data Acquisition and Control

The rotation of the rudder is measured by an incremental encoder which connected to the rudder hinge axis at the fin root. A second encoder is used to measure the rotation of the motor-fork assembly. Three uniaxial accelerometers

Table 1: T-Tail Specifications.

Characteristic	Notation	Units	Fin
Airfoil			NACA 0015
Span	b	m	0.405
Area	S	m^2	0.212
Root Chord	C_r	m	0.593
Tip Chord	C_e	m	0.452
Mean Aerodynamic Chord	cma	m	0.526
Aspect Ratio	λ		0.774
Sweep Angle	Λ	deg	36.75
Taper	r		0.762
Rudder Root Chord	C_{r_t}	m	0.165
Rudder Tip Chord	C_{e_t}	m	0.125

have been placed at the center, at the leading and trailing edge of the fin tip. A triaxial accelerometer has been placed at the tip of the reduced horizontal tail plane. All data acquisition and system control has been implemented by using the RealTime Application Interface, an open-source real time operating system developed at Politecnico di Milano.

3 Numerical model

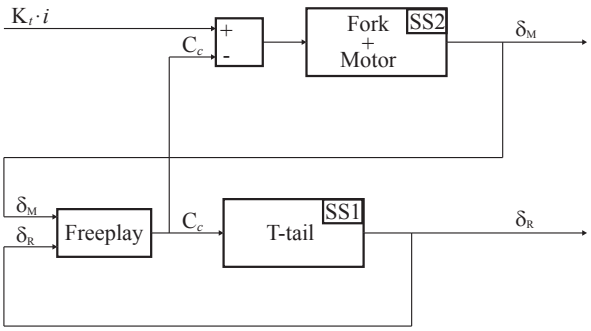


Figure 5: Nonlinear simulator: Sub-systems block scheme.

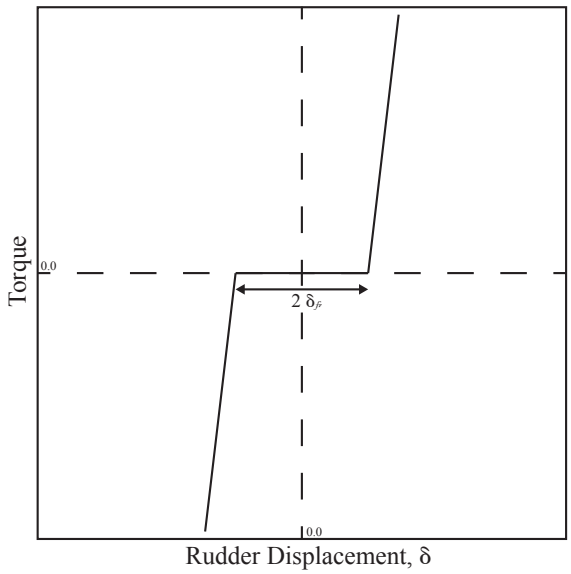


Figure 6: Nonlinear simulator: Freeplay fork-contact torque curve.

The nonlinear system has been modeled as two connected dynamics linear sub-systems, i.e. the T-tail and the fork-motor assembly. The T-tail sub-system state space matrices are built using the structural Finite Element model and the aerodynamic Doublet Lattice Model, both developed in MSC.NASTRAN. The frequency domain data are transformed into a finite state space realization by using the algorithm presented in [7].

The aeroelastic system is described in Eq. 1:

$$\mathbf{M}_s\ddot{\mathbf{q}} + \mathbf{C}_s\dot{\mathbf{q}} + \mathbf{K}_s\mathbf{q} = q\mathbf{f}_a + \mathbf{f}_m \qquad (1)$$

where $\mathbf{M}_s$, $\mathbf{C}_s$ and $\mathbf{K}_s$ are respectively the mass, damping and stiffness structural matrices and $\mathbf{f}_a$ is the vector of Generalized unsteady Aerodynamic Forces (GAFs), and $\mathbf{f}_m$ represents the effect of the second subsystem on the rudder structure

$$\mathbf{f}_m = \begin{Bmatrix} 0 \\ C_c \end{Bmatrix}. \qquad (2)$$

The vector of state variables of the structural system is composed by

$$\mathbf{q} = \begin{Bmatrix} \mathbf{q}_m \\ \delta_R \end{Bmatrix}, \qquad (3)$$

where $\mathbf{q}_m$ are the modal degrees of freedom and δ_{SS1} the rudder rotation angle. Usually the GAFs are computed in the reduced frequency k domanin by using the classical DLM, i.e.

$$\mathbf{f}_a = \mathbf{H}_{am}(k, M_\infty)\mathbf{q}. \qquad (4)$$

The model of Eq. 4 can be fitted into a state space system by using the classical Roger's expansion [10]

$$\mathbf{H}_{am}(p) \simeq \mathbf{D}_0 + \mathbf{D}_1 p + \mathbf{D}_2 p^2 + \sum_{i=1}^{N} \frac{p}{p-\beta_i}\mathbf{E}_i, \qquad (5)$$

where $p = \frac{sc}{2V}$ is the non dimensional Lapalce variable. This expression can be easily transformed into a classical state space time domain system

$$\begin{aligned} (\tfrac{c}{2V})\dot{\mathbf{x}}_a &= \mathbf{A}_a\mathbf{x}_a + \mathbf{B}_a\mathbf{q} \\ \mathbf{f}_a &= \mathbf{C}_a\mathbf{x}_a + \mathbf{D}_0\mathbf{q} + (\tfrac{c}{2V})\mathbf{D}_1\dot{\mathbf{q}} + (\tfrac{c}{2V})^2\mathbf{D}_2\ddot{\mathbf{q}}. \end{aligned} \qquad (6)$$

The electro-mechanical part of the second sub-system in represented by Eq. 7,

$$\begin{aligned} L\frac{di}{dt} &= -Ri - K_e\omega + v \\ (J_m + J_f)\frac{d\omega}{dt} + B\omega &= K_t i - C_c \end{aligned} \qquad (7)$$

where i is the current, v is the voltage and ω is the angular velocity of the motor axis; L and R are the motor's inductance and resistance and K_e is the rotor speed constant; J_m and J_f are respectively the inertia moments of the motor and the connected fork; B represent the effect of friction, and K_t is the motor torque constant. The dynamics of motor electronics has been neglected because is too fast compared to the aeroelastic one. A PID system is used to control the motor current in order to obtain the required rudder position.

When the pin is on the boundary of the fork, the two sub-systems are connected and the stiffness of the movable surface is controlled by the electric motor. Instead, when the pin is not on limit sides of the fork, and travels freely into the gap, the two sub-systems act as independent dynamical models.

The modeling of this switch system is managed by using a penalty function approach [11], that computes a force exchanged between the two sub-systems as described in Eq. 8:

$$\begin{cases} C_c = 0 & \text{if} \quad |\delta_M - \delta_R| \le \delta_{fr} \\ C_c = K_\delta((\delta_M - \delta_{fr}) - \delta_R) + K_{\dot\delta}(\dot\delta_M - \dot\delta_R) & \text{if} \quad |\delta_M - \delta_R| > \delta_{fr}. \end{cases} \qquad (8)$$

This approach causes a switch that represent a discrete change in the connection stiffness, as shown schematically in Figure 6.

The gains K_δ and $K_{\dot\delta}$ have been chosen, in agreement with [12], to minimize both the penetration as well the rebound between the two subsystems.[1]

The simulations have been performed using Simulink with Dormand-Prince method. It is crucial for the system to be integrated by an algorithm that is able to compute exactly the switching point. The importance of locating this switching point was addressed in Ref. [13].

[1] See Par. 4.2 for a sensibility analysis about these gains.

4 Preliminary Test

4.1 Finit Element Model update

A preliminary test on the flutter model has been used to identify the natural frequencies. A network of accelerometers has been placed on the tail and an experimental modal analysis has been perofrmed by using a classical impact hammer technique. The difference between numerical and experimental first and second bending mode has been minimized by tuning the stiffness properties of the main spar. Figure 7 show the cross MAC matrix between numerical and experimental model, while Table 2 summarizes the final frequencies and mode shapes.

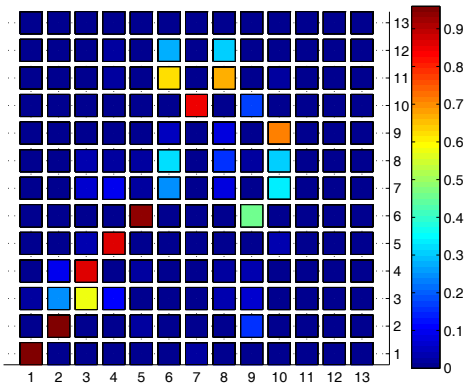


Figure 7: Cross MAC numerical-experimental model.

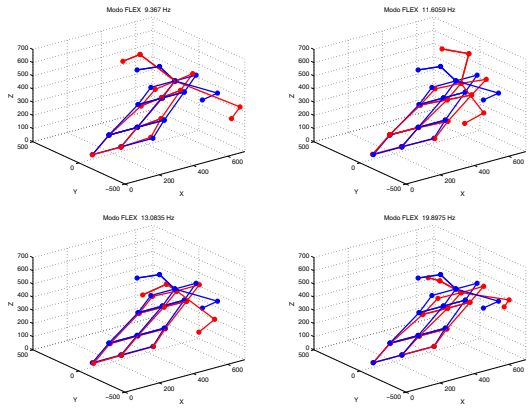
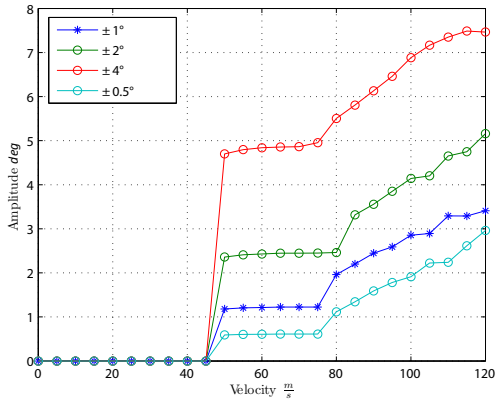


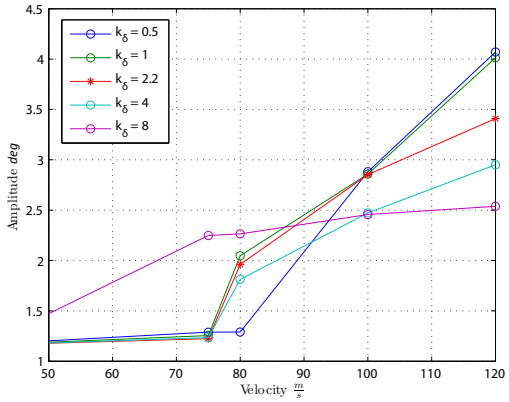
Figure 8: Mode shapes.

4.2 Sensitivity of the Numeric model

To verify the strength of the numerical model, a sensitivity analysis has been made around the K_δ parameter, and the freeplay gap, as shown in Figure 9.



(a) freeplay gap sensitivity



(b) K_δ sensitivity

Figure 9: Sensitivity analysis.

5 Results

The studied configuration has a freeplay amplitude of $\pm 1°$. The initial angle of attack of the model is zero. The T-tail is placed in the freestream such that the control surface is resting in the middle of the freeplay region. Transients are allowed to decay, and the rudder rotation angle and the accelerometers signal are recorded. Then, the freestream speed is slowly increased to a new value and the process is repeated.

Table 2: T-Tail Frequencies and Mode shapes.

Numerical model		Experimental model	
Mode Shape	Frequency [Hz]	Mode Shape	Frequency [Hz]
1 - 1st Fin Torsion Mode	9.36	1 - 1st Fin Torsion Mode	9.37
2 - 1st Fin Bending Mode	11.63	2 - 1st Fin Bending Mode	11.61
-	-	3 - 1st Ballast Dummy Mode	12.42
3 - 1st Tailplane Bending Mode	13.07	4 - 1st Tailplane Bending Mode	13.08
4 - 1st Tailplane In-Plane Mode	18.18	5 - 1st Tailplane In-Plane Mode	18.17
5 - 2nd Fin Bending Mode	23.22	6 - 2nd Fin Bending Mode	19.90
-	-	7 - 2nd Ballast Dummy Mode	28.64
-	-	8 - 3rd Ballast Dummy Mode	30.94
-	-	9 - 4rd Ballast Dummy Mode	36.39
6 - 1st Fin In-Plane Mode	36.80	-	-
7 - 1st Tailplane Torsion Mode	38.45	10 - 1st Tailplane Torsion Mode	38.42

5.1 Numerical simulation

Figure 10 shows the numeric model of the T-tail with the half rudder configuration. The numerical model excitation is generated by an impulsive rotation of the rudder; the freestream speed explored ranges form 10 m/s to 120 m/s, with an increasing step of 5 m/s that is reduced to 1 m/s near the switching point. The zero position of the fork is kept by the PID controller.

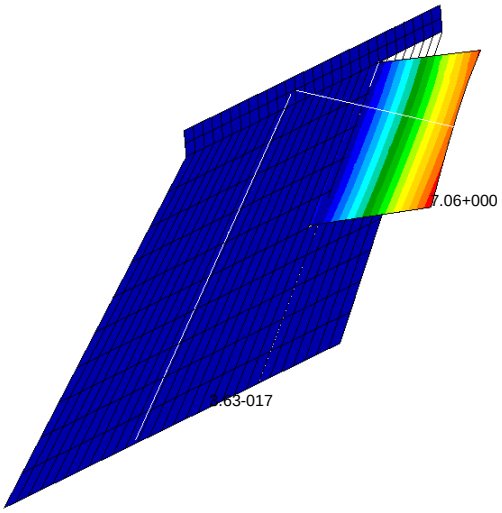


Figure 10: Numeric model of T-Tail with half rudder configuration.

The numerical V-g flutter diagrams, computed without the nonlinearity, are shown in Figure 11. The first (11.63 Hz) and the second (23.22 Hz) fin bending modes cross the zero damping line respectively at 47 m/s and 78 m/s.

The numeric model results, obtained by a direct time marching integration of the nonlinear model, show two LCO regions (see Figure 13). The first region is the consequence of the flutter of the first bending mode, while the second is a combination of the first and the second mode flutters.

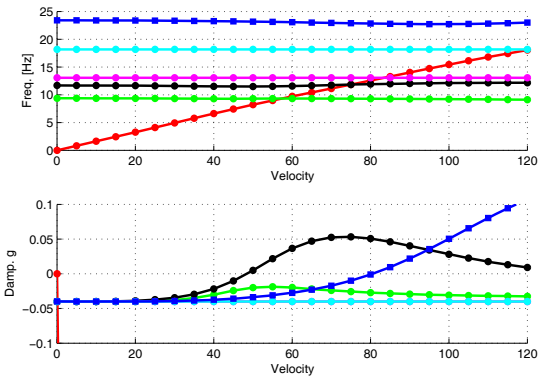
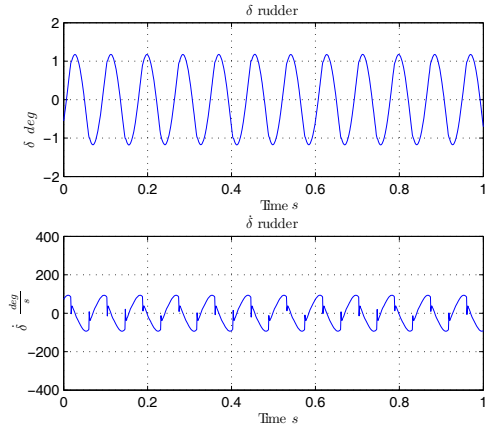
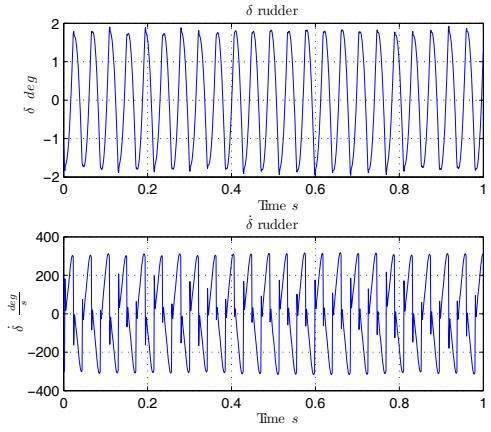


Figure 11: V-g V-f diagram for T-tail.

Figure 12 shows the time histories for the low and hight frequency LCOs. Note that all the time series printed in this section are for an arbitrary $1s$ time window considered when all transient are damped out and only the LCO regimes are evident. The main frequency that dominate in each region is shown in Figure 15 and, as expected, is related to the associated structural mode. The flap motion is similar in the two LCOs, there is only an increase in amplitude, as shown in the phase diagrams of Figure 14.



(a) low frequency LCO 50 m/s



(b) high frequency LCO 80 m/s

Figure 12: Time history of rudder rotation angle and velocity.

The frequency analysis, see Figure 15, shows the typical path of an odd nonlinearity with all the odd multiple of the basic frequency that are clearly visible in the response. The FFT of the high frequency LCO reveals that both bending modes are participating into the response, even thought the dominant mode is the second bending mode.

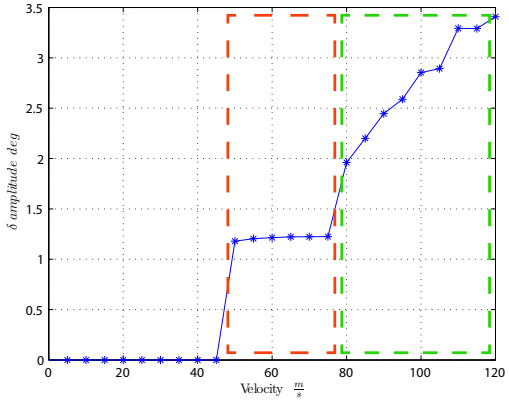
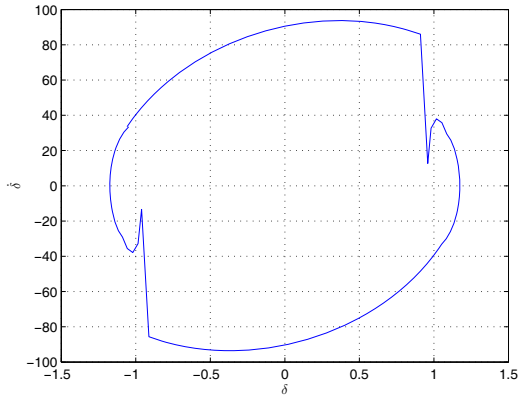
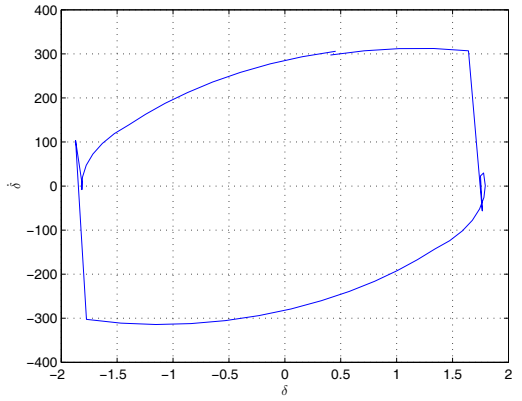


Figure 13: LCO Velocity vs δ (deg).

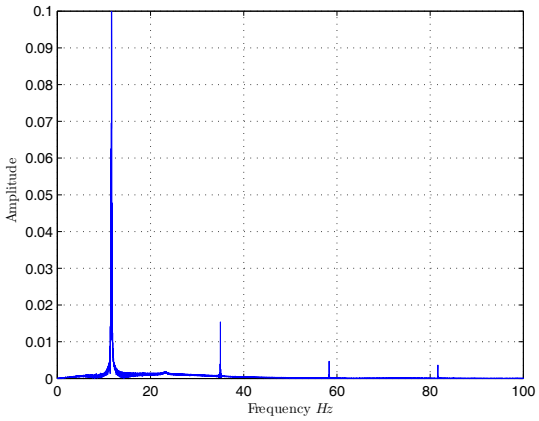


(a) low frequency LCO 50 m/s

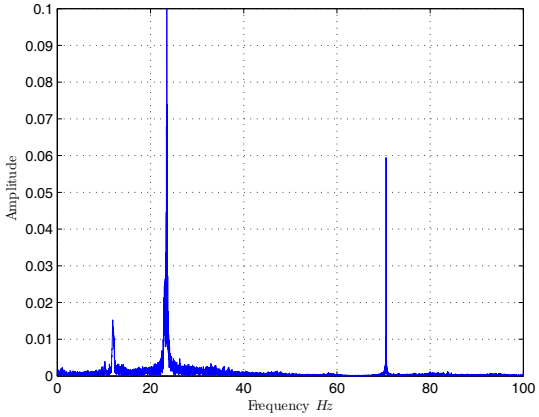


(b) high frequency LCO 80 m/s

Figure 14: LCO phase portraits, $\pm 1°$.

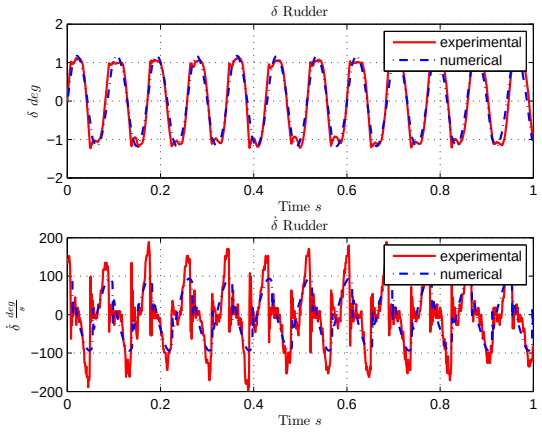


(a) low frequency LCO 50 m/s

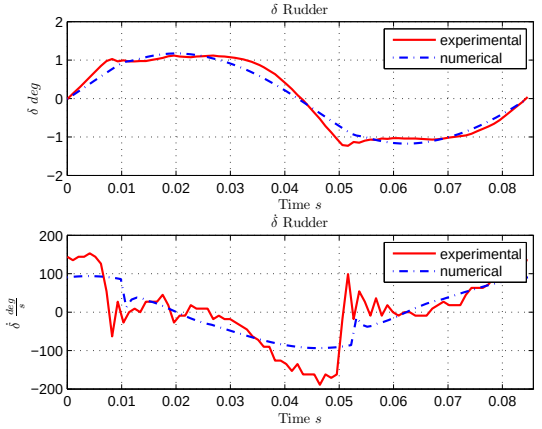


(b) high frequency LCO 80 m/s

Figure 15: LCO signals FFT, $\pm 1°$.



(a) low frequency LCO 50 m/s, one second



(b) low frequency LCO 50 m/s, one period

Figure 16: Rudder LCO, $\pm 1°$ time series comparison.

5.2 Experimental Results

A similar approach has been used experimentally, however, in this case, the maximum freestream speed has been 50 m/s in relation to the maximum PSWT speed. Two runs have been performed, increasing and decreasing the wind tunnel speed by a speed step of 1 m/s, respectively. The experimental model does not require any external excitation because of the intrinsic small imperfections are able to switch-on the LCO. After a preliminary validation of the motor PID controller, the fork has been fixed at the zero position. The range of velocity explored shows only the low frequency LCO; the compared rudder angular position between numerical and experimental model are shown in Figure 16.

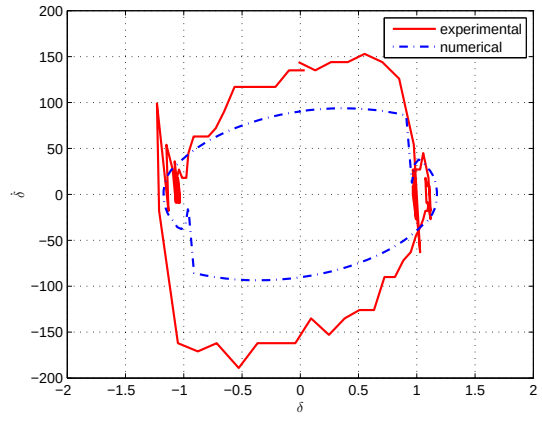


Figure 17: LCO phase portraits, $\pm 1°$, 50 m/s comparison.

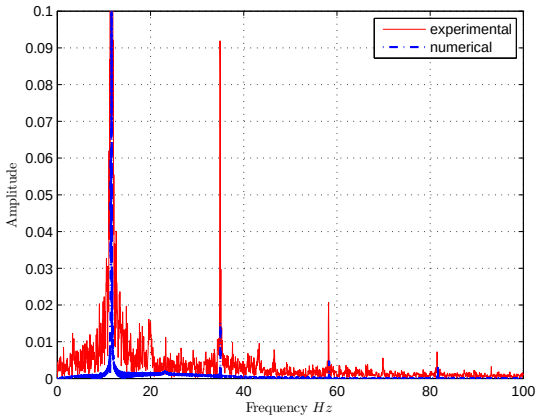
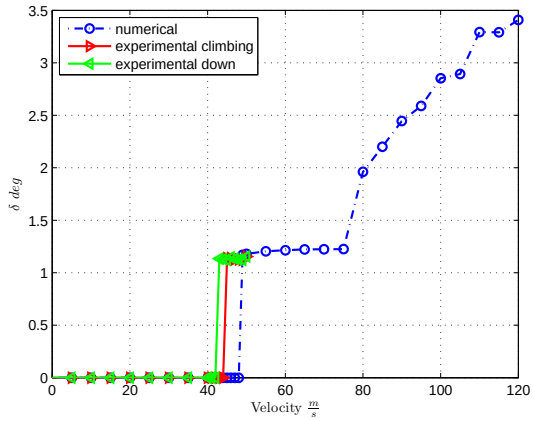
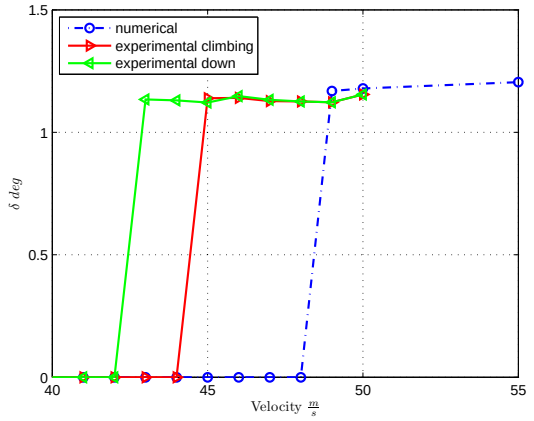


Figure 18: LCO signals FFT, $\pm 1°$, 50 m/s comparison.

In agree with the numerical results, also the LCO signals FFT shows the characteristic path of an odd nonlinearity (see Figure 18); the basic frequency of the first banding mode is clearly visible and the odd multiple too. Figure 19 shows the relation between LCO amplitude and the freestram speed; otherwise the numerical model, where the LCO starting and the stopping velocity is the same, in the experimental one, the LCO starting point occur at 45 m/s and the stopping velocity is 43 m/s. This difference is more evident for the second test case, shown in Figure 22. It can be shown analytically that the response of the nonlinear systems scale with the size of the freepaly region. To test this assumption, the $\pm 2°$ configuration has been studied too.

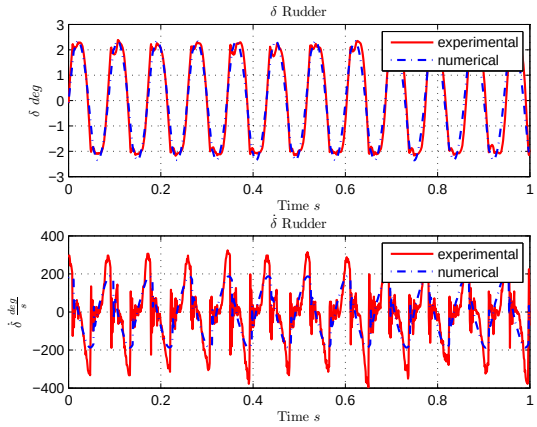


(a) whole

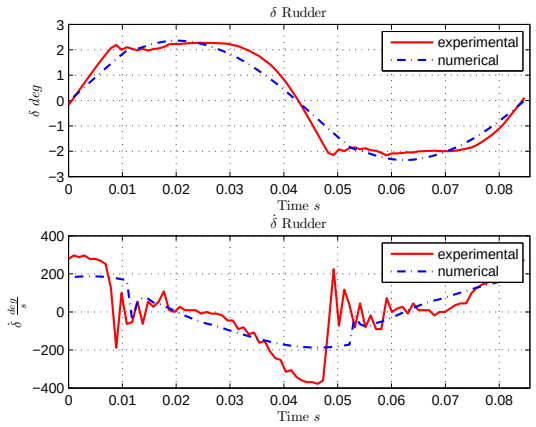


(b) zoom

Figure 19: LCO Velocity vs δ deg, $\pm 1°$ comparison.



(a) low frequency LCO 50 m/s, one second



(b) low frequency LCO 50 m/s, one period

Figure 20: Rudder LCO, $\pm 2°$ time series comparison.

In Figure 20 and 21 the time series and the phase portraits at 50 m/s compared are shown.

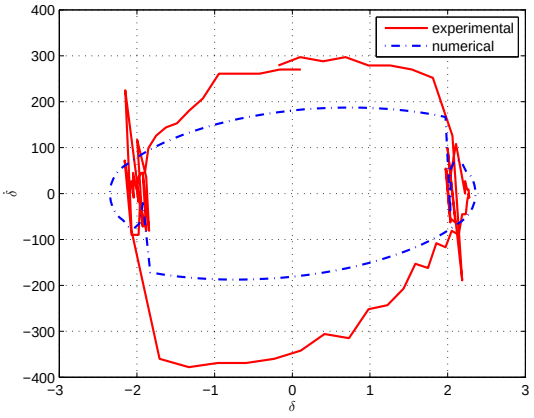
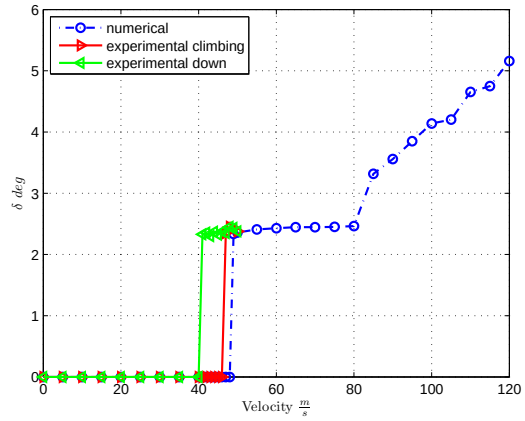
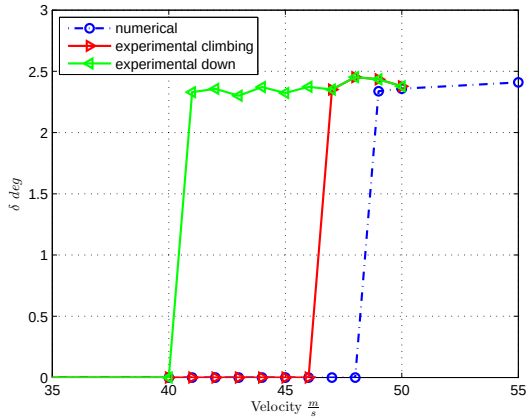


Figure 21: LCO phase portraits, $\pm 2°$, 50 m/s comparison.



(a) whole



(b) zoom

Figure 22: LCO Velocity vs δ deg, $\pm 2°$ comparison.

6 Conclusions

This paper has presented an experimental-numerical comparison for a T-tail in presence of freeplay nonlinearities in the rudder. An appropriate kinematical chain has been developed in order to obtain a mechanism that allows for a variable freeplay amplitude.

The comparison of responses for both experiments and corresponding numerical simulations has been shown and the numerical model has been validated by an experimental one.

It should be noted that the works presented here serve as verification and foundation for the next stage of this project, which is focused on the effects of structural nonlinearities on a flutter control system. This is why a control system has been added to the model to allow for prescribed movement of the rudder.

Acknowledgments

The authors would like to thank Paolo Mantegazza and Moti Karpel for their generous advice, Alessandro Scotti for his support during the experimental model construction.

References

[1] Dowell, E. H., Edwards, J., and Strganac, T., "Nonlinear Aeroelasticity," *Journal of Aircraft*, Vol. 40, No. 5, 2003, pp. 857–874.

[2] Croft, J., "Airbus Elevator Flutter: annoying or dangerous?" *Aviation Week and Space Technology*, Vol. 155, No. 9, August 2001, pp. 41.

[3] Dowell, E. H. and Tang, D., "Nonlinear Aeroelasticity and Unsteady Aerodynamics," *AIAA Journal*, Vol. 40, No. 9, 2002, pp. 1697–1707.

[4] Dowell, E. H., "Some recent advances in nonlinear aeroelasticity: fluid-structure Interaction in the 21st Century," *Proceedings of the 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Orlando, FL, 12–15 April 2010.

[5] Ricci, S., Scotti, A., and Zanotti, D., "Control of an all-movable foreplane for a three surfaces aircraft wind tunnel model," *Mechanical Systems and Signal Processing*, Vol. 20, July 2006, pp. 1044–1066.

[6] Ricci, S., Scotti, A., De Gaspari, A., and Riccobene, L., "Active Aeroelastic Control over a Multisurface Wing: Modelling and Wind-Tunnel Testing," *AIAA Journal*, Vol. 47, No. 9, 2009, pp. 1995–2010, doi: 10.2514/1.34649.

[7] Mattaboni, M., Quaranta, G., and Mantegazza, P., "Active Flutter Suppression for a Three Surface Transport Aircraft by Recurrent Neural Networks," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 4, 2009, pp. 1295–1307.

[8] Sandford, M. C., Ruhlin, C. L., and Yates, J. C. E., "Subsonic and Transonic Flutter and Flow Investigations of the T-Tail of a Large Multjet Cargo Airplane," Technical note D-4316, NASA, Langley Research Center Langley Station, Hampton, Va., February 1968.

[9] Sandford, M. C. and Ruhlin, C. L., "Experimental Parametric Studies of Transonic T-Tail Flutter," Technical note D-8066, NASA, Langley Research Center Langley Station, Hampton, Va., Decembrer 1975.

[10] Roger, K. L., "Airplane Math Modeling Methods for Active Control Design," Tech. Rep. CP–228, AGARD, August 1977.

[11] Fichera, S., *Non linearità di tipo freeplay in un modello di coda a T aeroelastico. Studio numerico e sperimentale.*, Master's thesis, Politecnico di Milano, Dipartimento di Ingegneria Aerospaziale, Milano, 2010, In italian.

[12] Hunt, K. and Crossley, F., "Coefficient of Restitution Interpreted as Damping in Vibroimpact," *Journal of Applied Mechanics*, Vol. 42, No. 2, June 1975, pp. 440–446.

[13] Lin, W. B. and Cheng, W. H., "Nonlinear flutter of loaded lifting surfaces (I) and (II)," *Journal of the Chinese Society of Mechanical Engineers*, Vol. 14, No. 5, 1993, pp. 446–466.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Robust aeroservoelastic analysis for the investigation of rotorcraft pilot couplings

V. Muscarello*, P. Masarati* and G. Quaranta*
Politecnico di Milano - Dipartimento di Ingegneria Aerospaziale, Italy
e–mail: *muscarello,masarati,quaranta@aero.polimi.it

Keywords: Rotorcraft Aeroservoelasticity, Robust Analysis, Rotorcraft Pilot Coupling.

Abstract

Aircraft-Pilot Couplings, namely the adverse interaction of the pilot with the aeromechanics of aircraft, represent a dangerous problem. The detection of an aircraft's proneness to this type of phenomenon is usually difficult prior to flight testing. Their occurrence in rotorcraft, called Rotorcraft-Pilot Couplings, has received less attention until recent years. This work applies robust stability analysis techniques, mutuated from robust control theory, to develop a simple and versatile technique that allows to investigate the boundaries of the pilot's transfer function that guarantee the stability of the coupled system regardless of the complexity of the vehicle model. The proposed method is applied to the analysis of a detailed model of a medium weight helicopter, and the results are discussed.

1 Introduction

Adverse interactions between pilot and aircraft include a wide class of phenomena identified since the very early days of aeronautics [1]. A recent review of the literature on the subject has been proposed in [2].

The investigation of Rotorcraft-Pilot Couplings (RPC) requires the capability to model aeroservoelastic phenomena as well as the dynamic behavior of the pilot. In detail, aeroservoelastic RPCs, often indicated as Pilot Assisted Oscillations (PAO), may occur in the frequency band between 2 and 8 Hz [3], and thus require to model the biodynamics of the pilot [4]. In the frequency range above 2 Hz the pilot may be modeled as a set of mechanical impedances between the motion of the seat and the resulting actuation of the control inceptors, since no voluntary action can be envisaged. Experimental results obtained so far have shown how pilot arms' response to vibrations is characterized by an high level of variability [4, 5]. As a consequence, it should be considered as an highly uncertain element in the dynamic modeling of this kind of problems.

The objective of this work is to exploit the robust control framework for the analysis of rotorcraft aeroservoelastic stability in presence of uncertain elements [6], where the biodynamic response of the pilot is considered uncertain. It is worth noticing that the 2–8 Hz frequency range contains characterizes the most significant rotorcraft dynamics, from flight mechanics to rigid-blade modes (see Figure 1). As a consequence, detailed aeroservoelastic models of the rotorcraft are necessary, and they need to cover all aspects of aeromechanics, like main and tail rotor aeroelasticity, structural dynamics, hydraulic control system, drive train and flight control system dynamics.

The paper presents results of RPC proneness

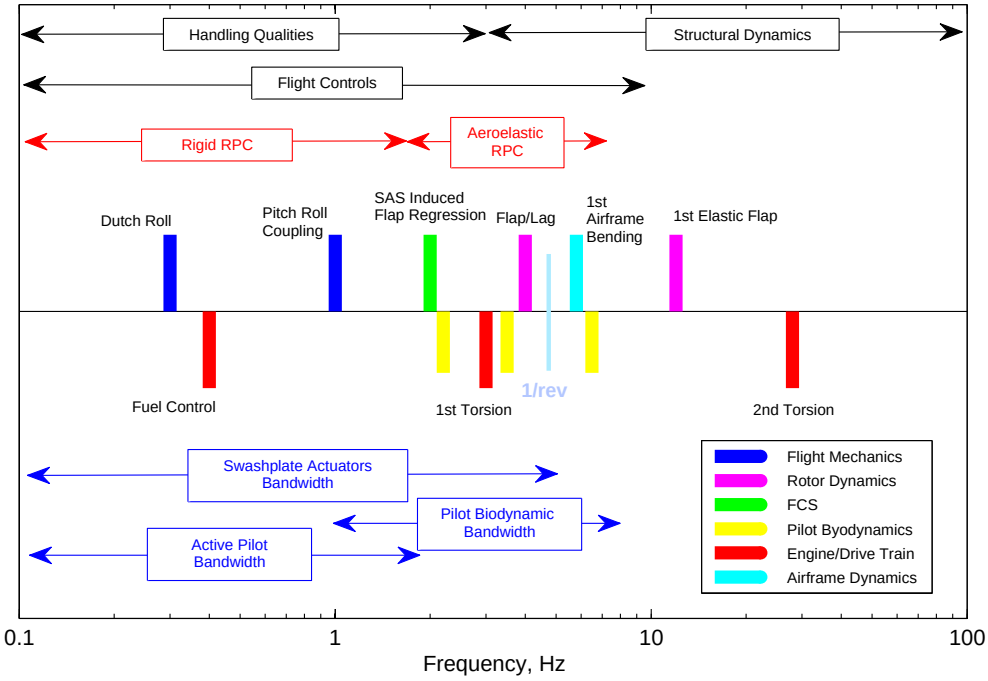


Figure 1: Typical rotorcraft frequencies.

analysis obtained using the robust stability analysis paradigm. The capability and the effectiveness of the proposed approach for the design of rotorcraft is considered, highlighting pros and cons of the different criteria used to compute the stability margins.

2 Robust Stability Margin Computation

Robust analysis allows to consider the effect of uncertainties on stability and performances of the system under investigation. More generally, using the definition given by Friedland [7], the degree of robustness can be defined as the level of immunity to uncertainty or changes of the plant. In the last thirty years, under the general framework of robust analysis, a plethora of methods have been investigated, especially in the automatic control community, that converged into the most renowned theories of $\mathcal{H}_\infty$ and μ [8]. Many authors applied these techniques to the aeroelastic stability of fixed wing aircraft in presence of uncertainty, for example [9, 10].

Parametric uncertainty description is usually considered. All system characteristics are assumed to depend on a set $\mathcal{D}$ of unknown but bounded parameters collected in a vector $\mathbf{p}$, often normalized for convenience, so that

$$\mathcal{D} = \{\mathbf{p} : \ \mathbf{p} = \mathbf{p}_0 + \delta\mathbf{p}, \ |\delta\mathbf{p}|_\infty \leq 1\}, \qquad (1)$$

where $\mathbf{p}_0$ corresponds to the nominal aircraft model without uncertainty. Physical insight and engineering judgment usually allow to estimate adequate boundaries of the parameters [9]. Consolidated numerical approaches for robust analysis, like μ-analysis, allow to compute worst case scenarios based on the given set $\mathcal{D}$.

Analyses of this kind are of utmost importance for control system design. In fact, to avoid expensive and time-consuming redesign at late stages of the development process it is always a good idea to consider from the beginning the unavoidable parameter variation of the controlled plant, and to design a control system capable of operating correctly by delivering the required performances in all the admissible aircraft configurations.

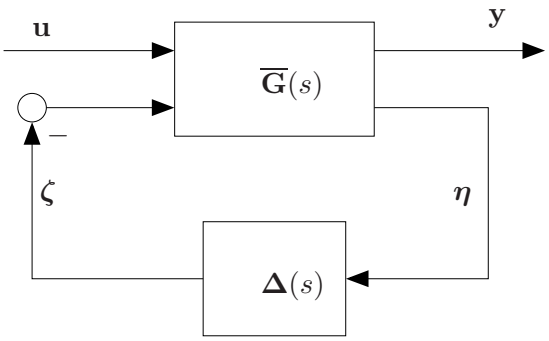


Figure 2: Negative feedback loop between nominal plant and uncertainty operator

However, when attempting to exploit the same type of tools for the design of the aircraft, a slightly different point of view could be more productive. Rather than looking for the worst case stability boundaries given the set $\mathcal{D}$, it would be preferable to have a tool capable of estimating the broadest set $\mathcal{D}$ of possible system parameter changes that can be accepted without compromising the stability of the system, or preserving a prescribed stability margin. This point of view will be exploited in the present work.

Classical robust analysis stems from the possibility to recast any uncertain dynamical, finite-dimensional, linear and time-invariant system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{2a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{2b}$$

as a feedback loop between the nominal plant $\overline{\mathbf{G}}(s)$ and the uncertainty, formulated as a complex matrix function $\boldsymbol{\Delta}(s)$, as shown in Figure 2. Different approaches can be followed, depending on whether only global $\mathcal{H}_\infty$ bounds of the $\boldsymbol{\Delta}(s)$ are known (these are usually termed *unstructured* uncertainty cases), or a more detailed knowledge on the structure of $\boldsymbol{\Delta}(s)$ and on scalar or block coefficients bounds is used (these are usually termed *structured* uncertainty cases).

This format can be recovered by using the Linear Fractional Transformation (LFT) matrix operations [8, 11], as soon as the uncertainties appear in rational polynomial form. After partitioning the nominal plant $\overline{\mathbf{G}}(s)$ with respect to the two subsets of input ($\mathbf{u}$, $\boldsymbol{\zeta}$) and output ($\mathbf{y}$, $\boldsymbol{\eta}$) signals (Fig. 2),

$$\begin{Bmatrix} \mathbf{y} \\ \boldsymbol{\eta} \end{Bmatrix} = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \boldsymbol{\zeta} \end{Bmatrix} \tag{3}$$

the input-output relationship between the signals $\mathbf{y}$ and $\mathbf{u}$ can be written as

$$\mathbf{y}(s) = \left(\mathbf{G}_{11} - \mathbf{G}_{12}\boldsymbol{\Delta} \left(\mathbf{I} + \mathbf{G}_{22}\boldsymbol{\Delta} \right)^{-1} \mathbf{G}_{21} \right) \mathbf{u}(s) \tag{4}$$

considering $\boldsymbol{\zeta} = -\boldsymbol{\Delta}\boldsymbol{\eta}$. If the nominal system $\overline{\mathbf{G}}$ is stable, the only potential source of instability in Eq. (4) is the term $(\mathbf{I} + \mathbf{G}_{22}\boldsymbol{\Delta})^{-1}$, so the stability problem simply requires to evaluate the stability of the $\mathbf{G}_{22}$-$\boldsymbol{\Delta}$ loop, noticing that both transfer matrices are square, since $\boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathbb{R}^n$. In the following the subscript $(\cdot)_{22}$ is dropped for clarity.

As for the classical Single Input Single Output (SISO), the stability analysis for a generic Multiple Input Multiple Output (MIMO) feedback system can be accomplished using the Generalized Nyquist Criterion (GNC) [12]. The uncertainty operator $\boldsymbol{\Delta}$ can be written as the combination of a nominal value, $\boldsymbol{\Delta}_0$ and an uncertainty part $\delta\boldsymbol{\Delta}$, namely

$$\boldsymbol{\Delta} = \boldsymbol{\Delta}_0 + \delta\boldsymbol{\Delta}, \tag{5}$$

such that the loopback plant is

$$\mathbf{H} = \mathbf{G}\boldsymbol{\Delta}. \tag{6}$$

Considering the multivariable feedback system, under the assumption that

(i) the open loop transfer matrix $\mathbf{G}(s)$ is stable,

(ii) the nominal system $\mathbf{H}_0$ is stable under unit feedback,

(iii) the number of poles of the perturbed system in the right half of the complex plane does not change,

the GNC states that the stability limit of the system is

$$\det\left(\mathbf{I} + \mathbf{H}(s)\right) = 0, \qquad \text{for } s = j\omega. \tag{7}$$

In other words, to avoid instability the eigenvalues of the uncertain system must not intersect

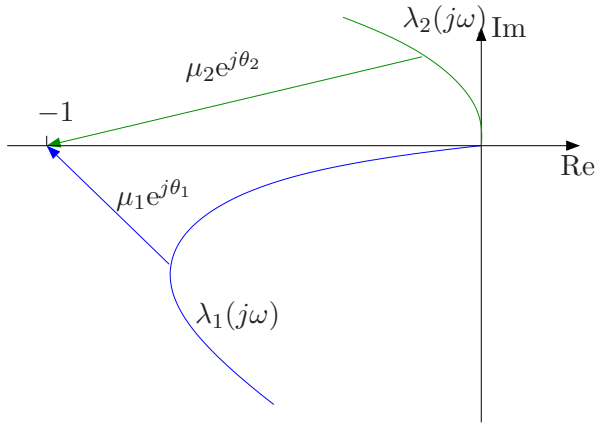


Figure 3: Nyquist eigenloci: identification of distance vectors from point $(-1+j0)$

the critical point $(-1+j0)$ on the complex plane, i.e.

$$\lambda_i\left(\mathbf{H}(j\omega)\right) \neq -1, \qquad \forall\omega \text{ and } i=1,\ldots,n \quad (8)$$

The meaning of this condition is clearly illustrated in a Nyquist plot, as shown in Figure 3. The uncertainty perturbs the nominal eigenloci curves $\lambda_i(j\omega)$, that show the evolution of the eigenvalues of the open loop transfer function as functions of the frequency ω, by adding a disturbance $\mu e^{j\theta}$. The stability margin of the generic eigenvalue $\lambda_i(j\omega)$ at each specific frequency ω is described by its distance $\mu_i e^{j\theta_i}$ from the critical point $(-1+j0)$.

Conversely, considering a stable plant, the maximum level of uncertainty that can be accepted without driving the system to instability can be evaluated.

The determinant of Eq. (7) is zero if any of the eigenvalues of the perturbed system is equal to -1. By parametrizing the uncertain eigenvalues with

$$z_i = \lambda_i(\mathbf{H}_0(j\omega)) + \mu e^{j\,\theta}, \quad (9)$$

the values of $\mu \in \Re$ and $\theta \in \Re$ that drive the i-th eigenvalue to instability, μ_i and θ_i, can be computed by setting $z_i = -1$ in Eq. (9),

$$\mu_i(\omega) e^{j\,\theta_i(\omega)} = -\left(\lambda_i(\mathbf{H}_0(j\omega)) + 1\right). \quad (10)$$

So the stability of the system is ensured whenever for

$$\theta \equiv \theta_i(\omega) = -\tan^{-1}\left(\frac{\Im\left(\lambda_i(\mathbf{H}_0(j\omega))\right)}{\Re\left(\lambda_i(\mathbf{H}_0(j\omega))\right)+1}\right) \quad (11)$$

the condition

$$\mu(\omega) < \mu_i(\omega) = \|\lambda_i(\mathbf{H}_0(j\omega)) + 1\| \quad (12)$$

is met. Of course, of all eigenvalues only the one closest to the point $(-1+j0)$ in the complex plane at each frequency, i.e. the one with the smallest μ_i, determines the stability margins at that frequency:

$$\mu_s(\omega) = \min_i \mu_i(\omega). \quad (13)$$

The quantity $\mu_s(\omega)$ represents a stability margin index, because when it goes to zero the perturbed system hits the stability boundary.

Other stability margin indexes can be obtained if the following parametrization is used for the uncertain eigenvalues

$$z_i = \lambda_i(\mathbf{H}_0(j\omega))\, \rho e^{j\,\phi}. \quad (14)$$

In this way the values of $\rho \in \Re$ and $\psi \in \Re$ that drive the i-th eigenvalue to instability, ρ_i and ψ_i, can be computed by setting $z_i = -1$ in Eq. (14),

$$\rho_i(\omega) e^{j\,\phi_i(\omega)} = -\frac{1}{\lambda_i(\mathbf{H}_0(j\omega))}. \quad (15)$$

Also in this case the stability of the system is ensured whenever for

$$\phi \equiv \phi_i(\omega) = -\arg\left(\frac{1}{\lambda_i(\mathbf{H}_0(j\omega))}\right) \quad (16)$$

the condition

$$\rho(\omega) < \rho_i(\omega) = \left\|\frac{1}{\lambda_i(\mathbf{H}_0(j\omega))}\right\| \quad (17)$$

is met. Of course, of all eigenvalues only the one closest to the point $(-1+j0)$ in the complex plane at each frequency, i.e. the one with ρ_i closest to 1, determines the stability margins at that frequency:

$$\rho_s(\omega) = \min_i |\rho_i(\omega) - 1|. \quad (18)$$

In this case ρ_s is closely related to the the classical *gain margin* defined for SISO systems [7]. However, this margin is frequency dependent since the role of the *loop gain* is taken by a quantity that is frequency dependent, i.e. the uncertainty $\mathbf{\Delta}(s)$. However, at the frequency ω where

$\phi_{\rho_s}(\omega) = -180$ deg, the quantity $|\rho_s(\omega) - 1|$ is exactly the gain margin. Similarly, $\phi_{\rho_s}(\omega)$ is related to the *phase margin* [7], since it expresses the phase lag required to drive the system unstable. In detail, for the value of ω where $|\rho_s(\omega)-1| = 0$, $180+\phi_{\rho_s}(\omega)$ is exactly the phase margin.

This information is very useful as it quantifies how the perturbation is prone to driving the system unstable, thus intrinsically accounting for the sensitivity of the stability margin of the system on the uncertain parameters. It should be more useful to designers than the bare knowledge of the entire eigenspectrum of the time realization of the plant model of Eq. (3), since the latter may contain modes with small damping but whose dependence on the uncertain parameters is negligible, and thus are not critical with respect to stability, while the damping of other modes, nominally higher than that of the previous ones, could show a significant sensitivity on the uncertain parameters. To obtain a similar amount of information from the eigenspectrum of the problem, for the generic i-th eigensolution of the original problem of Eq. (2), augmented by an eigenvector normalization equation,

$$(\lambda_i \mathbf{I} - \mathbf{A})\,\mathbf{X}_i = \mathbf{0} \tag{19a}$$

$$\frac{1}{2}\mathbf{X}_i^* \mathbf{X}_i = 1, \tag{19b}$$

where $(\cdot)^*$ indicates the complex conjugate transpose, one should evaluate the sensitivity with respect to the uncertain parameter p_j,

$$(\lambda_i \mathbf{I} - \mathbf{A})\,\mathbf{X}_{i/p_j} + \mathbf{X}_i \lambda_{i/p_j} = \mathbf{A}_{/p_j}\mathbf{X}_i \tag{20a}$$

$$\mathbf{X}_i^* \mathbf{X}_{i/p_j} = 0 \tag{20b}$$

and then, using a continuation algorithm, determine what value of the parameter perturbation results in a null or purely imaginary value of λ_i.

In order to maximize the information gained, the *Generalized Stability Margins* (GSM) μ_s and θ_{μ_s}, or the corresponding ρ_s and ψ_{μ_s}, can be translated into bounds of the physical quantities that are included in the structured uncertainty operator $\mathbf{\Delta(s)}$. This can be easily obtained by considering that imposing that the eigenvalues of the perturbed system be $z_i = -1$ for all i corresponds to the condition

$$\mathbf{H}_0(j\omega) + \delta\mathbf{H}(j\omega) = -\mathbf{I}. \tag{21}$$

Remembering the definitions of right ($\mathbf{v}$) and left ($\mathbf{u}$) eigenvalues, $\lambda_i \mathbf{v}_i = \mathbf{H}_0 \mathbf{v}_i$, $\mathbf{u}_i^T \lambda_i = \mathbf{u}_i^T \mathbf{H}_0$ with $\mathbf{u}_i^T \mathbf{v}_j = \delta_{ij}$, $\delta_{ij} = 1$ when $i = j$, 0 otherwise, Eq. (21) can be multiplied by $\mathbf{u}_i^T$ on the left and $\mathbf{v}_i$ on the right, obtaining

$$\lambda_i(\omega) + \mathbf{u}_i^T(\omega)\delta\mathbf{H}(j\omega)\mathbf{v}_i(\omega) = -1. \tag{22}$$

Eqs. (11) and (10) yield

$$\mu_i(\omega) e^{j\theta_i(\omega)} = \mathbf{u}_i^T(\omega)\delta\mathbf{H}(j\omega)\mathbf{v}_i(\omega). \tag{23}$$

Consequently, the stability of the system is guaranteed when the inequality

$$\mathbf{u}_i^T(\omega)\mathbf{G}(j\omega)\delta\mathbf{\Delta}(j\omega)\mathbf{v}_i(\omega) < -1 \tag{24}$$

holds.

Equation (24) is a complex coefficient inequality, that results in two constraints between the parameters $\mathbf{p}$ that must be satisfied to have instability at frequency ω. These two inequalities define, at each frequency ω, a region of the parameter space of acceptable parameters. The knowledge of this region can be very helpful for the designer, especially when dealing with elements that present a wide range of characteristics that depend on several factors like environmental or operating conditions, wear and so on.

A significant advantage of this type of analysis is the fact that it is performed in the traditional, often familiar frequency-domain. In fact only frequency response data are necessary, opening the door to complementing numerical models with experimental data.

During the design process this type of analysis can be proficiently performed considering only the uncertainty on the parameters that are the main focus of that specific design phase. However, other parameters of the system are usually uncertain as well; moreover, the actual behavior of the system may differ from the linear model used in the analysis by higher-order terms neglected in the analysis. As a consequence, it is not a good idea to design the system too close to the stability margin. On the contrary, it may be

desirable to maintain a minimal distance from zero for the GSM. In this case it is useful to draw the lines representing all points in the parameter space that have an assigned amplitude of the stability margin $\tilde{\mu}$. This can be obtained by computing the *critical direction* unit vector $\mathbf{d}$

$$\mathbf{d}(j\omega) = \frac{\lambda_i(\mathbf{H}_0(j\omega)) + 1}{|\lambda_i(\mathbf{H}_0(j\omega)) + 1|} \tag{25}$$

first, and then imposing that the eigenvalue z_i of the uncertain system discussed in (9) is equal to $z_i = -\tilde{\mu}\mathbf{d}$. As a consequence a new stability margin can be defined

$$\hat{\mu}_i(\omega)\mathrm{e}^{j\hat{\theta}_i(\omega)} = -\left(\lambda_i(\mathbf{H}_0(j\omega)) + \tilde{\mu}\mathbf{d}\right) \tag{26}$$

yielding new constraints

$$\mathbf{u}_i^T(\omega)\mathbf{G}(j\omega)\delta\mathbf{\Delta}(j\omega)\mathbf{v}_i(\omega) < -\tilde{\mu}\mathbf{d} \tag{27}$$

for the physical parameters.

3 Multidisciplinary Helicopter Models for Aeroservoelastic Analysis

The type of analyses described in the previous paragraph require reliable numerical models. They must allow to investigate the interaction among different subsystems like deformable mechanical components, servo-hydraulic elements, unsteady aerodynamic forces, pilot models, control logic and so on. Instead of developing yet an entirely new rotorcraft aeroservoelastic simulation software, a simulation tool called MASST (Modern Aeroservoelastic State-Space Tools) [13, 14] has been built on top of a popular general-purpose mathematical environment. The tool can perform massive analyses of relatively simple, yet complete modular models of complex linearized aeroservoelastic systems. Each separate component of the model consists of submodels collected from well-known, reliable and possibly state-of-art sources, rather than derived from first principles equations; they are blended together in a mathematical analysis environment. Heterogeneous sub-components can be imported from different sources, to model: (1) deformable airframe structural dynamics; (2) airframe unsteady aerodynamics; (3) rotor aeroelasticity; (4) drive train; (5) servo-actuators; (6) sensors and filters; (7) Flight Control System; (8) pilot biomechanics. Items 1–4 provide the core of basic aeroelastic analysis capabilities. Items 1–5 provide aeroservoelastic analysis capabilities. All items provide closed-loop aeroservoelastic capabilities.

3.1 Model Description

As a test bed for RPC analysis, an aeroservoelastic model of the SA330 Puma has been built in MASST, using Refs. 15 and 16 as data sources. The hover condition at sea level has been considered.

The *airframe structural model* is characterized by 6 rigid body modes (Fore/Aft, Lateral, Plunge, Roll, Pitch and Yaw). The model is composed only by the mass, damping and stiffness modal matrices plus the displacements at few noteworthy points like the center of the main and tail rotor hubs and the sensors at the pilot and co-pilot seats.

The *main rotor aeroelastic model* is derived using a comprehensive rotorcraft solver, able to generate the rotor linear matrices considering different equilibrium configurations. The rotor aeroelastic models are represented as

$$\mathbf{A}_2\ddot{\mathbf{q}}_r + \mathbf{A}_1\dot{\mathbf{q}}_r + \mathbf{A}_0\mathbf{q}_r = \mathbf{B}_g\mathbf{v}_g + \mathbf{f}_c. \tag{28}$$

where $\mathbf{q}_r$ are global rotor degrees of freedom chosen using the Ritz decomposition, matrices $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{B}_g$ are Linear Time Invariant (LTI), computed using coefficient averaging to eliminate any periodicity whenever the rotor is not in axial flow conditions, and vector $\mathbf{f}_c$ represents the forces applied by the servo-actuators on the rotor to control the blade collective and cyclic pitch angles. The state vector $\mathbf{q}_r$ contains the bending and torsion rotor modes, expressed in the non-rotating reference frame using the multiblade transform, plus six rigid body motion modes used as *constraint modes*, to connect the rotor to the corresponding airframe model with the Craig-Bampton Component Mode Synthesis (CMS) approach [17].

Collective, cyclic and scissor modes have been considered for the 4 bladed soft-in-plane main

rotor. In detail, the rotor has been modeled considering 3 bending modes, 2 torsion modes, and the 6 hub/pylon rigid modes. The complete model is characterized by 26 degrees of freedom.

The main rotor aeroelastic model has been verified by comparing the blade frequencies in vacuum and the rotor performances in hover with analogous results obtained using the general-purpose multibody solver MBDyn http://www.mbdyn.org/. Figure 4 shows the

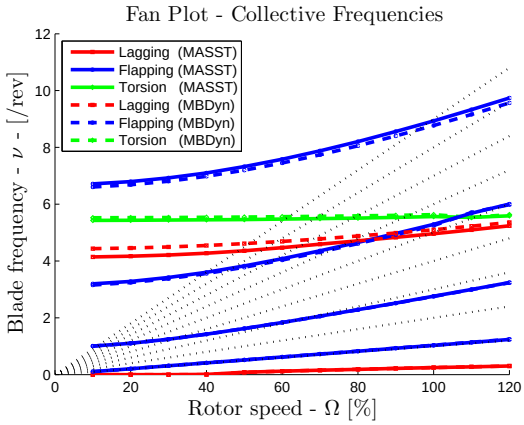


Figure 4: Fan Plot - Main rotor collective frequencies

blade natural collective frequencies compared with the blade frequencies obtained in MBDyn (dashed lines). These results correlate well also with those reported in [15], obtained by various comprehensive rotorcraft codes. Figure 5 shows the main rotor polar in hover.

The rotor trim to find the collective pitch needed to match the rotorcraft weight is obtained using the multibody MBDyn model.

The *tail rotor* is modeled by a dummy feedback loop characterized by a static gain scheduled with the flight speed, that represents the control derivative of the side force generated by the tail rotor as a function of the tail rotor collective pitch control. The tail rotor lateral control derivative $Y_{\theta_{0T}}$ (see [16]) is shown in Figure 6, dimensionalized with the rotorcraft mass, as a function of the flight speed. In hover, only the control derivative at $V_\infty = 0$ is used.

Typical linear *servo actuator* transfer functions are defined for the three actuators of the main rotor swashplate and for the single actua-

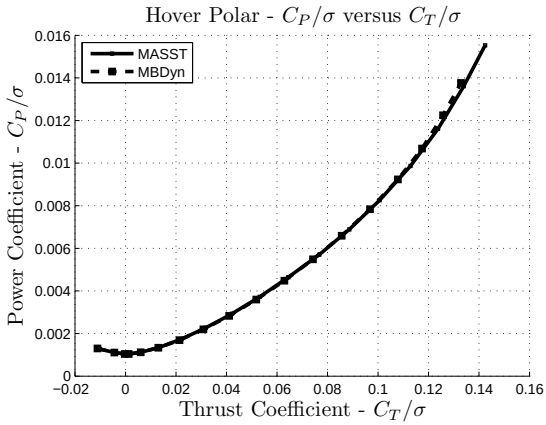


Figure 5: Main Rotor Performances - Hover polar

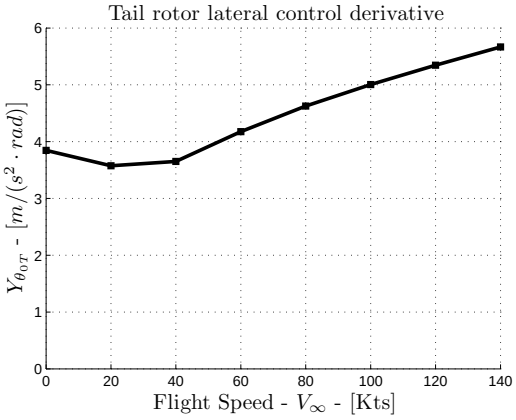


Figure 6: Tail Rotor - Lateral control derivative

tor of the tail rotor. The transfer functions of a servo actuator describe the displacement of the servo actuator piston x, as a function of the requested displacement x_c, and of the generalized reaction force applied by the dynamics of the rotor controls, f_c, namely

$$x = H_x(s)\, x_c + H_f(s)\, f_c. \tag{29}$$

In general, both the servo-valve and compliance dynamics are fully represented by transfer functions $H_x(s)$ and $H_f(s)$. In this work, the servo-valve transfer functions of main and tail rotor actuators are modeled as second order low-pass filters

$$H_x(s) = \frac{1}{(1+s\tau_1)(1+s\tau_2)} \tag{30}$$

with $\tau_1 = 0.1$ [sec] and $\tau_2 = 0.016$ [sec], while only the static contribution of the dynamic compliance, written as a function of the static stiffness, has been taken into account

$$H_f(s) = \frac{1}{K_{act}} \tag{31}$$

with $K_{act} = 3 \cdot 10^7$ [N/m] for the main rotor servo actuators and $K_{act} = 6.6 \cdot 10^6$ [N/m] for the single actuator of the tail rotor. A gear ratios matrix, called *"Servo To Blade"*, is used to convert the actuator displacements into collective and cyclic commands.

The model includes six *sensors* for the accelerations at the pilot and co-pilot seats in longitudinal, lateral and vertical direction. An additional dummy sensor is defined to measure the tail rotor collective pitch generated by the servo actuator. Its output is used as the input of the tail rotor controller. An *external force* is used to model the lateral force applied at the tail rotor hub and driven by the tail rotor dummy controller.

4 Robust Stability for Rotorcraft-Pilot Coupling

The robust stability of the SA300 Puma model has been analyzed considering the aircraft in hover and investigating the feedback loop between the vertical acceleration at the pilot seat and the collective pitch introduced by the pilot. As a consequence the uncertain operator is a scalar, and represents the transfer function of the pilot biodynamics.

The baseline helicopter model contains two slightly unstable modes, the phugoid mode around 0.05 Hz and the Dutch roll mode around 0.07 Hz. Since these modes are in a frequency range where the passive biodynamic response of the pilot is not effective, they have been eliminated from the numerical model during a preliminary model reduction phase. As a result, only stable modes are present in the model; this allows to comply with the stable open loop plant hypothesis.

Additionally, a nominal pilot model has been considered, consisting of a simple constant gain $\Delta_0 = 0.0077$ chosen to ensure the satisfaction of the stability of the nominal plant for unit feedback loop. The Nyquist plot of the nominal plant for the frequency range between 0.0 and 40 Hz is shown in Figure 7.

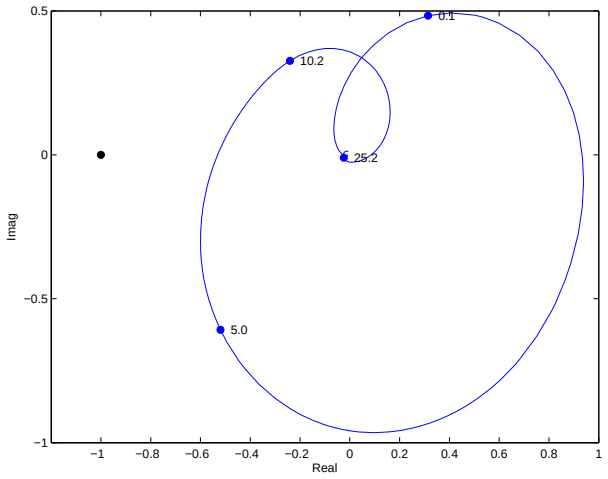


Figure 7: Nyquist plot of the SA330 Puma transfer function between the collective pitch and the vertical acceleration at pilot's seat.

This diagram allows to compute the stability margin index defined in Eq. (12), which is shown in Figure 8. The limit values of $\delta\Delta$ at each frequency can be computed using Eq. (24).

Finally the limit values for the pilot biodynamic frequency response at each ω can be computed as $H_{pb}(j\omega) = \Delta_0 + \delta\Delta$. The Nyquist plot of this frequency response is shown in Fig-

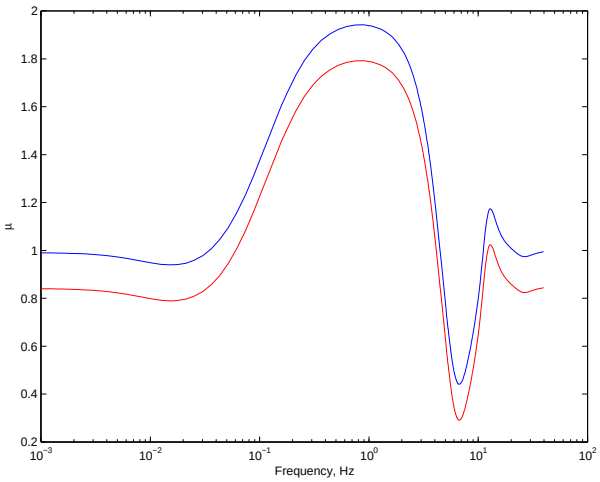


Figure 8: Stability margin index $\mu(\omega)$; the blue line represents the stability margin while the red line represents the margin obtained by imposing a distance of 0.15 from the point $(-1. + j0)$

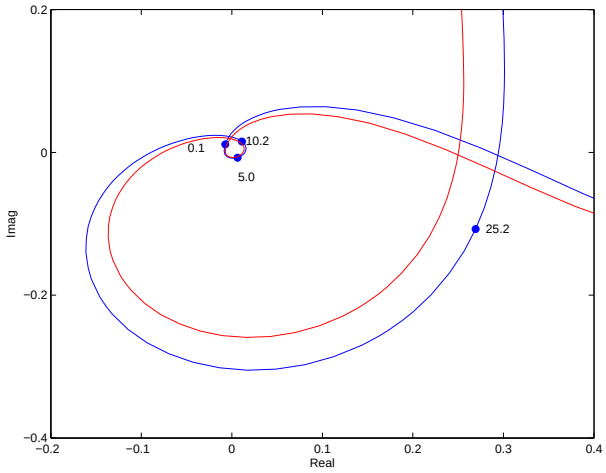


Figure 9: Nyquist plot of the limit frequency response for the pilot model; the red line represents the limit frequency response obtained by imposing a distance of 0.15 from the point $(-1 + j0)$.

ure 9. This meaningful frequency response function can be better described by looking at the amplitude and phase diagrams shown in Figure 10 as functions of the frequency. It is interesting to note that in the frequency range that characterizes the pilot's biodynamics, from 1 Hz on, the limit amplitude is quite flat. This indicates that in this entire range of frequency a problem may be encountered. The points of the limit pilot transfer function that are at zero phase represent a sort of gain margins. Looking at Figure 10 one can see that there are three points of this kind. The one with minimal margin is close to 7.3 Hz, with a limit pilot amplitude of 0.0144. As a consequence, if we consider a pilot model composed by a simple scalar gain, i.e. $H_p(s) = G_p$, any pilot that shows a gain higher than $G_p = 0.0144$ will drive the system towards a PAO instability. This effect is visible in the Nyquist diagram of Figure 11, where the Nyquist plots for scalar gain pilot model at gain values slightly below and slightly above the limit value. The figure shows how the latter is unstable.

It is interesting to see that this gain is not equal to the minimum amplitude of the limit pilot transfer function. In fact, the minimum point is at 2.53 Hz with an amplitude of 0.0077 (i.e. half the margin at 7.3 Hz). However, the sys-

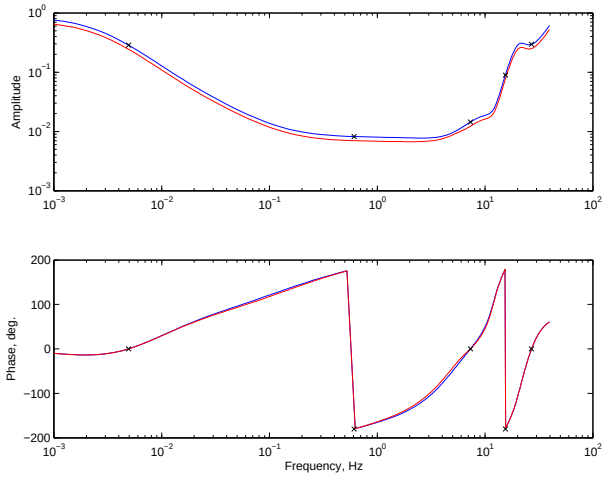


Figure 10: Amplitude and phase for the limits of pilot biodynamic models; the red line represents the limit frequency response obtained by imposing a distance of 0.15 from the point $(-1 + j0)$.

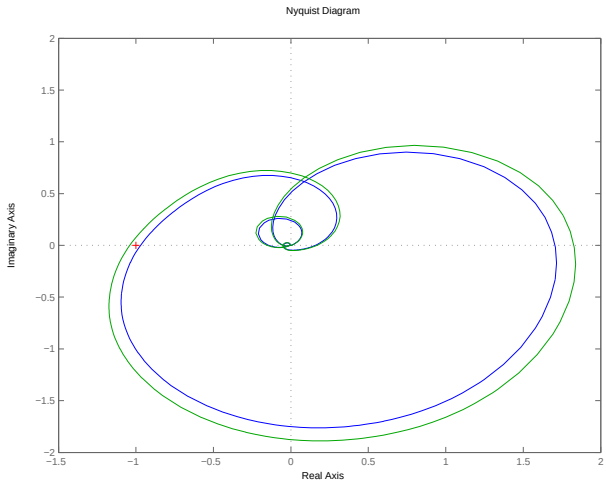


Figure 11: Nyquist plots of rotorcraft-pilot system for scalar gain pilot models. Blue line: pilot gain $G_p = 0.014$; Green line: $G_p = 0.015$.

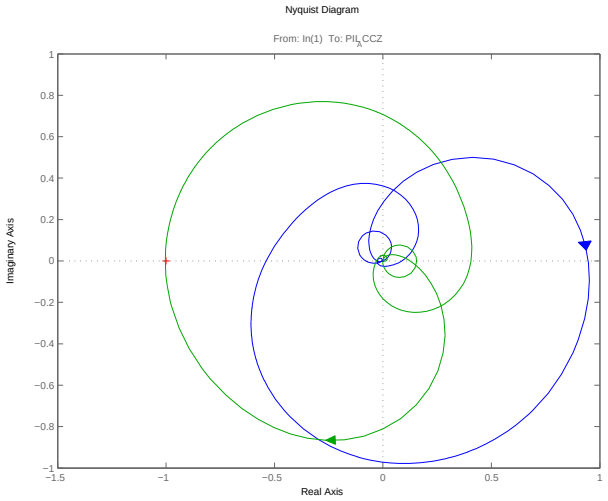


Figure 12: Nyquist plots of rotorcraft-pilot system for simple pilot models. Blue line: pilot gain $H_p = 0.0077$; Green line: $H_p = 0.0077\mathrm{e}^{j\psi}$, ψ corresponding to -121 deg.

tem in this case will be unstable only if the pilot model also introduces a phase delay of about 121 deg. Figure 12 shows how a scalar gain pilot with a gain equal to $G_p = 0.0077$, leads to a robustly stable system. However if a delay term is added, which means that the pilot model becomes $H_p = G_p\mathrm{e}^{j\psi}$ with ψ corresponding to -121 deg, the Nyquist plot crosses the point $(-1+j0)$.

The results shown in Figure 10 can be even more useful if a more complex pilot model with a behavior that changes at each frequency is considered. Looking only at the amplitude, it is possible to obtain a sufficient stability condition by stating that if the pilot frequency response is at all frequencies below the limit pilot amplitude, the pilot will not cause an instability. On the contrary, if the pilot frequency response crosses the limit pilot amplitude curve, the system may become unstable if the appropriate phase lead/lag is introduced.

In order to understand how to use this design criterion let us consider the pilot transfer function proposed by Mayo [18]. The so-called 'Ectomorphic' (smaller size) model gives the absolute acceleration of the hand that holds the collective lever with respect to the acceleration of the pilot's seat:

$$H^{\mathrm{ecto}}(s) = \frac{\ddot{z}_{\mathrm{coll}}}{\ddot{z}_{\mathrm{seat}}} = \frac{5.19s + 452.3}{s^2 + 13.70s + 452.3}. \quad (32)$$

To obtain the relative rotation of the collective bar this transfer function must be integrated twice, after subtracting the acceleration of the seat, and divided by the collective bar length L, and by the angle θ_L (in radians) spanned by the lever to go from $\eta = 0\%$ to $\eta = 100\%$, and multiplied by the gain G_c that represent the gear ratio between the collective lever percentage and the variation of the blade collective pitch,

$$\begin{aligned} H^{\mathrm{ecto}}_{\eta\ddot{z}_{\mathrm{seat}}} &= \frac{G_c}{s^2\theta_L L}\left(\frac{\ddot{z}_{\mathrm{coll}}}{\ddot{z}_{\mathrm{seat}}} - 1\right) \\ &= \frac{G_c}{s^2\theta_L L}\frac{-s(s+8.51)}{(s^2 + 13.70s + 452.3)}. \end{aligned} \quad (33)$$

However, the presence of a double pure integrator yields a drifting behavior when $s \to 0$. Considering that the low frequency behavior would be characterized by the already presented active pilot model, which by the way is needed to control the unstable phugoid and Dutch roll modes, the function is high-pass filtered by changing the integrator poles into real ones at a typical value of pilot crossover frequency $\omega_c = 3.10$ radian/s

for the control task on the vertical axis [19]:

$$H_{\theta_c \ddot{z}_{\text{seat}}} = \frac{-G_c(s^2 + 8.51s)}{\theta_L L(s+\omega_c)^2(s^2 + 13.70s + 452.3)}. \tag{34}$$

Let us consider a collective bar $L = 350$ mm long, that travels for 35 deg, such that $\theta_L = (35/180)\pi$, and a blade collective pitch range of 18 deg, yielding $G_c = \pi/10$. Figure 13 shows the plot of Mayo's Ectomorphic pilot transfer functions, modified according to Eq. (34) and multiplied respectively by 1.6 and by 0.2, compared to the limit pilot diagram. This represent a range of uncertain pilot models, where the uncertainty is related to the gain. In the case of large pilot

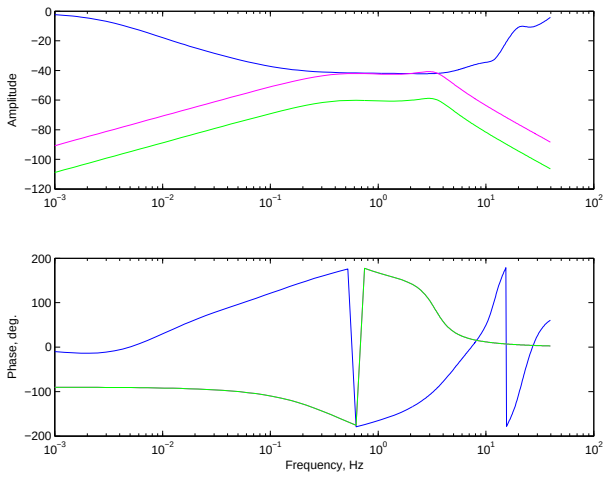


Figure 13: Amplitude and phase for the limits of pilot biodynamic models with Mayo's transfer functions overlapped. Green line: Mayo's model with $G_c = .2\pi/10$; Magenta line: $G_c = 1.6\pi/10$

gain, the pilot transfer function crosses the limit pilot curve. A conservative design must consider the possibility to reduce the gain, for instance by increasing the length of the collective bar, or its angular range, compatibly with ergonomic considerations.

The Nyquist plot for the rotorcraft model in series with Mayo's pilot model at gain $G_c = 1.6\pi/10$ is shown in Figure 14. In this case the system is not unstable but very small stability margins are shown. A gain increase of 5% destabilizes the system, because a point of Mayo's pilot frequency response function is coincident in terms of amplitude and phase with one of the limit pilot curve. In this case the instability is at a very low frequency (about 0.6 Hz, close to the crossover frequency ω_c).

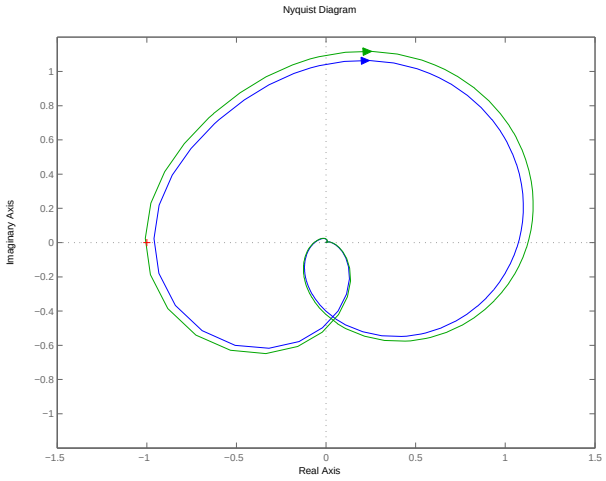


Figure 14: Nyquist plots of rotorcraft-pilot system for Mayo's pilot models. Blue line: pilot gain $G_c = 1.6\pi/10$; Green line: $G_c = 1.05 \times 1.6\pi/10$.

5 Concluding Remarks

This work presents a novel approach for rotorcraft design that makes extensive use of robust control theory to determine parameter bounds that may help the designer when highly uncertain elements, like pilots' biodynamic response, must be taken into account. The work discusses how limits on pilot models can be inferred from a rotorcraft model, and how this information can be useful for design. Further design criteria with different levels of conservativeness can be considered. Up to now, the information related to the phase angle of the limit pilot model have not been significantly exploited. Future research will consider how this additional piece of information can be exploited to further reduce the conservativeness of the approach.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme

[FP7/20072013] under grant agreement no. ACPO-GA-2010-266073.

References

[1] D. G. Mitchell, D. H. Klyde. Recommended practices for exposing pilot-induced oscillations or tendencies in the development process. Proceedings of the USAF Developmental Test and Evaluation Summit, 1–20. Woodland Hills, California, November 16–18 2004. AIAA 2004–6810.

[2] M. D. Pavel, J. Malecki, B. DangVu, P. Masarati, G. Quaranta, M. Gennaretti, M. Jump, H. Smaili, A. Ionita, L. Zaicek. Aircraft and rotorcraft pilot coupling: a survey of recent research activities within the European project ARISTOTEL. Proceedings of the 3rd CEAS Air & Space Conference. Venice, Italy, October 24–28 2011.

[3] M. Gennaretti, J. Serafini, P. Masarati, G. Quaranta, O. Dieterich. Aeroelastic and biodynamic modeling for stability analysis of rotorcraft-pilot coupling phenomena. Proceedings of the 34th European Rotorcraft Forum. Liverpool, UK, September 16–19 2008.

[4] M. Jump, S. Hodge, B. DangVu, P. Masarati, G. Quaranta, M. Mattaboni, M. Pavel, O. Dieterich. Adverse rotorcraft-pilot coupling: The construction of the test campaigns at the University of Liverpool. Proceedings of the 34th European Rotorcraft Forum. Liverpool, UK, September 16–19 2008.

[5] M. Mattaboni, G. Quaranta, P. Masarati, M. Jump. Experimental identification of rotorcraft pilots' biodynamic response for investigation of PAO events. Proceedings of the 35th European Rotorcraft Forum, 1–12. Hamburg, Germany, September 22–25 2009.

[6] P. Masarati, V. Muscarello, G. Quaranta. Robust aeroservoelastic stability of helicopters: Application to the air/ground resonance problem. Proceedings of the American Helicopter Society 67th Annual Forum. Virginia Beach, VA, USA, May 3–5 2011.

[7] B. Friedland. *Control System Design. An Introduction to State Space Methods*. McGraw Hill, New York, NY, 1987.

[8] K. Zhou, J. Doyle. *Essentials of robust control*. Prentice Hall, Upper Saddle River, NJ, 1998.

[9] R. Lind, M. Brenner. *Robust Aeroservoleastic Stability Analysis*. Advances in Industrial Control. Springer, 1999.

[10] D. Borglund. The μ-k method for robust flutter solutions. *Journal of aircraft*, **41**(5), 1209–1216, 2004. ISSN 0021-8669.

[11] J. Cockburn, B. Morton. Linear Fractional Representations of Uncertain Systems. *Automatica*, **33**(7), 1263–1271, 1997.

[12] C. Desoer, Y. Wang. On the generalized Nyquist stability criterion. *Automatic Control, IEEE Transactions on*, **25**(2), 187–196, 1980. ISSN 0018-9286.

[13] P. Masarati, V. Muscarello, G. Quaranta. Linearized aeroservoelastic analysis of rotary-wing aircraft. Proceedings of the 36th European Rotorcraft Forum, 099.1–10. Paris, France, September 7–9 2010.

[14] P. Masarati, V. Muscarello, G. Quaranta, A. Locatelli, D. Mangone, L. Riviello, L. Viganò. An integrated environment for helicopter aeroservoelastic analysis: the ground resonance case. Proceedings of the 37th European Rotorcraft Forum, 177.1–12. Gallarate, Italy, September 13–15 2011.

[15] W. G. Bousman, C. Young, F. Toulmay, N. E. Gilbert, R. C. Strawn, J. V. Miller, T. H. Maier, M. Costes, P. Beaumier. A comparison of lifting-line and CFD methods with flight test data from a research Puma helicopter. *TM 110421*, NASA, October 1996.

[16] G. D. Padfield. *Helicopter Flight Dynamics: the Theory and Application of Flying Qualities and Simulation Modeling*. Blackwell, Oxford, UK, 1996.

[17] R. R. Craig, Jr., M. C. C. Bampton. Coupling of substructures for dynamic analysis. *AIAA Journal*, **6**(7), 1313–1319, July 1968.

[18] J. R. Mayo. The involuntary participation of a human pilot in a helicopter collective control loop. Proceedings of the 15th European Rotorcraft Forum, 81.1–12. Amsterdam, The Netherlands, 12–15 September 1989.

[19] D. T. McRuer, H. R. Jex. A review of quasi-linear pilot models. *Human Factors in Electronics, IEEE Transactions on*, **HFE-8**(3), 231–249, September 1967. Doi:10.1109/THFE.1967.234304.

CEAS 2011

ATM Airport Operations 1

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Flight testing of noise abating RNP procedures and steep approaches

Dr. Helmut H. Toebben, Vilmar Mollwitz, Dr. Bernd Korn, Prof. Dr. Dirk Kügler
DLR German Aerospace Center, Institute of Flight Guidance, Lilienthalplatz 7, 38108 Braunschweig, Germany

Keywords: *steep approach, curved approach, noise*

Abstract

To test different types of noise abating approach procedures the Institute of Flight Guidance performed flight tests on the 6th September 2010 with a Boeing 737-700. In total 13 approaches to the Research Airport in Brunswick were flown while the approach area of the airport was equipped with six noise measurement microphones. Brunswick airport is equipped with an experimental ground based augmentation system (GBAS) which allows the implementation of 48 ILS lookalike precision approach procedures with different approach angles simultaneously.

1 Introduction

One of the major concerns regarding the expected growth in air traffic is the increase of air pollution and the related climate change as well as the increase of noise especially in the vicinity of airports.

Actually there are two large research programmes in Europe which address this development. On the one hand it is the Joint Technology Initiative "Clean Sky" which will develop breakthrough technologies related to the aircraft itself to reduce environmental impact. On the other hand it is the SESAR (Single European Sky ATM Research) programme which is the technological and operational dimension of the Single European Sky (SES). SESAR is trying to make flying more environmentally friendly from the air traffic management point of view. Both programmes look for steep changes in air transport with major improvements but also with a relatively large time horizon.

Minor improvements can be reached already nowadays by implementing new approach procedures that can be flown by many of today's aircraft.

This paper describes the design of new approach procedures for Frankfurt airport which were implemented at the research airport Brunswick and flight tested with a Boeing 737-700. The approach procedures consisted of steep approaches [6] with approach angles from 4.5° over 5° to 5.5° as well as of marginal steeper approaches with 3.2° approach angle instead of the widely use 3.0° as well as area navigation (RNAV) procedures and required navigation performance (RNP) procedures.

In order to fly the different approach angles under precision approach conditions the experimental ground based augmentation system (GBAS), which is in operation at the research airport Brunswick since 2009,

delivered the necessary navigation performance.

To guarantee the highest precision all the approaches where flown in 0.10-nmi RNP mode.

2 Procedure Design

The RNAV procedures that have been validated in the flight trials have originally been developed by the "forum flughafen und region" in cooperation with the German Air Navigation Service Provider Deutsche Flugsicherung GmbH (DFS) for the use at Frankfurt airport [8]. The procedures, which can be seen in Fig. 1 (green lines), are designed to avoid the densely populated area of Offenbach, which lies under the extended centrelines (blue line) of the two main runways of Frankfurt (25L and 25R) at a distance of about 14km or 8NM from the thresholds. The newly designed procedures will lead the aircraft around Offenbach in the south and onto the extended centreline at the waypoint OBERA. Here, the aircraft intercept the ILS-Approach at an altitude of 2000ft, about 1650ft above the thresholds. As these procedures should be evaluated at Braunschweig airport, they were transferred to Braunschweig as can be seen in Fig. 2 [1][2][3][4][5]. The distance and bearing from the threshold of Braunschweig's runway 26 to the different waypoints are exactly the same as to the waypoints from Frankfurt's runway 25L. Hence the transferred procedure looks exactly the same as the original one, only tilted to fit the different runway orientation at Braunschweig airport.

Fig. 2 also shows a third procedure in the south leading from waypoint VE907 via VE906 and ENTSD to the runway. This procedure has not been designed by DFS but was added by DLR to investigate the feasibility of RNP-approaches, where the aircraft is lead onto the extended centerline in a fixed-radius turn. When the aircraft intercepts the extended centerline, it has a height of just 1000ft above the threshold, while already descending on a constant flight path angle of 3°.

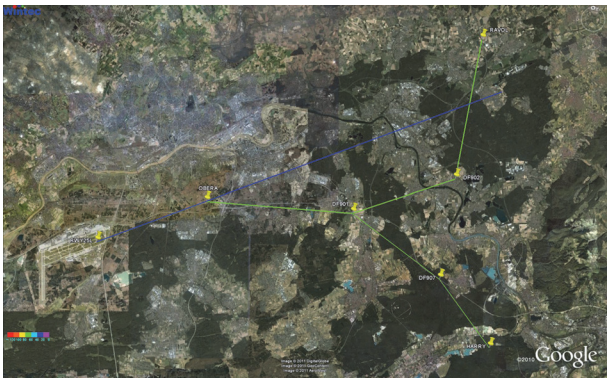


Fig. 1 Straight in (blue line) and RNAV (green lines) procedure for runway 25 in Frankfurt (EDDF)

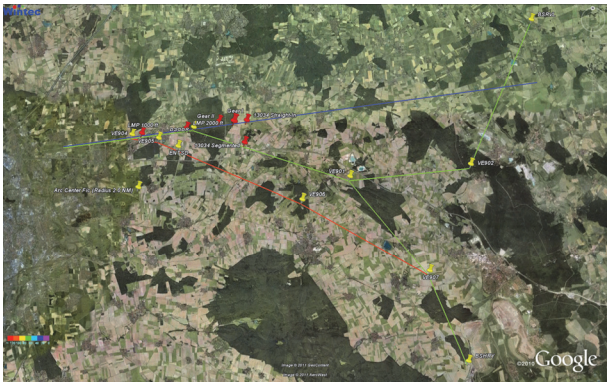


Fig. 2 Straight in (blue line), RNAV (green lines) and RNP (red line) procedure for runway 26 in Brunswick (EDVE)

The three procedures can also be seen in detail in Fig. 3 and 4. The abovementioned third procedure leads from VE906 to the waypoint ENTSD, which is the Final Approach Fix (FAF), but does not lie on the extended centerline. At ENTSD the aircraft will intercept at an altitude of 2000ft a glide slope of 3° leading constantly to the threshold of runway 26. Behind ENTSD, the aircraft - now in a constant descent of 3° - enters at VE905 a turn with a fixed radius of 2.0NM which ends at VE904 exactly on the extended centerline. At VE904 the height above threshold on the 3° glide slope is precisely 1000ft. From there, the aircraft continues on the centerline and on the constant glide path to the runway, as on any ordinary ILS-approach.

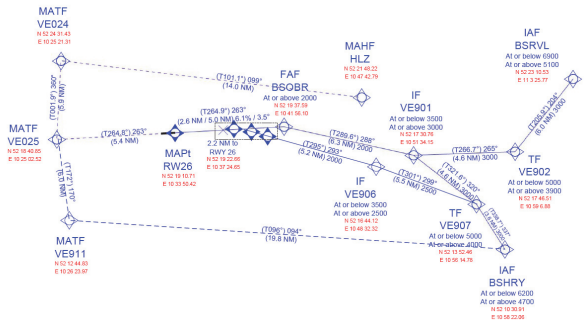


Fig. 3 RNAV and RNP procedures for Brunswick (EDVE) in detail

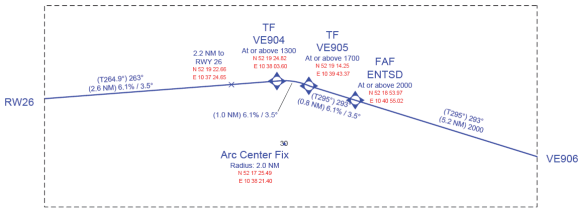


Fig. 4 Fixed radius turn of RNP procedure for Brunswick (EDVE) in detail

Besides the RNAV and RNP procedures also slightly steeper approaches with 3.2 degree approach angle on a 15 nm long straight final were tested that were also developed by the forum flughafen und region in cooperation with the German Air Navigation Service Provider Deutsche Flugsicherung GmbH (DFS) for the use at Frankfurt airport [7]. On top DLR tested steppt approaches at 4.5, 5.0 and 5.5 degree on straight 15 nm long final

3 Test

For the flight test and the noise measurements a Boeing 737-700 from Air Berlin was chartered performing 13 different approaches to the research airport in Brunswick. The aircraft was flown by Captain Marc Altenscheidt (chief pilot of the 737 fleet) and Captain Tim Techt (training captain of the 737 fleet). The flight trials were conducted in a series of two legs.

Table 1 Test matrix 1st leg

	Approach Transition	Final Approach	Remarks
BWE – BWE	RNAV Approach BSRVL	GLS 4.0° GPA	Low Approach
BWE – BWE	RNP Approach BSHRY	ILS 3.5° GPA	Low Approach
BWE – BWE	RNAV Approach BSHRY	GLS 4.0° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 3.0° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 3.2° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 4.5° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 5.0° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 5.5° GPA	Low Approach

Table 2 Test matrix 2nd leg

	Approach Transition	Final Approach	Remarks
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 3.0° GPA	Low Approach
BWE – BWE	Traffic Pattern 26, 15NM Final	GLS 3.2° GPA	Low Approach
BWE – BWE	RNAV Approach BSRVL	GLS 3.0° GPA	Low Approach
BWE – BWE	RNAV Approach BSHRY	GLS 3.0° GPA	Low Approach
BWE – BWE	RNP Approach BSHRY	ILS 3.5° GPA	Low Approach

The approach area was equipped with a set of 6 noise measurement microphones. The position of the microphones can be seen in Fig. 2. The two positions called “13034 straight in” and “13034 segmented” are located in a distance of 13034 meters to the runway threshold and correspond to a noise measurement point which also exists at Frankfurt airport and is located in the city center of Offenbach. The two points called “Fahrwerk I” and “Fahrwerk II” are located in an area where the gear should be down. The measurement point “BSOBR” is located at a point where the flaps setting should be finalized. At the measurement point

"1000 ft" the aircraft should be established on final approach especially after following the RNP route.

4 Results

4.1 Precision

The precision with which the aircraft follows the predefined flight track is a crucial item in the Required Navigation Performance concept (RNP). As the name implies, RNP requires the aircraft to show navigational performance, i.e. precision, within a certain value, e.g. 0.3 nautical miles (NM), which is then called RNP0.3. Accordant to the RNP concept, an ANP of for instance 0.05 NM means it is assured the aircraft is within a radius of 0.05 NM around the indicated position 95% of flight time. The on-board navigation systems of the aircraft constantly monitor the Actual Navigation Performance (ANP). Whenever the ANP is above the RNP, in this example worse than 0.3 NM, the procedure for which the certain RNP is required has to be aborted. The ANP itself is continually calculated on-board by the navigational systems depending on data availability and general assumptions about drift rates as well as data integrity under different circumstances. The mentioned assumptions are based on experience obtained during the certification process of a certain system used for navigation in the aircraft or general rules and formulae outlined in the certification guidelines.

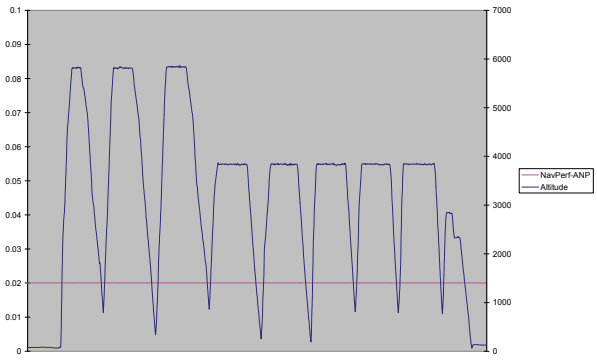


Fig. 5 Actual Navigation Performance (ANP) and altitude during 1st leg

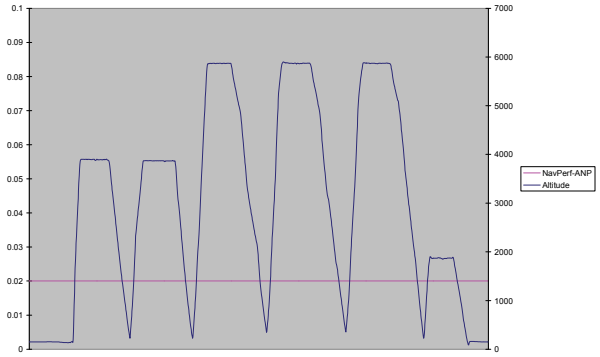


Fig. 6 Actual Navigation Performance (ANP) and altitude during 2nd leg

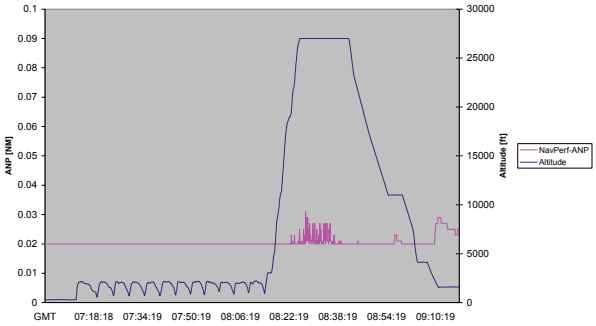


Fig. 7 Actual Navigation Performance (ANP) and altitude during approach phase and cruise flight

Fig.s 6 to 8 show the ANP during the flight evaluations in Braunschweig. Fig. 6 shows ANP during the first leg, containing several approach procedures. Fig. 7 shows the ANP during the second leg, also containing several approach procedures. The simple result derived from these Fig.s is that the ANP always remained at the value of 0.02 NM. GPS and an inertial navigation system were used to achieve this value. It has been stated by the flight crew involved in the flight evaluations that from their experience, the ANP is almost always 0.02 NM near the ground and increases only during cruise flight at higher altitudes. As evidence for this, Fig. 3 shows ANP and altitude during a flight that first contained several short approaches, i.e. stayed near the ground, thereafter one cruise flight at a typical cruise altitude of 27000 ft. Here, ANP is only higher than 0.02 NM when the aircraft is above 10000 ft and during taxiing after completion of the flight. Nevertheless, ANP never exceeded 0.031 NM during this flight and thus always remained below the RNP

of 0.1 NM required for approaches by a fair margin.

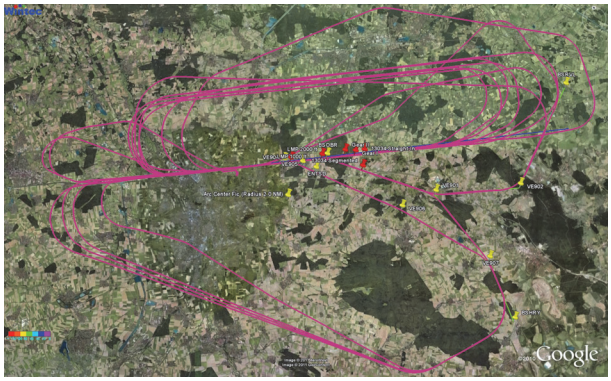


Fig. 8 Position of the microphones (red pinboard needles) and GPS Track of the test flights

4.2 Noise

The outcome of the noise measurements has to be viewed with respect to the fact that the measurements were single events. As the conditions always vary slightly during such trials, especially the wind, one would need a much higher number of flights to get a good statistic. At least the values can show a tendency.

Table 3 and 4 show all measured values. As indicator for the noise the *A*-frequency and *S*-time weighted maximum sound level LAS_{max} was measured.

The RNAV and RNP procedure reduce the noise in the sensitive area of Offenbach by 6 to 8 decibel (A) compared to a straight in approach which is a reduction of 50%. The turn on final approach at the end of the RNAV and RNP procedure did not show any increase in noise compared to a straight in approach. There are minor variations between the different approaches due to different approach speeds.

The slightly steeper approaches at an angle of 3.2 degree are a little bit less noisy than the 3.0 degree approaches. A reduction of 0.4 decibel (A) was measured during the first leg. During the second lag an increase in noise was measured on the 3.2 degree approach. This increase results from a higher approach speed (50 km/h) due to a slightly stronger tailwind.

Much better reductions of the noise level of course were achieved with the much steeper approaches at angles of 4.5, 5.0 and 5.5 degree. At an angle of 4.5 degree a reduction of 2.1 decibel (A) was measured compared to a 3.0 degree approach. At an angle of 5.0 degree the reduction was 3.6 decibel (A) and at 5.5 degree angle it was as high as 4.6 decibel (A).

5 Conclusion

The noise measurement results for the segmented RNAV and RNP approaches obviously showed a high reduction of the noise level for the areas lying in the east of Frankfurt airport like Offenbach and Mainz. Therefore these routes are in operation by Frankfurt airport since 10th February 2011 in the time between 11.00 pm and 5.00 am.

At approach angles of 3.2 degree the tailwind might become a problem sometimes. To reduce the speed the aircraft has to put the flaps earlier which might lead to a higher noise level than sticking to the 3.0 degree approach.

Much steeper approaches of course show a much higher reduction in noise level but actually can be flown down to landing only by a small number of aircraft, primarily turboprops. Therefore DLR is going to investigate segmented steep approaches in the near future. This means the approach will start at high approach angles and during the approach the approach angle will be reduced to 3.0 degree. This would reduce the noise level at least in the region with the steep approach angle.

Acknowledgments

Special thanks go to the flight crew of Airberlin and the team of people of Fraport and DLR that operated the noise measurement stations as well as to the Deutsche Flugsicherung DFS.

Table 3 Noise measurements 1st leg

	LASmax [dB(A)]	Altitude [m] SL	Speed [km/h]	Flaps [°]	Gear
RNAV RVL GBAS 4°					
13034 "segmented"	65,8	798	333	5	up
Fahrwerk I	no data				
Fahrwerk II	59,1	733	327	5	up
BSOBR	71,4	637	310	15	down
1000 ft	71,3	479	269	15	down
RNP HRY ILS 3.5°					
BSOBR	58,8	632	302	15	down
1000 ft	72,5	443	279	15	down
RNAV HRY GBAS 4°					
13034 "segmented"	67,1	805	317	10	up
Fahrwerk I	57	783	315	10	up
Fahrwerk II	63,3	741	312	10	up
BSOBR	69,8	678	299	15	down
1000 ft	73,7	482	253	15	down
3.0° Fraport					
13034 "straight in"	63	789	327	5	up
LMP Fahrwerk I	65,4	744	330	5	up
LMP Fahrwerk II	68,1	689	326	15	down
BSOBR	72,1	583	307	15	down
1000 ft	73,3	387	260	15	down
3.2° Fraport					
13034 "straight in"	62,5	846	331	5	up
Fahrwerk I	64,1	794	334	5	up
Fahrwerk II	68	733	337	5	up
BSOBR	71,6	626	312	15	down
1000 ft	73,8	417	266	15	down
GBAS 4.5°					
13034 "straight in"	64,8	1124	264	40	down
Fahrwerk I	65,4	1049	263	40	down
Fahrwerk II	64,8	972	260	40	down
BSOBR	65,5	839	259	40	down
1000 ft	70,7	536	250	40	down
GBAS 5.0°					
13034 "straight in"	63,1	1225	276	30	down
Fahrwerk I	62,2	1194	268	40	down
Fahrwerk II	65,3	1118	264	40	down
BSOBR	65,1	923	273	40	down
1000 ft	68,3	580	261	40	down
GBAS 5.5°					
13034 "straight in"	62,3	1234	283	30	down
Fahrwerk I	62	1232	268	40	down
Fahrwerk II	63,4	1180	259	40	down
BSOBR	64	1012	254	40	down
1000 ft	67,4	638	267	40	down

Table 4 Noise measurements 2nd leg

	LASmax [dB(A)]	Altitude [m] SL	Speed [km/h]	Flaps [°]	Gear
3.0° Fraport					
13034 "straight in"	62,5	805	336	5	up
Fahrwerk I	64,5	758	338	5	up
Fahrwerk II	68,9	670	332	15	down
BSOBR	70,9	599	314	15	down
1000 ft	74,2	400	276	15	down
3.2° Fraport					
13034 "straight in"	66,9	852	388	5	up
Fahrwerk I	68,3	804	388	5	up
Fahrwerk II	69,6	736	389	5	up
BSOBR	73,8	633	369	15	down
1000 ft	75,2	420	316	15	down
RNAV RVL GBAS 3°					
13034 "segmented"	67,1	791	326	5	up
Fahrwerk I	57,6	767	324	5	up
Fahrwerk II	61,2	729	315	5	up
BSOBR	70,3	660	292	15	down
1000 ft	74,1	385	286	15	down
RNAV HRY GBAS 3°					
13034 "segmented"	69,4	747	336	15	down
Fahrwerk I	58,3	715	326	15	down
Fahrwerk II	60,7	648	318	15	down
BSOBR	70,1	565	302	30	down
1000 ft	73,6	391	277	30	down
RNP HRY ILS 3.5°					
BSOBR	57,6	638	327	30	down
1000 ft	72,2	432	294	30	down

References

[1] ICAO, Doc 9613 – Performance – Based Navigation (PBN) Manual, International Civil Aviation Organisation, 3ed., 2008

[2] ICAO, Doc 8168 – Procedures for Air Navigation Services – Aircraft Operations – Volume I – Flight Procedures, International Civil Aviation Organisation, 5 ed., 2006.

[3] ICAO, Doc 8168 – Procedures for Air Navigation Services – Aircraft Operations – Volume II – Construction of Visual and Instrument Flight Procedures, International Civil Aviation Organisation, 5 ed., 2006.

[4] EUROCONTROL, RNAV Approaches European Organisation for the Safety of Air Navigation, http://www.ecacnav.com/rnav, 2010.

[5] ICAO, Doc 9905 – Required Navigation Performance Authorization Required (RNPAR) Procedure Design Manual, International Civil Aviation Organisation, 1ed., 2009.

[6] Geister R., "Segmented steep precision approaches based on GLS", CEAS Aeronautical Journal, http://dx.doi.org/10.1007/s13272-011-0016-6, 2011, pp. 1 – 8

[7] Forum Flughafen und Region, http://www.forum-flughafen-region.de/uploads/media/Hintergrund_erhoehter_Anfluggleitwinkel_3_2_Grad_Layout.pdf, 2010

[8] Forum Flughafen und Region, http://www.forum-flughafen-region.de/uploads/media/Hintergrund_segmented_Layout.pdf, 2010

Real Time ATC simulation of Time Based CDA operations

R. de Muynck and T. Bos
National Aerospace Laboratory NLR, Netherlands

A. Kuenz and C. Edinger
German Aerospace Center DLR, Germany

L. Rappich and S. Törner
LFV Swedish Air Navigation Services, Sweden

Keywords: *Continuous Descent Approach, Time Based Operation, Arrival Management, Controlled Time of Arrival*

Abstract

The expected future growth in aviation requires more efficient and environmentally sustained operations. The present paper presents the results of a concept evaluation to increase the rate of Continuous Descent Approaches. This work was carried out within the frame of the EU sponsored 6FW project ERAT (Environmentally Responsible Air Transport). The project facilitates in enhanced climb and descent services as defined by SESAR and aims for operational improvements for implementation in 2015. The concept of operation attempts to achieve medium to high density CDA operations applying time based operations. It is based on an advanced arrival management (AMAN) system that supports Controlled Time of Arrival (CTA) operations.

1 Introduction

Continuous efforts are made worldwide to achieve a sustainable growth of the air transport industry with an acceptable burden to society. As such, it is necessary that improvements are made, both by introducing more economic and silent aircraft, but also by a more efficient and environmentally friendly operation of the entire air traffic system.

To reduce the environmental burden to communities around airports and to improve the economic and environmental efficiency during the initial and final phases of flight, many countries have already invested in researching and introducing improved departure and arrivals operations.

At present, the frequent use of CDAs, flown from top-of-descent to final approach, is not yet feasible during daytime operations at major airports. Such Continuous Descent Approaches are currently too unpredictable for existing air traffic management systems, which means it is difficult for air traffic controllers to merge them within the tightly orchestrated arrival flow around busy airports. On the other hand, with respect to departing traffic, there is also the goal to provide for continuous climbing departure profiles, as any restriction made to departing traffic due to the facilitation of continuous descents may easily offset any benefit gained on the arrival side.

Within the EU sponsored FP6 project ERAT (Environmentally Responsible Air Transport), the project partners teamed up to prepare and validate the operational context for enhanced climb and descent services as defined in SESAR for two major airports. As part of this project,

Sweden's Air Navigation Service provider LFV, with the aid of DLR, NLR and the other project partners, have developed and validated an operational concept, controller tools and procedures for the implementation of Continuous Descent Approach operations under higher traffic conditions.

Based on previous experiences for operating timed (CDA) approaches in other projects, such as CASSIS [1], TMA2010+ [2,3,4], SARA, OPTIMAL [5], and Sourdine II [6], work was continued to further develop and validate a suitable operational concept for time based continuous descent approaches under medium-high traffic conditions at Stockholm Arlanda airport. The main elements of the concept are described in the following section.

Furthermore, prototypes of essential concept elements, such as a suitable integrated controller working environment and air traffic management tools capable for time based arrivals needed to be developed.

The concept, system support and associated operational procedures were evaluated under realistic simulated air traffic conditions during real-time ATC simulations of Stockholm Arlanda airport, as described in section 3. These simulations and results are described in the final sections.

2 ERAT Operational Concept for Arlanda

The concept of time based CDA operations, applied to the Stockholm Arlanda airport and airspace environment, applies the assignment of Controlled Times of Arrival (CTA) to arriving aircraft in order to achieve accurate sequencing for the final approach. The CTA control point is a dedicated waypoint on the arrival route for the inbound arriving aircraft.

Earlier projects already gained considerable experience with the application of different CTA waypoint locations, e.g. TMA entry, around 20-40 NM or on the runway. Based on the obtained feedback, it was regarded most optimal to use a CTA distance to the runway of 30NM with the present state of technology.

An overview of the operational concept for a typical flight is indicated in fig. 1. The CTA time constraint to be achieved by the inbound aircraft is determined by the ground based arrival management system (AMAN), which determines the landing sequence. For such a nominal arrival the sequence of events is the following.

Once within planning range, arriving aircraft for Arlanda airport are taken into account in the AMAN landing sequence planning. The aircraft are issued a standard arrival route (STAR) to the landing runway in use. Approximately 30-35 minutes prior to estimated runway touchdown, the AMAN system finalises the planned landing sequence. The scope of this concept applied to Stockholm Arlanda has been visualised on the map in fig. 2.

Based on the frozen runway arrival time in the landing sequence planning, the estimated time over the 30NM control waypoint on the assigned arrival route can be provided as a 4D time constraint for the arriving aircraft.

The en-route (ACC) ATC controller is notified and may then issue this time constraint as a CTA clearance to the aircraft. The flight crew verifies and reprograms the active flight plan in the onboard Flight Management System (FMS) with the assigned CTA time.

Once confirmed by the flight crew, the aircraft now will try to achieve passing the CTA waypoint with ±30 sec accuracy using guidance provided by the 4D capable FMS. In addition, the ACC controller will mark the CTA controlled flight in the system, indicated by a special 🕒 symbol with the aircraft flight label on the radar screen.

Taking into account other en-route traffic, the ACC controller provides progressive descent clearances to allow the continued execution of the timed based idle descent. Once approaching the TMA boundary, the descending aircraft is then transferred to the next approach (APP) controller. Further on, once passing the 30NM control waypoint with sufficient accuracy, the aircraft will continue on the idle optimum descent with further time constraints monitored by the approach controller and/or final director.

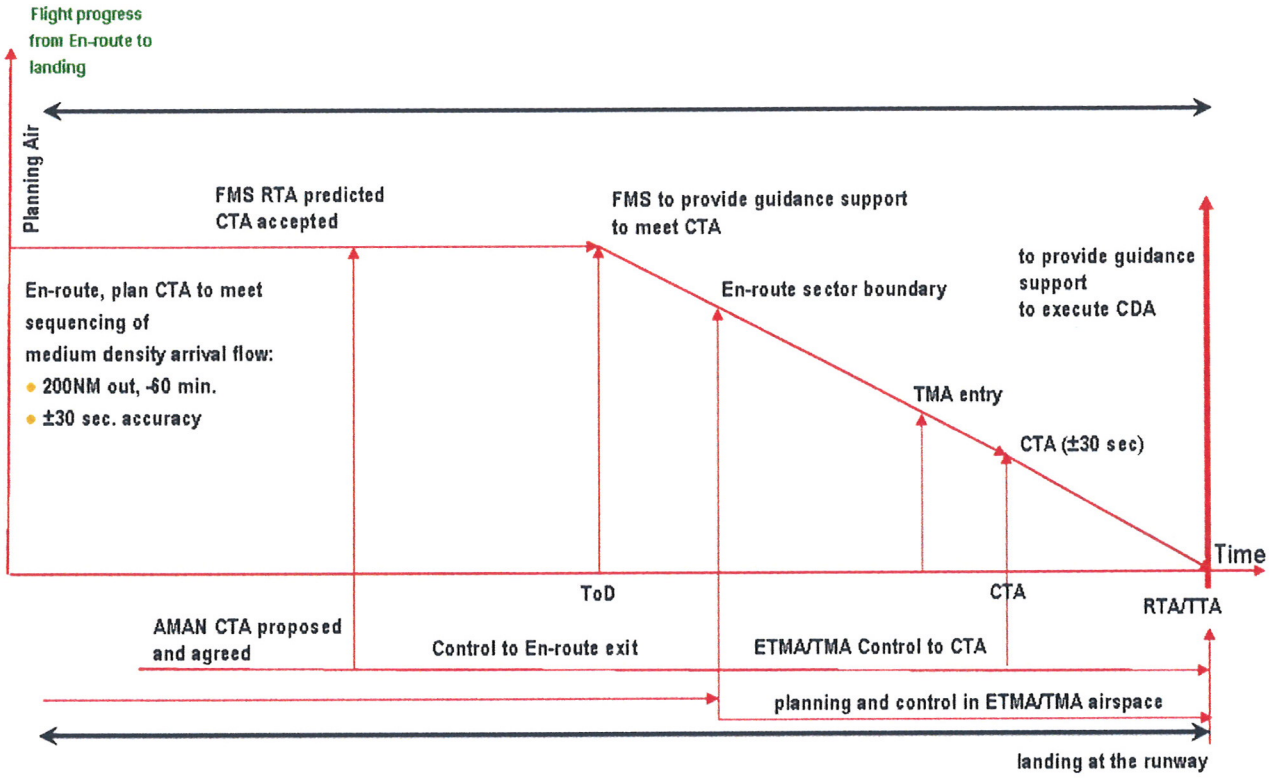


Fig. 1 Overview of operational concept

It should be noted that any controller may issue speed, heading or altitude instructions at all times to achieve a safe operation of flight.

In addition to the nominal flights operating on a CTA clearance, the operational concept and support tools also considered handling of other traffic not operating on a time constraint on their own for various reasons. In addition, the concept and system should also be robust to disturbances in the traffic flow.

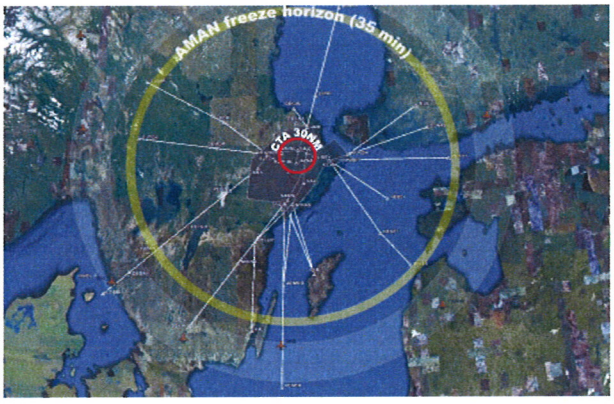


Fig. 2 Scope of arrival management at Arlanda airport

2.1 Non-CTA based arrivals

In reality there will always be a percentage of the inbound traffic that will not be able to comply with a CTA clearance, e.q. due to onboard equipment limitations or due to aircraft performance (too early/late to achieve). In such cases, the ACC controller will continue to give conventional instructions to achieve the target time over the CTA waypoint. However, to support the controller in achieving the planned time over the CTA point with increased accuracy, additional system support is provided by means of time to loose or gain for that flight.

2.2 Short flights

Ideally, the landing sequence planning will be scheduled with all inbound flights becoming active in the AMAN planning system at cruise altitude. However, in reality special handling will be necessary for short flights with flight times at or around the AMAN planning horizon.

The en-route and arrival times of such flights can only be determined with sufficient accuracy after becoming airborne. This uncertainty will also affect final sequence planning of those arrivals just before or after these flights.

Within the developed concept, short flights are taken into the planning before take-off with a provisional arrival time slot, which is used to

coordinate a target take-off time before departure.

3 Concept evaluation

To evaluate the feasibility and performance of the concept on the various operational aspects a series of real time simulations were prepared. Following extensive preparations and testing, ATC simulator evaluations were carried out during two weeks in Malmö and Amsterdam in 2010 with the participation of active controllers from Stockholm ATC. For the ATC evaluations, the NARSIM ATC simulation platform was used configured with en-route and terminal airspace around Stockholm Arlanda airport. The DLR developed TrafficSim was used to generate air traffic able to accurately fly time-based CDA approaches.

3.1 Airspace and routes

The scope of the arrival management process was already indicated in fig. 2. In reality this scope surpasses multiple ATC en-route sectors, in particular to the East with the neighboring Finnish airspace. For the assignment of arrival time constraints this will normally require coordination. However for evaluation of the concept this was not regarded essential and not taken into account. The simulated en-route control sectors were therefore enlarged to encompass the entire CTA allocation process.

A single landing runway configuration was simulated at Stockholm Arlanda, resulting in using runway 19L for arriving traffic and 19R for departures. In addition, for traffic in and out of nearby Bromma airport, the runway in use for both take-off and landing was runway 12. The standard instrument departure (SID) routes were used as officially published, although with a modification to the initial SID altitude.

Draft P-RNAV and RNP based standard arrival routes to runway 19L were designed for the present concept evaluation, since at present no published RNAV arrival routes exist for this runway. In addition, the special CTA control waypoints at 30NM were added to the procedure. Fig. 3 provides an overview of the final version of the arrival routes. Note that either the P-RNAV or RNP based arrival routes were used, not simultaneously.

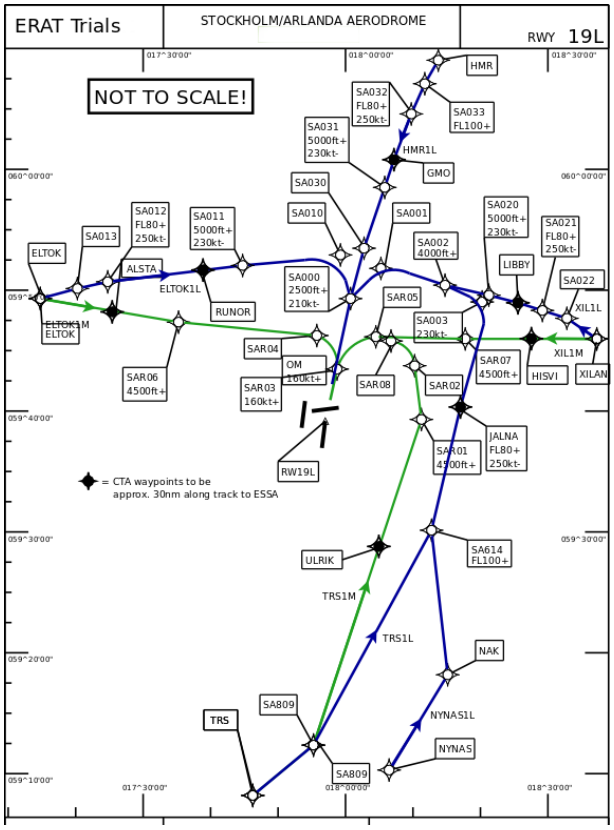


Fig. 3 Overview of P-RNAV and RNP arrival procedures in the Stockholm Arlanda terminal area

No alternations were made to the route structure outside the Stockholm terminal area for accommodating the RNAV STARs.

3.2 ATC working positions

The Controller Working Positions for area and approach control resembled those currently used at Arlanda, however with specific adaptations for the concept evaluations. Six controller working positions were prepared for the following roles:

- APP West executive controller
- APP East executive controller
- APP final director
- Approach coordinator
- ACC executive controller sector 3 (W)
- ACC executive controller sector 1 (E)

A number of actual en-route airspace sectors were combined to reduce the number of controllers required for the simulation. While this could result in a higher workload than usual for these positions, this was compensated by using appropriate traffic samples for these sectors.

Rotation of the participating controllers was applied, when applicable, to collect opinions of the concept from different participants' view points. It was expected that familiarisation with the traffic samples would also be minimised as controllers were rotated.

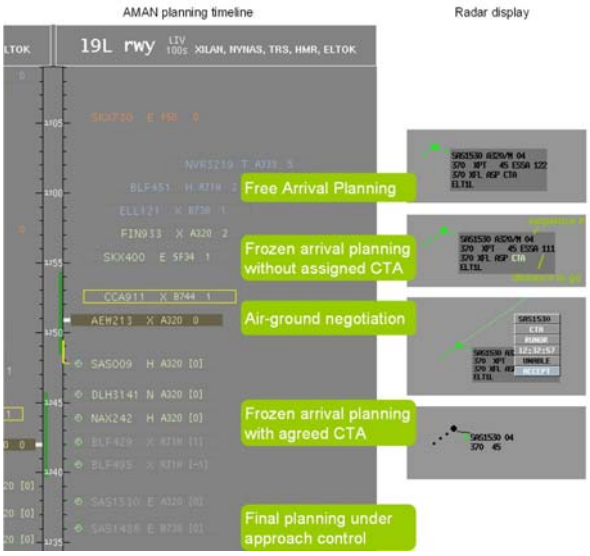


Fig. 4 Integrated Controller Working Position

In addition to the "normal" radar screens showing traffic for the relevant control sectors, an Arrival Sequence window was also available to the controllers. This presented the planned arrival sequence by means of a "time-line" with aircraft labels. In addition, this sequence and time information was also available on the radar screens in the aircraft label or on mouse-over as shown in fig. 4.

3.3 Experimental factors

To make a validated comparison between conventional and time based operations, the following conditions were compared:

- A baseline, reference condition reflecting present day operations
- Time based CDA concept with P-RNAV based arrival routes
- Time based CDA concept with shorter RNP based arrival routes

The time based CDA operations with CTA point at 30NM were supported by enhanced AMAN, and conventional speed control and vectoring as necessary.

The following factors were randomised over the different validation runs

- Traffic sample
- Wind condition
- Controller seating

The simulated traffic scenarios consisted of both inbound and outbound traffic for Arlanda including short flights, flights to/from nearby Bromma airport as well as overflying en-route traffic. The amount of traffic into Arlanda was set for a medium-high landing rate.

4 Results

During the two weeks of real time ATC simulations in total 27 measured evaluation runs were made. The first week of evaluation run (RTS1) was carried out at LFV Malmö in June 2010. A second week of extensive simulations (RTS2) was completed in September 2010 at NLR in Amsterdam. Feedback and comments received during the RTS1 sessions were incorporated in the overall simulation setup, support tools and traffic samples. In addition, the second series also included the evaluation of RNP arrival routes, which provided a further reduction in track miles compared to the P-RNAV based STARs which are more comparable to present day vectored tracks.

During and after each run subjective assessments, data logging and observations were recorded. Following each run direct feedback and comments were given by the participating controllers and pilots.

4.1 CDA arrival profiles flown

A direct indication of the feasibility of the developed concept is the number of successfully completed CDA approaches during the different experimental conditions.

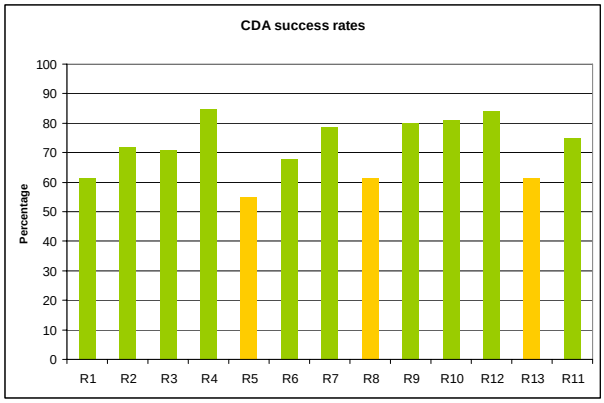


Fig. 5 comparing CDA success rate between baseline and time based concept

Fig. 5 provides an overview of successfully completed CDA's from top-of-descent to landing observed during the first week of evaluation sessions (RTS1). Clearly visible is the fact that the time based CDA concept (green) shows a marked improvement compared to the baseline situation, where already good score is achieved by the controllers using conventional operations (yellow).

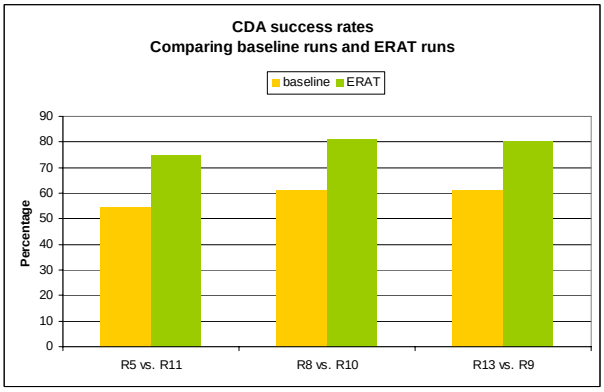


Fig. 6 comparing CDA success rate between baseline and ERAT time based concept (comparable runs)

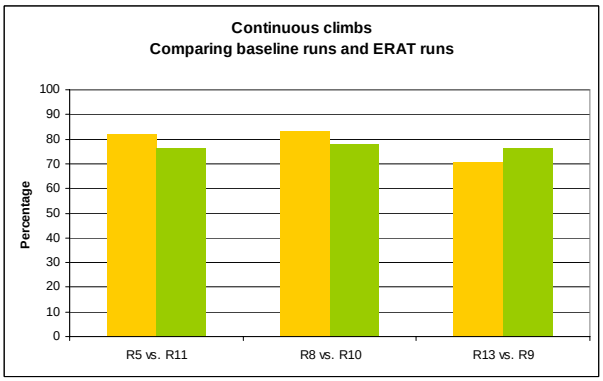


Fig. 7 Amount of continuous climb departures

Focusing within the overall results on true comparable runs with identical initial conditions in traffic and weather, the figures 6 and 7 show a clear increase in the amount of CDA's without significantly impacting the efficiency of aircraft departures.

4.2 Workload and Situation Awareness

While the definite increase in the amount of completed CDA profiles could be noted, the delegation of 4D trajectory control to arriving aircraft to achieve the target time did not appear to negatively affect the workload of the controllers. During the simulations, through a pop-up menu on a separate touch screen, controllers were asked every few minutes to give a self assessment of their current workload. Averaged results are indicated in figure 8 for the approach controllers. The comparison does not show significant differences with the baseline situation. Similar results were also obtained from the en-route controllers.

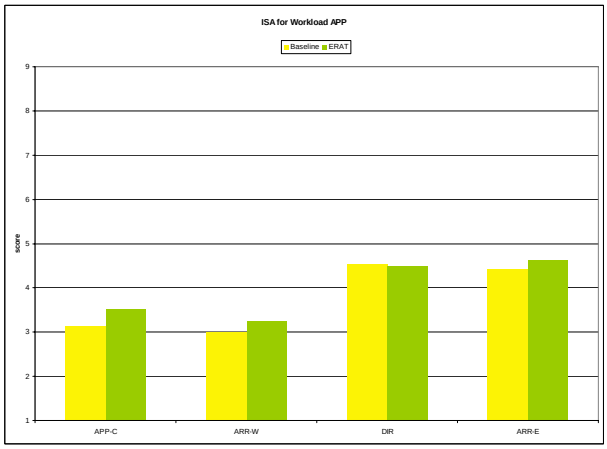


Fig. 8 Instantaneous Self Assessment of Workload experienced by APP Controllers

During the post-run questionnaires the ATC controllers were requested to provide feedback on the perceived situation awareness (SA), ref. [9]. The resulting SA scores are shown in figure 9, which do not show significant differences compared to the baseline situation (yellow).

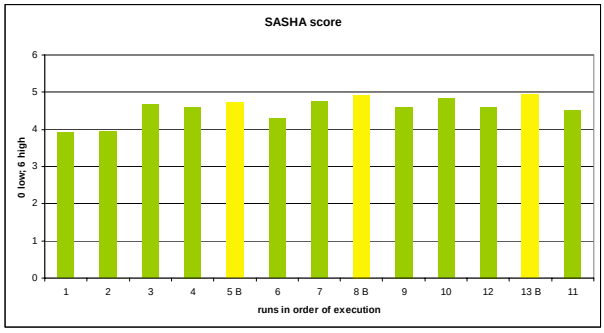


Fig. 9 Average SASHA score for each run (RTS1)

Nevertheless, the en-route controllers sometimes noted difficulty to maintain situation awareness of flights that received a CTA time clearance. It was sometimes difficult to estimate the applied FMS speed strategy, in particular in relation to surrounding other traffic.

4.3 Handling of short flights

Apart from evaluating the feasibility of the time based CDA concept, special interest went to developing procedures for better handling of short flights, with flight times of 25-40 minutes, which occur on a regular basis at Arlanda. The short flights are often operated by turboprop aircraft, which also adds an additional challenge in handling within the more homogeneous arrival flow of jet traffic.

Since take-offs will not be performed with the same accuracy as is achieved for CTA time constraints when already in flight, planning the final landing sequence around such a short flight required special coordination. The flights were sequenced in advance using preliminary arrival slots and target take-off times.

Although sequencing uncertainty remained until initial climb, the flexibility of the operation proved to be sufficient to deal with short flights and turboprop behaviour.

4.4 RNP based arrivals

Given the promising results obtained during RTS1 when operating on P-RNAV STARs that proved to be well compatible with vectoring, a step ahead was made to evaluate shorter RNP based STARs during the second evaluation in RTS2. While the P-RNAV STARs mixed very well with vectoring when necessary, these also required relatively long 10NM final approach. Use of RNP based STARs with a shorter 5NM final approach should bring additional efficiency due to reduced overall track miles to fly.

Results with the (limited) amount of runs with RNP approaches showed that the number of successful CDAs was lower compared to the more standard P-RNAV based CDA routes using the ERAT concept. Analysis showed that aircraft were vectored more often, in particular at lower altitudes. In hindsight, the amount of dedicated training of the RNP runs should have been more and a definite learning curve was visible during the trials. It is expected that better results are possible with further RNP STAR and SID route refinements as well as more operational experience of both pilots and controllers. Nevertheless, based on simulation results it was expected that a landing throughput of up to 30 per hour should be feasible when using RNP STARS, slightly lower compared to the regular P-RNAV STARs.

5 Concluding remarks

Results clearly indicated that the developed time based CDA concept with the CTA point at 30NM is well feasible with the offered medium traffic density of around 32 landings per hour with higher temporary peaks. About 20% more CDA’s were flown using the P-RNAV arrivals with the time based concept compared to conventional operating procedures during the experiment baseline conditions.

In addition, the operation with RNP arrivals appears to yield additional efficiency in terms of fuel consumption. Given the limited amount of familiarisation, results are promising and the concept with shorter RNP routes is considered feasible in medium traffic density

The experiences gained with the real-time ATC simulations have considerably supported the validation of the time based CDA concept at Stockholm Arlanda and improved the concept as well as supporting technology. The knowledge and experience obtained will be

applicable to other airports to profit from time-based CDAs in daily practice.

Acknowledgements

The authors wish to thank the project team members who participated during preparation and execution of the real-time ATC simulation of Stockholm Arlanda airport. Also, much appreciation is given to the participating LFV controllers and simulator pilots who gave very valuable feedback and suggestions both during prototyping, testing and the trials.

References

[1] Klooster J.K, Del Amo, A., Manzi P., "Controlled Time-of-Arrival Flight Trials", *Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)*, Napa, 2009.

[2] Rundberg C., Mullan, C., "LFV/TMA2010+ Real-Time Simulation 1", LFV/Eurocontrol, 2007.

[3] Mullan C., "TMA2010+ Project North RTS2 Validation Report", LFV/Eurocontrol, 2009.

[4] Mullan C., Pickup S., Rundberg C., "TMA2010+ Project North RTS3 Exercise Report," Eurocontrol, 2009.

[5] Roux Y, et. al., "OPTIMAL final publishable report", Airbus, 2008

[6] Boer R. den, et. al., "Sourdine II Final Report", NLR, 2006

[7] Anon., "SASHA, Situation Awareness for SHAPE Questionnaire, shortened version", Eurocontrol, 2006

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Airport Operations Simulation using Particle Swarm Optimization

S. Bagassi, D. Francia, F. Persiani
University of Bologna, Italy

Keywords: *Airport Operation, Shortest Path Problem, Particle Swarm Optimization*

Abstract

The optimization of airport operations is recognized as a challenge that aims at finding the best trade-off solution in order to maximize the airport capacity and minimize both pollution and noise.

In a previous study we presented the processing, by Dijkstra algorithm, of airport data in order to find the shortest path between two assigned waypoints.

However this approach did not provide the best solution in terms of computational time, that become of major importance in a decision making process such as airport operation optimization.

In this paper a new approach to the optimum path search is undertaken through the meta-heuristic Particle Swarm Optimization (PSO).

PSO simulates the behaviors of bird flocking. Each single solution is a "bird" in the search space and it is called "particle". The best solution is reached by following the bird which is nearest to the objective.

1 General introduction

Along the whole mission of aircraft, airport operations strongly affect the pollution and noise produced, impacting on the environment and people living near airports. Airport area, and more generally the Terminal Aerodrome Zone, is the bottleneck of Air Traffic System.

In the last few years Air Traffic Management (ATM) system aim to manage congestion and capacity constraints at major European airports relating to sustainable development: this will lead to a greater emphasis on the environmental performance of the European ATM system. The overall capacity of the European ATM system is to a significant extent determined by the capacity of its airports and en-route sectors.

An increasingly significant constraint on European airport development and operation is the environmental issue, leading to the emergence of the concept of environmental capacity.

In this context a trade off solution that aims to maximize airport capacity and to minimize pollution and noise is a big issue. ATM is considered to organize data on airport surface operations, the management of all airport vehicle activities on or near runways, including the movement of aircraft, baggage vans, fuel trucks, catering vehicles, security personnel, and

any other ground traffic. Furthermore, the effectiveness of the airport surface traffic control is mainly reflected on the runway, taxiway and gate levels.

The complexity of the Air Traffic Management decision making processes could be mitigated by the adoption of optimal system-wide control strategies to manage airport congestion. In this paper we propose an approach to the optimization of the taxi process in the context of airport operation.

The taxi process can be defined as the surface movements of the aircraft between the runways and the parking positions and the optimization we propose aims to find the best possible way to go from a waypoint to another. Essentially the question is to find the shortest way to go from an initial waypoint, that corresponds to two spatial coordinates on a 2D map, to a final one, respecting some constraints and by choosing the best way in order to optimize an assigned cost function.

This is commonly known as the Shortest Path (SP) problem.

In a previous work we proposed the same issue and we offer the solution by means of the Dijkstra algorithm [1].

The airport operations model we proposed is mainly composed of:

1) data collection;

2) processing;

3) visual module.

In this paper we mainly focused on the processing module.

SP problem involves finding the shortest path between two pairs of vertices in a graph. A graph G=(V,E) comprises a set V of N vertices and a set E of edges connecting vertices in V.

A graph can be represented as an adjacency matrix *A* in which each element *(x,y)* represents the edge between element *x* and *y*. There is a cost $C_{x,y}$ associated to each couple of elements *(x,y)*. The total cost to go from a starting vertex to a final one, by visiting a certain number of vertices in the graph, has to be minimized while choosing the best sequence of visited vertices.

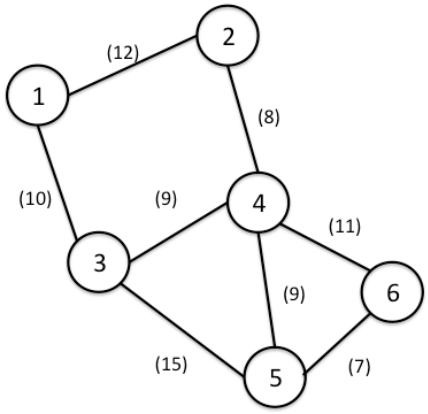


Fig. 1 An example of graph: 6 connected vertex and assigned cost functions, in brackets.

To solve this optimization problem modeling and simulation are the most powerful tools: different approaches have been proposed for the SP routing problem based on classical combinatorial optimization problem [2, 3].

SP problem has widespread applications in the field of communication networks, routing transportation systems and path planning and they usually find solutions by means of the Bellman's dynamic programming algorithm, the Bellman–Ford algorithm for networks [4] and the Dijkstra algorithm.

According to Dijkstra approach, as proposed in [1], all the main ground waypoints are collected in a list Wp = (1,2,..,N) . To each waypoint i∈Wp is associated its position, in terms of Latitude and Longitude. In a square matrix A relationships between the couple of way points that are connected and those that are not linked each other are displayed. In order to assign a weight to the path between the couple of waypoints i, j, distances between them are calculated and a square symmetric matrix C is created. The two matrices A and C, the starting point s∈Wp and the end point e∈Wp are the inputs for Dijkstra algorithm. It gives as outputs the shortest path SP = (s,....,e) and its cost E(SP).

As major shortcoming this algorithm search only for the shortest route, but it cannot determine any other similar/non-similar short routes and it exhibits high computational complexity for real-time communications.

To overcome the problem of computational efficiency many different approaches have been proposed, among them the most prominent seems to be neural networks (NNs) [5–7] and evolutionary algorithms, such as Genetic Algorithms (GA) and Tabu Search (TS) meta-heuristics [8–12].

However, these approaches involve a large number of iterations: in this paper we proposed a different approach to solve the SP problem by means of Particle Swarm Optimization (PSO).

2 PSO Approach

Particle Swarm Optimization is a stochastic optimization method based on the simulation of the social behavior of bird flocks or fish schools. It was originally developed by Kennedy and Eberhart [13] to efficiently find optimal or near-optimal solutions in large search spaces.

The PSO is a Swarm Intelligence method in which the population is not manipulated through operators inspired by evolutionary procedures as in the GA: in PSO, each individual, named particle, of the population, called swarm, adjusts its trajectory toward its own previous best position, and toward the previous best position attained by any member of its topological neighborhood. In the global variant of PSO, the whole swarm is considered as the neighborhood.

However, PSO shares many similarities with evolutionary computation techniques: the system is initialized with a population of random solutions and searches for optima by updating generations. In PSO, the potential solutions fly through the problem space by following the current optimum particles.

The PSO algorithm utilizes Swarm Intelligence in order to find the best place in the search space. The actual position of a particle is associated with a particular design vector x. The trajectory of the i-th particle at iteration n can be described with the position update equations.

The single generic particle i at time t hold a position X_{it} and a velocity V_{it} .

$$X_i^{t+1} = X_i^t + V_i^{t+1} \qquad (1)$$

$$V_i^{t+1} = w \cdot v_i^t + c_1 \cdot \varphi_1 \cdot \left(p_{ib}^t - X_i^t\right) + c_2 \cdot \varphi_2 \cdot \left(g_b^t - X_i^t\right) \qquad (2)$$

where:

w = inertia weight

c_1, c_2 = constant weight factors

φ_1, φ_2 = random factors in the [0,1] interval

Usually both c1 and c2 are taken to be 2.0, yielding good results [13].

The values *p-best* , p_{ib}, and *g-best*, g_b, are the ones each particle can record and it refers to the best position so far reached by the particle and to the best position that neighbors reached, respectively: p_{ib} is a local value whereas g_b is a global one. At every iteration, each particle is updated by following the described two best values.

Basic criteria to reach the objective during the optimum search are: the separation, that aims to avoid collision with neighbors; the cohesion, that aims to head the mean neighbors' position; the alignment, that aims to head the bird flock's position.

All of particles are characterized by fitness values, that are evaluated by the fitness function to be optimized, and by velocities which direct the flying of the particles.

A typical procedure for general algorithmic flow of PSO can be summarized as shown in Fig.2.

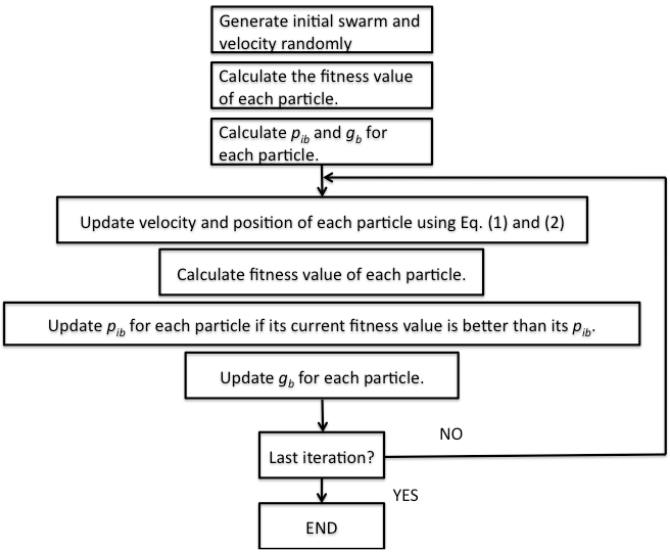


Fig. 2 Procedure of a standard PSO algorithm.

2.1 SPSO (Shortest-path PSO) algorithm

As previously noticed, algorithms usually adopted to solve the SP problem are those used for the classical combinatorial optimization problem. In this section we describe the new approach based on the meta-heuristic PSO by means of some adjustment to be applied to the SP problem.

The thorniest problem in applying PSO (and GA) to the SP problem is due to the encoding of a path in a network graph into a particle in PSO. This encoding affects the effectiveness of a solution/search process.

2.1.1 The path encoding

The procedure we followed is inspired to path encoding techniques for solving the SP problem using GA.

The encoding technique follows the indirect scheme for chromosome representation proposed in [11].

As first, a priority vector has to be generated that guides the choice of the nodes of the graph that have to be visited while defining the path.

The priority in this vector has to be assigned randomly at the first iteration.

The path is generated by choosing the highest priority for each available node connected to the source node and leading to the destination node.

Starting from the graph structure of the assigned network, the priority vector assumes the form shown in Fig. 3 where the cell position correspond to the node in the graph: into the cell is assigned the priority related to each node.

1	10	28	8	26	24

1	10	-N	8	26	24

1	10	-N	8	-N	24

Fig. 3 The priority vector for a 6-nodes graph.

This choice of the next-to-be node according to the priority in the x vector is repeated until the destination node is reached.

Referring to the graph illustrated in Fig.1, let's suppose to search for the optimal path to go from the node 1 to the node 6.

Once the node 1 is assigned as starting node, the choice of the second node is guided by the priority in the vector of Fig. 3. The highest priority is the value 28, so the node 3 is chosen to be the next-to-be in the path. The highest priority, no more considering 28, is 26, so the node 5 is chosen to be the next-to-be in the path. Finally the highest priority in the vector is 24 and the final node 6 is so reached.

Thus the complete path {1, 3, 5, 6} is obtained. Two warinesses are suggested in order to overcome situations when the path does not terminate at the destination node, thus leading to an invalid path [14].

The first consists into assigning a large negative priority value, -N as displayed in Fig. 3, to a node already included in the path so that the node is highly unlikely to be selected again.

The second consists in reducing the possibility of building a backward path (hence, possible loop formation) by means of the introduction of a heuristic operator incorporated as follows.

A node is selected as to be next node in the path if its ID is larger than the present node ID by a certain specified value:

$$ID_{\text{next node}} - ID_{\text{present node}} > -M \qquad (3)$$

where M is a positive integer.

Once the particle has been created by means of the encoding strategy, the PSO starts to run and it follows the procedure described in the scheme of Fig. 2.

A typical fitness function for a SP problem can be the one expressed in Eq. 4 as follows:

$$f_i = \left(\sum_{j=1}^{Ni-1} Cyz \right)^{-1} \qquad (4)$$

The fitness function will return a cost value for each particle of the swarm, the maximum value the fitness will find is relative to the shortest path in the research space.

3 Case Study

To test the methodology proposed, we referred to the real case study of Rome Fiumicino airport, here referred as LIRF.

The airport is described through its gates, taxiways and runways.

This airport can hold all a/c categories, it has 4 runways, 3 parallel each other and 1 perpendicular to the others, as shown in Fig.4.

The traffic in this airport reaches the maximum of 1157 ground movements a day. On the average, it is about 340.000 ground movements a year, 45% due to national air traffic and 55% to the international one.

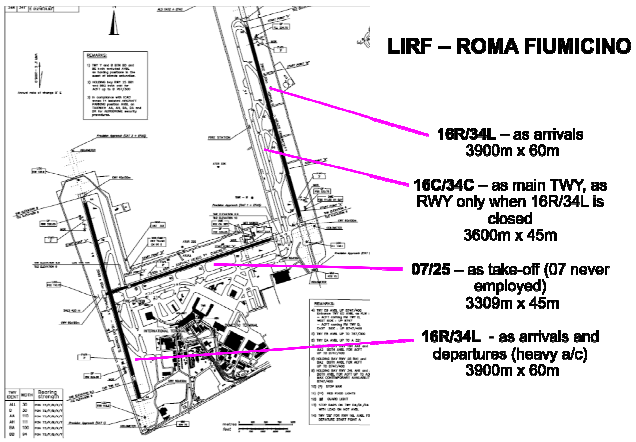


Fig. 4 The Rome Fiumicino airport: runways are highlighted.

In the airport space taxiways, that have to run from a starting point to a destination, are optimized.

The network of all the accessible points of the airport manoeuvring space is described by means of 45 waypoints, as listed in Table 1.

Each waypoint is characterized by two values on a Cartesian plane that correspond to the latitude and longitude of the point on a digital map, as shown in Fig. 5.

Point	*Stroke*	*Position*
1	Joint DA	Starting Joint
2		Middle Joint
3		Starting Runway
4	Joint DB	Starting Joint
5		Middle Joint
6		Starting Runway
7	Runway 16L	Touch Down Point
8	Joint DC	Starting Joint
9		Middle Joint
10		Starting Runway
11	Joint DF	Runway Point at DF
12		Joint End Crossing TWY D
13	Taxiway D	D1
14		D2
15		D3
16		D4
17		D5
18		D6
19	Waiting Point	BA2
20		BA1
21	Waiting Point	RBA1
22		RBA2
23	Runway 25	Take-off Point BA
24	Cross	B/N
25	Waiting Point	BB1
26		BB2
27	Runway 25	Take-off Point BB
28	Cross	B/P
29		NG/EB
30		EG/EA
31		EG/EB
32		EG/G
33		EG/H
34		H/NG
35		NG/G
36		NG/NF
37		NF/NC
38		NF/NH
39		NH/H
40		NC/H
41	Stand 303 IN	I303
42	Stand 303	S303
43	Stand 303/6 OUT	U303/306
44	Stand 306 IN	I306
45	Stand 306	S306

Table 1: the list of the significant waypoint on LIRF

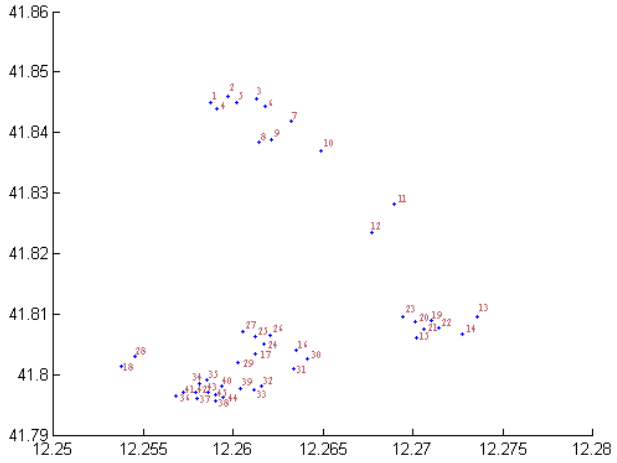


Fig. 5 The LIRF waypoints on a Cartesian plane

The first step is to generate the priority vector that will guide the choice of points at the first iteration of the encoding phase. This is a 45 cells' vector and it will be updated at each time a node will be chosen as next-to-be node.

In the application here proposed the priority vector has been defined taking into account the real accessibility between two nodes of the graph.

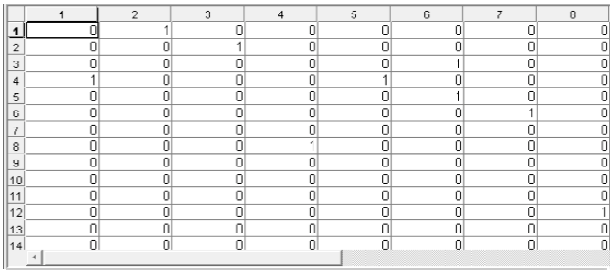
A high priority of the node depends on two factors: the first is the evaluation of the node connection to a large number of other nodes, the second is the evaluation that the node could be closer or not to a straight line joining the starting point to the destination.

The square matrix A previously described has been generated in order to distinguish between the couple of way points that are connected and those that are not linked each other.

The generic element A(*x,y*) represents the condition of the segment between the couple of way points *x,y*.

If the path between *x* and *y* is connected by a taxiway that can be covered in the direction from *x* to *y* then the value assigned to A(*x,y*) is 1. Otherwise, if the path between *x* and *y* is not connected by a taxiway or the taxiway cannot be covered in the direction from *x* and *y*, then the value assigned to A(*x,y*) is 0.

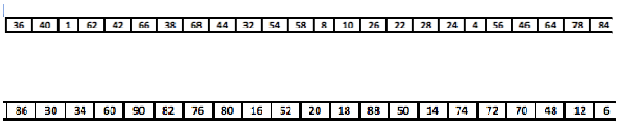
Fig. 6 shows an extrapolation of the total matrix A.



	1	2	3	4	5	6	7	8
1	0	1	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0
3	0	0	0	0	0	1	0	0
4	1	0	0	0	1	0	0	0
5	0	0	0	0	0	1	0	0
6	0	0	0	0	0	0	1	0
7	0	0	0	0	0	0	0	0
8	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	1
13	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0

Fig. 6 A part of the connection's matrix A.

The sequential procedure for the path encoding described in paragraph 2.1.1 has been applied and the priority vector has been created: the 45 cells priority vector is shown in Fig. 7 as follows. It will be automatically updated at each iteration.



36	40	1	62	42	66	38	68	44	32	54	58	8	10	26	22	28	24	4	56	46	64	28	84

36	30	34	60	90	82	76	80	16	52	20	18	88	50	34	74	72	70	48	12	6

Fig. 7 The priority vector, displayed in two lines.

For this application, 45 considered waypoints, the heuristic operator M has been set to the value 10.

Once the priority vector is implemented into the PSO algorithm, the fitness function calculate the cost relative to a first attempt particle: this fitness function has been slightly modified to best fit the proposed application.

The matrix A has been synthesize the constraints of the real case study of the Rome Fiumicino airport. A similar matrix A' has been created in which the 0 values are inverted to a large positive value. Then product between the matrix A' and the matrix C of the real costs of each node's connection becomes the argument of the fitness function to be maximized.

The large positive number inserted in the matrix A' means that two nodes are not connected, this means that the fitness resulting from a path that considers not feasible connections assumes a value close to 0.

4 Conclusion

A new meta-heuristic approach is proposed in this paper in order to overcome the problems of computational efficiency proper of the classical combinatorial optimization problem.

A path encoding has been formulated in order to make the path optimization evaluable by means of Particle Swarm Optimization.

A real case study is proposed to test the SP problem with a new approach. As first attempt, only a comparison between the SPSO solution and the solution given by Dijkstra Algorithm presented in [1] is covered in this paper. We compared two best routes found between a stand waypoint, considered as starting point, and the waypoint on a joint of the runway, considered as arrival point. In particular the considered waypoints are the 45th and the $4^{th.}$ The comparison highlights that the sequence of visited nodes is the same and that the computational times are comparable in the two cases.

Further tests will be undertaken to enhance the methodology proposed and to confirm the validity of the Shortest-path Particle Swarm Optimization (SPSO) solutions.

References

[1] S. Bagassi, D. Francia, F. Persiani, 'Simulating airport operations in a synthetic environment', Proceedings of IMProVe 2011, International Conference in Innovative Methods in Product Design, 2011.

[2] F.B. Zahn, C.E. Noon, Shortest path algorithms: An evaluation using real road networks, Transport. Sci. Vol.32, 1998, pp. 65–73.

[3] G. Desaulniers, F. Soumis, 'An efficient algorithm to find a shortest path for a car-like robot', IEEE Trans. Robot. Automat. Vol.11, No. 6, 1995, pp. 819–828.

[4] E.L. Lawler, 'Combinatorial Optimization: Networks and Matroids', Holt, Rinehart, and Winston, New York, 1976, pp. 59–108.

[5] M.K. Ali, F. Kamoun, 'Neural networks for shortest path computation and routing in computer networks', IEEE Trans. Neural Netw. Vol. 4, 1993, pp. 941–954.

[6] J.Wang, 'A recurrent neural network for solving the shortest path problem', Proceedings of the IEEE International Symposium on Circuits and Systems, 1994, pp. 319–322.

[7] F. Araujo, B. Ribeiro, L. Rodrigues, 'A neural network for shortest path computation,' IEEE Trans. Neural Netw. Vol. 12, No. 5, 2001, pp. 1067–1073.

[8] M. Munemoto, Y. Takai, Y. Sato, 'A migration scheme for the genetic adaptive routing algorithm', Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 1998, pp. 2774–2779.

[9] J. Inagaki, M. Haseyama, H. Kitajima, 'A genetic algorithm for determining multiple routes and its applications', Proceedings of the IEEE International Symposium on Circuits and Systems, 1999, pp. 137–140.

[10] C.W. Ahn, R.S. Ramakrishna, 'A genetic algorithm for shortest path routing problem and the sizing of populations', IEEE Trans. Evol. Comput. Vol. 6, No. 6, 2002, pp. 566–579.

[11] M. Gen, R. Cheng, D.Wang, 'Genetic Algorithms for solving shortest path problems', Proceedings of the IEEE International Conference on Evolutionary Computation, 1997, pp. 401–406.

[12] J. Kuri, N. Puech, M. Gagnaire, E. Dotaro, 'Routing foreseeable light path demands using a tabu search meta-heuristic', Proceedings of the IEEE Global Telecommunication Conference, 2003, pp. 2803–2807.

[13] J. Kennedy, R.C. Eberhart, 'Particle swarm optimization', Proceedings of the IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.

[14] A. W. Mohemmed, N. C. Sahoo , T. K. Geok 'Solving shortest path problem using particle swarm optimization', Applied Soft Computing, Vol. 8, No. 4, 2008, pp. 1643-1653.

CEAS 2011

Biofuels and Bioenergy

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Bioenergy in Aviation: The Question of Land Availability, Yields and True Sustainability

Florian Riegel and Jutta Steinsdörfer
Bauhaus Luftfahrt e.V., Lyonel-Feininger-Strasse 28, 80807 Munich, Germany

Keywords: *aviation, bioenergy, biofuel yields, land availability, sustainability*

Abstract

The aviation industry has identified the development of bio-derived hydrocarbon fuels as one of the major ways it can supply its needs whilst reducing its greenhouse gas emissions significantly. Due to recent advances in research and technology, aviation biofuels are likely to be available for commercial use within the next few years. A sufficient biomass resource base and a well-performing biomass market are basic preconditions to realize such an ambition.

A wide range of studies have been conducted in the field of biomass potentials. The assessments draw completely different conclusions about the role of biomass in the global energy supply. For this reason, Bauhaus Luftfahrt has developed a novel approach to estimate the technical potential for bioenergy on the basis of spatially explicit data and considering a comprehensive set of sustainability criteria.

This conference paper goes into detail about assumptions and methodologies and presents preliminary results of the analysis.

1 Introduction

One cannot disregard the fact that fossil fuels are being depleted and the major cause of global warming. Especially the transport sector is challenged and shaped by a high dependency on fossil fuels as high density, liquid energy carriers.

The advances the aviation industry has made in reducing its per passenger energy consumption and its relative CO_2 emissions are significant: The aerodynamics of aircraft, the performance and efficiency of aero engines and the operational improvements by airlines, airports and the air traffic system, all combined, have made air transport more than 70 percent more fuel efficient over the past 40 years [1].

Nevertheless, the aviation industry is willing to achieve further improvements in the way aircraft are constructed and how they are flown. In June 2009, IATA airlines took a landmark decision to adopt a set of ambitious targets: 1) a cap on aviation CO_2 emissions from 2020 ("carbon-neutral growth"); 2) a reduction in aviation CO_2 emissions of 50 percent by 2050 as compared to 2005 levels [2].

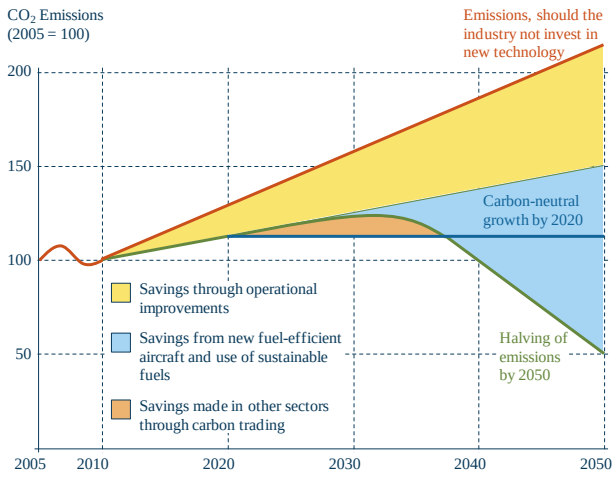


Figure 1. The IATA CO_2 emission reduction targets (schematic; based on [3]).

However, emission mitigation measures such as new aircraft designs, new engine advances or the development of more efficient aircraft operations will be insufficient to compensate for the forecast growth in air travel (Fig. 1). To fill this gap, the aviation industry has recognized the development of sustainable "drop-in" fuels as one of the major ways it can supply its needs whilst reducing its carbon footprint significantly.

Bio-derived hydrocarbon fuels provide an attractive renewable energy option: spatially well-distributed, easy to access and carbon neutral at first sight. On the other hand, biofuels can cause more damage than they prevent when, for example, tropical rainforests are cleared for new energy crop plantations.

As a result of recent advances in research and technology, aviation biofuels are likely to be available for commercial use within the next few years. A sufficient biomass resource base and a well-performing biomass market that can guarantee a reliable, sustainable and long-lasting biomass supply are essential preconditions to realize such an ambition.

2 State of the Art

In recent years, a considerable number of studies have been conducted in the field of biomass potentials. The assessments provide detailed and well-founded insight into the most important aspects and parameters for estimating the bioenergy potential.

In principle, recent studies

- feature *global* or, at most, *regional* coverage. Thus, none of the assessments focuses on a national scale.
- solely estimate the *geographical* potential for primary bioenergy; that is the theoretical potential at land available for the production of primary biomass, expressed in Exajoules (10^{18} Joules) per year (EJ yr^{-1}). Thus, none of the assessments focuses on the technical potential for bioenergy; that is the geographical potential minus losses due to the conversion of biomass to transportation fuels.
- derive scenarios for the *situation in 2050*. Thus, none of the assessments focuses on the current bioenergy potential.

Despite these similarities, recent studies draw completely different conclusions about the possible contribution of biomass to the future global energy supply: The highest estimate of 1,442 EJ yr^{-1} is almost 44 times higher than the lowest estimate of 33 EJ yr^{-1} (Fig. 2).

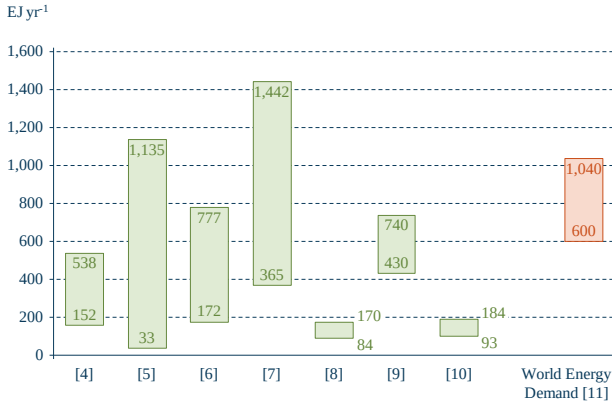


Figure 2. Global potentials for primary bioenergy in 2050 and the projected world energy demand, expressed in EJ yr^{-1}.

The main reason for the difference among the assessments is that their scope, in respect of biomass categories included, varies as much as methodological and scenario assumptions, primarily about land availability and biomass yields.

3 Approach

Bauhaus Luftfahrt – a Munich-based research institution dealing with the longer-term future of aviation – has developed a novel approach to estimate the *technical* potential for bioenergy on the basis of spatially explicit data and considering a comprehensive set of sustainability criteria (Table 1).

Table 1. Sustainability criteria.

	Criterion
1	Biofuels shall contribute to *climate change mitigation* by reducing greenhouse gas emissions as compared to fossil fuels.
2	Biomass production shall not violate *human rights* or *labor rights*.
3	Biomass production shall not impair *food security*.
4	Biomass production shall avoid negative impacts on *biodiversity*.
5	Biomass production shall not result in *deforestation*.
6	Biomass production shall not result in *soil nutrient depletion*.
7	Biomass production shall not result in *fresh water depletion*.

Table 2. Relevant geographic information.

Spatial Phenomenon		Category	Data Source
Inland Water Bodies	Surfaces occupied by rivers, lakes and reservoirs	not suitable	*Global Land Cover Characteristics Database*
Forest Areas	Surfaces spanning more than half a hectare with trees higher than five meters and a canopy cover of more than ten percent	not accessible (cf. criterion 5)	*Global Forest Resources Assessment*
Constrained Habitats	Surfaces with a $LGP_{t=5}$ of less than 120 days[1]	not suitable	Average climate data for the period from 1961 to 1990
	Surfaces with a LGP of less than 60 days[2]	not suitable	Average climate data for the period from 1961 to 1990
	Surfaces with slopes greater than 30 percent	not suitable	*Global Digital Elevation Model*
Conservation Zones	Surfaces which receive protection because of their environmental or similar value	not accessible (cf. criterion 4)	*World Database on Protected Areas*
Settlement Areas	Surfaces dominated by all non-vegetative, human-constructed elements	not accessible	Combination of satellite imagery, census information and nighttime lights data

[1] $LGP_{t=5}$ is used to indicate the number of days with daily mean temperature above five degrees Celsius.
[2] LGP is used to indicate the number of days when moisture conditions are conducive to crop growth.

The bottom-up analysis is conducted on a *national* scale and addresses the potential for land-based biomass. Hence, no aquatic biomass is included in the study. Beyond that, the assessment restricts itself to dedicated crops on surplus agricultural land; that is land not needed for the production of food and feed.

One could think of further categories of land-based biomass, for example:

- dedicated energy crops on degraded, marginal land; that is land not suitable for the production of food and feed.
- surplus forestry; that is forest growth (natural increment) not required for industrial roundwood and fuelwood production.
- agricultural residues; that is residues (e.g. straw and oilcakes) released together with crop harvesting and processing.
- forestry residues; that is residues released together with wood harvesting (e.g. branches) and processing (e.g. sawdust).

However, none of the categories given above shall be considered in the analysis, since

1) biomass production on degraded, marginal land jeopardizes the important function of disturbed, degraded soils as a substrate for natural vegetation (cf. criterion 6).
2) surplus forestry piles the pressure on biodiversity and involves the risk of extensive deforestation (cf. criteria 4 and 5).
3) no less than half of the released crop and wood harvest residues have to be left in place to maintain soil fertility [6]. The remainder should be withheld to ensure the supply of traditional biomass for cooking and heating, primarily in least developed countries (cf. criteria 6 and 2).

At this stage, the study focuses on the *current* (i.e. 2007) bioenergy potential. Later on, sensitivity analyses will provide a closer look at global bioenergy potentials for the year 2050.

4 Methods and Preliminary Results

The assessment applies a three-step methodology: In the first step, starting from the total area of each individual country in the world, surfaces not suitable and/or not accessible for agriculture (Table 2) are removed to calculate the “net area” suitable and accessible for agriculture (4.1).

In the second step, referring to three different food consumption patterns (Table 5), the respective land requirements for a calorie- and protein-sufficient diet are taken into account to estimate the surplus of agricultural land potentially available for energy crop cultivation (4.2).

In the third step, based on an in-depth review of international scientific literature, an inventory of most relevant first generation and next generation feedstocks is established to derive technically feasible yields in biofuel production (4.3).

Table 3. Actual extent of relevant phenomena in km^2 (upper value) and percent area (lower value), measured to calculate the EU-15 net area suitable and accessible for agricultural activity.

Country	Total Area[3]	Inland Water Bodies	Forest Areas	Constrained Habitats	*considered to be forest areas*	Conservation Zones	*considered to be inland water bodies/ forest areas/ constrained habitats*	Settlement Areas	*considered to be constrained habitats/ conservation zones*	Net Area
AUT	83,870	1,420 1.69	38,720 46.17	44,584 53.16	24,469 29.17	822 0.98	216 0.26	2,005 2.39	307 0.37	21,311 25.41
BEL	30,530	250 0.82	6,747 22.10	– –	– –	64 0.21	1 0.00	3,234 10.59	4 0.01	20,240 66.30
DEU	357,100	8,430 2.36	110,760 31.02	4,291 1.20	3,367 0.94	13,553 3.80	5,792 1.62	24,504 6.86	441 0.12	205,162 57.45
DNK	43,090	660 1.53	5,380 12.49	– –	– –	2,301 5.34	675 1.57	791 1.84	6 0.01	34,639 80.39
ESP	505,370	6,260 1.24	176,452 34.92	75,797 15.00	43,369 8.58	3,280 0.65	1,220 0.24	4,424 0.88	654 0.13	284,400 56.28
FIN	338,420	34,330 10.14	221,570 65.47	30,912 9.13	11,051 3.27	11,750 3.47	10,407 3.08	460 0.14	– –	60,856 17.98
FRA	549,190	1,530 0.28	158,100 28.79	48,475 8.83	31,656 5.76	11,531 2.10	5,203 0.95	12,124 2.21	402 0.07	354,691 64.58
GBR	243,610	1,680 0.69	28,594 11.74	395 0.16	133 0.05	– –	– –	8,546 3.51	1 0.00	204,529 83.96
GRC	131,960	3,060 2.32	38,124 28.89	47,196 35.77	19,771 14.98	3,023 2.29	1,578 1.20	1,893 1.43	162 0.12	60,175 45.60
IRL	70,280	1,390 1.98	7,126 10.14	53 0.08	– –	897 1.28	236 0.34	282 0.40	3 0.00	60,771 86.47
ITA	301,340	7,200 2.39	89,150 29.58	82,260 27.30	54,223 17.99	8,091 2.69	6,260 2.08	13,416 4.45	1,102 0.37	162,808 54.03
LUX	2,590	– –	868 33.51	– –	– –	– –	– –	139 5.37	– –	1,583 61.11
NLD	41,540	7,780 18.73	3,650 8.79	– –	– –	2,008 4.83	775 1.87	4,474 10.77	32 0.08	24,435 58.82
PRT	92,090	620 0.67	34,446 37.40	4,193 4.55	3,091 3.36	1,405 1.53	856 0.93	778 0.84	8 0.01	54,603 59.29
SWE	450,300	39,960 8.87	282,030 62.63	82,002 18.21	29,499 6.55	41,015 9.11	36,963 8.21	1,226 0.27	6 0.00	70,535 15.66
EU-15	3,241,280	114,570 3.53	1,201,717 37.08	420,158 12.96	220,629 6.81	99,740 3.08	70,182 2.17	78,296 2.42	3,128 0.10	1,620,738 50.00

[3] obtained from http://faostat.fao.org/ (Effective 2007)

4.1 Calculation of Net Areas

Several ways exist to calculate the "net area" suitable and accessible for agriculture: Most studies rely on methods such as *land use allocation* and *land-claim exclusion*, ignoring that there are many uncertainties related to various factors. By contrast, the interpretation of remote sensing data and the usage of *Geographic Information Systems* is often considered too complex, in spite of the fact that today, Earth observation and weather satellites provide the highest spatial accuracy.

A Geographic Information System (GIS) is a computer-based tool capable of assembling, storing, analyzing and displaying georeferenced information; that is data identified according to its location on the globe. GIS datasets contain multiple themes for a common geographic area. Every theme can be managed as an information set

independent of other themes. Each has its representation as a collection of points, lines or polygons (features) or grid cells (rasters).

Geoprocessing forms an important part of the work with a GIS. There is a broad range of powerful spatial modeling tools for

- splitting global GIS datasets into multiple subsets of countries (*Extract by Mask*).
- distinguishing relevant from non-relevant geographic information (*Reclassify*).
- cross-checking spatial properties of a specific geographic location (*Combine*).
- measuring areas accurately on the Earth's surface.

In view of these apparent advantages, the authors have adopted a GIS-based approach to calculate agricultural land availability on a national scale. The goal is to identify and model surfaces not suitable and/or not accessible for agriculture (Table 2) at the highest available level of detail. Hence, the spatial resolution of the selected GIS datasets ranges between 15 arcseconds (the area a grid cell represents is 0.22 km^2 at the Equator) and five arcminutes (the area a cell represents is about 86 km^2 at the Equator).

As a matter of course, spatial phenomena are rarely discrete, but have the tendency to overlap one another. To give an example, a forested area may well be part of a conservation zone, etc. To ensure that all such surfaces are considered only once, the actual extent of each individual phenomenon is measured taking account of the immediately afore measured area/s (Table 3).

Table 4. Verification of spatially explicit data using the example of Germany.

		Spatially Explicit Data	Statistical Data[4]
Inland Water Bodies	in km^2	8,430	8,482
	in percent	2.36	2.38
Forest Areas	in km^2	110,760	107,349
	in percent	31.02	30.06
Constrained Habitats	in km^2	4,291	n/a
	in percent	1.20	n/a
Conservation Zones	in km^2	13,553	12,716
	in percent	3.80	3.56
Settlement Areas	in km^2	24,504	24,416
	in percent	6.86	6.84

[4] obtained from http://www.destatis.de/jetspeed/portal/cms/ (Effective 2007)

Table 3 shows preliminary results on the EU-15 net area suitable and accessible for agricultural activity. As can be seen using the example of Germany, the high resolution and close-to-reality strategy of the analysis leads up to results that assort well with solid statistical data (Table 4).

4.2 Estimation of Food Areas

As may be imagined, not all surfaces suitable and accessible for agricultural activity are available for biomass production as well. In fact, significant areas serve as cropland and grassland to supply food demand (cf. criterion 3).

First and foremost, food demand is subject to population and the per capita food consumption. This is why the study at hand distinguishes three different food consumption patterns (Table 5): a vegetarian diet with little or no animal protein, a moderate diet and finally, an affluent diet with a large share of dairy and meat products [12].

Table 5. Per capita food consumption for three different diets, expressed in kcal day^{-1} and in grain equivalents in kg day^{-1} [12].

		Vegetarian Diet	Moderate Diet	Affluent Diet
Caloric Intake	in kcal day^{-1}	2,412	2,412	2,747
Plant Products	in $kg_{gr\ eq}$ day^{-1}	1.05	0.90	1.13
Dairy Products	in $kg_{gr\ eq}$ day^{-1}	0.28	1.23	1.16
Meat Products	in $kg_{gr\ eq}$ day^{-1}	–	0.22	1.91
Total	in $kg_{gr\ eq}$ day^{-1}	1.33	2.35	4.20

The diets are all sufficient with regard to daily caloric intake and daily protein requirements, but differ strongly with respect to their composition. To make the diets comparable, they are expressed in grain equivalents (gr eq); that is measures for the amount of dry weight in grains utilized, directly or indirectly, in food consumption. Crops which are not cereals (e.g. fruit and vegetables) are translated to grains. Moreover, losses when converting grains and grasses to dairy and meat products are taken into account [12].

To estimate food demand on a national scale, the existing nutrition behavior of a country's po-

pulation is compared with the food consumption patterns described above, and the best fitting diet is selected. This allows to reproduce the actual shares of plant, dairy and meat products in a country whilst averting the risk of malnutrition.

The ratio between food demand (the per capita food consumption multiplied with the population size) and the country-weighted mean of irrigated/non-irrigated grain yields gives some indication of the country-specific land requirements for a calorie- and protein-sufficient diet. It is believed that any surplus of agricultural land could be used for energy crop cultivation.

For a variety of reasons, a country may reach a self-sufficiency below 100 percent. In this particular case, the remaining demand for food products is allocated to a country (or a region) with a remaining production potential.

4.3 Derivation of Yields

Dedicated energy crops, as defined by recent studies, are restricted to *woody biomass* (i.e. eucalyptus, poplar, willow) grown in short rotation coppice (SRC). For the study at hand, 26 *additional* first generation and next generation *feedstocks* are taken into account. With this in mind, more than 200 journal articles, conference papers and field reports have been reviewed.

Generally, first generation feedstocks include sugar (e.g. sorghum and sugar cane), starch (e.g. cassava and maize) and oilseed crops (e.g. jatropha and oil palm) that can be converted to bioalcohols, biodiesel or hydrotreated esters and fatty acids (HEFA). In contrast, next generation feedstocks comprise perennial herbaceous plant species (e.g. miscanthus and switchgrass) and short rotation woody crops which can be converted to cellulosic ethanol or biomass-to-liquids (BTL).

The in-depth literature review led to a consistent knowledge database on

- plant anatomy and plant chemistry.
- cultivation targets and specific habitat requirements.
- crop rotation and cultivar selection.
- fertilization and pest management.
- energy output and energy efficiency.
- water consumption and climate impact.

This Bauhaus Luftfahrt inventory [13] allows an exact matching of physiologically well-suited energy crops to a given cultivation area. Beyond that, it provides adequate information to exclude

1) feedstocks with a negative (deficit) water balance (cf. criterion 7).
2) feedstocks with life cycle greenhouse gas emission savings [14] of less than 35 percent as compared to fossil jet fuel (cf. criterion 1).

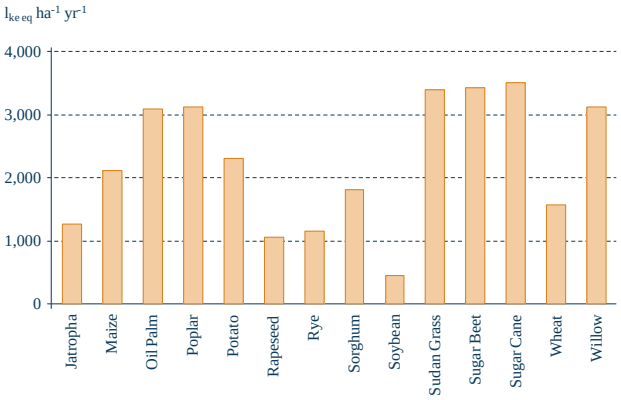


Figure 3. Average productivities of first and next generation feedstocks, expressed in kerosene equivalents in l ha^{-1} yr^{-1} (based on [13]).

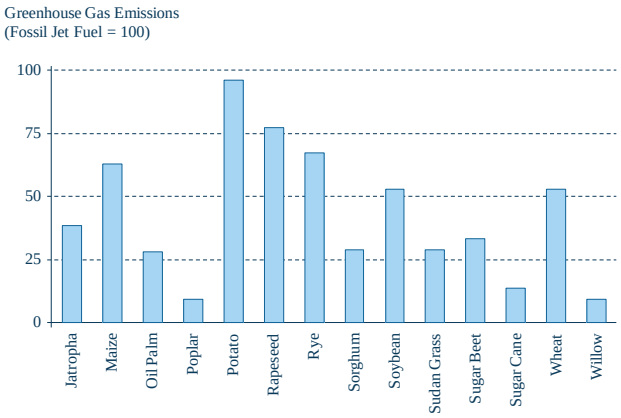


Figure 4. Average greenhouse gas emissions of first and next generation feedstocks as compared to fossil jet fuel. Life cycle assessments are all consistent with respect to system boundaries (based on [13]).

In the end, the expectable biofuel productivities of the remaining energy crops are decisive for the final allocation of one of these feedstocks to the available cultivation area. For this purpose, biofuel yields are translated to kerosene equivalents (ke eq); that is measures for the amount of fossil jet fuel it would take to equal the energy content of a certain biofuel.

As can be seen from preliminary results (Fig. 3 and 4), high-yielding energy crops like poplar, sudan grass or sugar cane might be able to contribute significantly to climate change mitigation as well, at least on global average. Lower yielding crops (e.g. rapeseed and soybean), however, are mostly accompanied by less valuable greenhouse gas emission savings.

The combination of land availability and biofuel yields clearly indicates the technically feasible potential for bioenergy on a national (by aggregation of cultivation areas), a regional/continental (through aggregation of country profiles) and finally, on a global scale (by aggregation of continent profiles).

At the close of the assessment, to ensure that biomass production is not only environmentally sound, but also socially fair, particular attention is payed to the traditions by which authority in a country is exercised. The *Worldwide Governance Indicators* (WGI) project provides individual governance indicators for six dimensions of governance (e.g. control of corruption and political stability). These indicators combine the views of a considerable number of enterprise, citizen and expert survey respondents [15]. Economies that violate applicable law and international conventions (cf. criterion 2) are removed from the analysis.

5 Conclusions

The study at hand estimates the current technical potential for bioenergy in an unprecedented depth and quality. Hence, the authors make a substantial contribution to the lively discussion about the amount of fossil jet fuel that could be substituted *sustainably* by bio-derived fuels.

Even though other industry sectors have feasible alternative energy solutions currently available – for example, the power generation industry may look to wind, hydro and solar technologies; cars and buses may run on hybrid, flexible fuel engines or electricity – the aviation industry will have to compete for bioenergy and the biomass resource base. This competition will be taken into account when interpreting and discussing the final results of the assessment (expected mid-year 2012).

References

[1] ATAG, “Beginner’s guide to aviation biofuels”, 2009.

[2] IATA, “A global approach to reducing aviation emissions”, 2009.

[3] IATA, “IATA 2010 report on alternative fuels”, 2010.

[4] Berndes G., Hoogwijk M. and van den Broek R., “The contribution of biomass in the future global energy supply: a review of 17 studies”, *Biomass and Bioenergy*, Vol. 25, 2003, pp. 1-28.

[5] Hoogwijk M. et al., “Exploration of the ranges of the global potential of biomass for energy”, *Biomass and Bioenergy*, Vol. 25, 2003, pp. 119-133.

[6] Hoogwijk M., Faaij A., Eickhout B., de Vries B. and Turkenburg W., “Potential of biomass energy out to 2100 for four IPCC SRES land-use scenarios”, *Biomass and Bioenergy*, Vol. 29, 2005, pp. 225-257.

[7] Smeets E., Faaij A., Lewandowski I. and Turkenburg W., “A bottom-up assessment and review of global bio-energy potentials to 2050”, *Progress in Energy and Combustion Science*, Vol. 33, 2007, pp. 56-106.

[8] Beringer T. and Lucht W., “Simulation nachhaltiger Bioenergiepotentiale”, WBGU, 2008.

[9] Dornburg V. et al., “Biomass assessment: assessment of global biomass potentials and their links to food, water, biodiversity, energy demand and economy”, Netherlands Environmental Assessment Agency, 2008.

[10] Greenpeace International and EREC, “A sustainable global energy outlook”, 2008.

[11] UNDP, UNDESA and WEC, “World energy assessment: energy and the challenge of sustainability”, 2000.

[12] Luijten J., “Sustainable world food production and environment”, Dutch Agricultural Research Department, 1995.

[13] Janzon A., “Eine geobotanische Studie über Bestand, Verfügbarkeit und Kultivierung global relevanter Energiepflanzen”, Diploma Thesis, Ludwig-Maximilians-Universität Munich, Germany, 2010.

[14] OPOCE, “Directive 2009/28/EC of the European Parliament and of the Council on the promotion of the use of energy from renewable sources”, *Official Journal of the European Union*, Vol. L 140, 2009, pp. 16-62.

[15] Kaufmann D., Kraay A. and Mastruzzi M., “The worldwide governance indicators: methodology and analytical issues”, World Bank, 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Object-Orientated Development of a Model Library for Fuel Cell Cooling Systems Using Modelica

Enno Vredenborg, Hauke Lüdders, Jan Grymlas and Frank Thielecke
TUHH - Hamburg University of Technology, Institute of Aircraft System Engineering, Neßpriel 5, 21129 Hamburg, Germany

Keywords: *Fuel cell, Modelica, aircraft cooling architecture, renewable energy*

Abstract

One aim of research projects for civil aircraft is the improvement of sustainability. Major concerns are the reduction of CO_2/NO_x emissions and noise pollution. The introduction of fuel cells as an independent energy source can help to reduce the environmental impact.

In order to develop fuel cell systems the model library FuelCellCooling in Dymola/Modelica is created. The library is especially used to study cooling systems for fuel cells. To cool fuel cells complex architectures with numerous components and different fluids are necessary. The library provides quasi-static and dynamic models to be able to meet different simulation objectives.

Object-orientated elements of Modelica are used to structure the model library. The essential elements are composition, inheritance and polymorphism. The library is extended by free available and commercial libraries, for example, to model fluid properties or to use general defined interfaces. The use of the library is demonstrated in this paper by an industry-related aircraft example.

1 General Introduction

The environmental impact of aircraft should be reduced as much as possible. Beside the influence on global warming, exhaust emissions and noise pollution are main topics discussed for future developments. A second driver for new aircraft generations are improvements of efficiency, weight and reliability. Especially on the overall aircraft level electrical systems pretend to meet these goals.

At the interface between these two major streams the integration of fuel cells is well-founded. A fuel cell delivers electrical power with a high efficiency, almost noiseless and without any pollutant emissions. System development focuses on polymer electrolyte membrane (PEM) fuel cells. In the fuel cell chemical energy from hydrogen and atmospheric oxygen is converted into electrical energy. The fuel cell electrical output power supplies the electrical network on ground and in flight and is therefore suitable to replace the gas turbine auxiliary power unit [1].

As in case of other electrical technologies the cooling system for a fuel cell is an additional complex attachment with a significant impact on weight. For designing, analyzing and assessing cooling technologies a model library in

Dymola/Modelica was created and is presented in this article.

2 Polymer Electrolyte Membrane Fuel Cell

PEM fuel cells are used for many mobile applications because of their high power density and a manageable working temperature of about 80 °C (High temperature PEM fuel cells up to 180°C). As Figure 1 shows, the hydrogen is added on the anode and transported to the catalyst layer via the gas-diffusion layer (GDL). The hydrogen reacts to protons and electrons. The protons pass the membrane, the electrons flow through an external circuit to the anode and can be used as electrical energy. On the cathode side the electrons and the protons react with the atmospheric oxygen. The reaction product is water [2].

One cell is only able to generate a maximum voltage of about 1.23 V which is not sufficient for many applications. To reach a higher output voltage several cells are combined to a so called fuel cell stack.

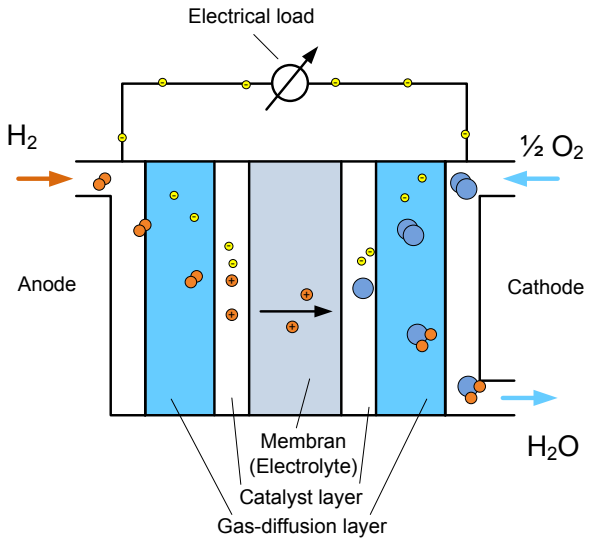


Fig. 1 Diagram of a PEM fuel cell [3]

3 Multifunctional Fuel Cell System

For the application of a fuel cell on-board of future aircraft as an independent energy source, a multifunctional approach is examined [4]. This system enables an optimal integration and prevents from weight disadvantages compared to conventional auxiliary power system. An essential aspect of this system is to use the actual exhaust gas. This gas has got a reduced oxygen concentration and contains the produced water. Thus it is a gas with firefighting properties.

The so called multifunctional fuel cell system (MFFCS) is able to fulfill the following functions:

- Independent energy supply on ground and in flight (especially for air conditioning, taxiing, commercial loads)
- Inert gas supply for fuel tanks
- Cargo bay fire suppression
- Emergency power supply
- Water generation for onboard-use

This system design is able to unify several functions, but it also leads to complex architectures with interfaces to several other systems. Figure 2 shows a schematic diagram with possible interfaces to relevant systems concerning thermal aspects.

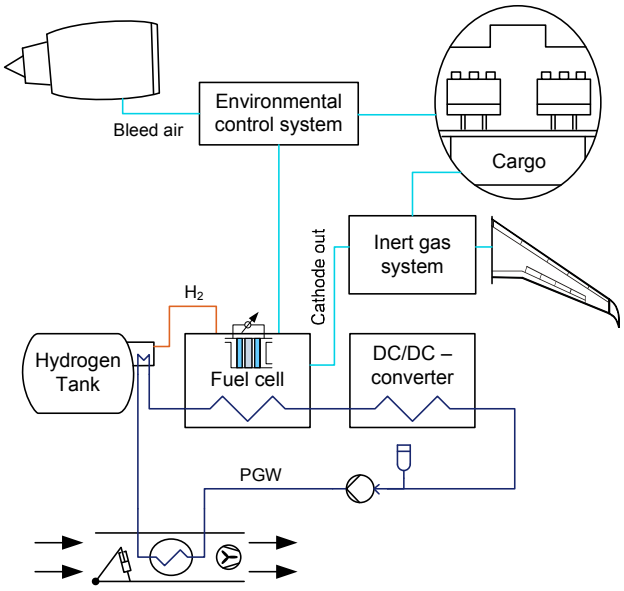


Fig. 2 Schematic MFFCS cooling system diagram

The system has to cool one or more fuel cells, connected power electronics and the inert gas system. Furthermore, a high interaction between the MFFCS and the environmental control system (ECS) needs to be taken into account. Possible heat sinks are ambient air and, if available, a liquid hydrogen tank.

The system consists of a variety of components, fluids and technical solutions, whereas some of them have not been used on commercial aircraft yet. For system design and assessment a model-based approach is necessary.

4 FuelCellCooling Library

To make the model-based approach possible a model library in Dymola/Modelica was created. The so called *FuelCellCooling* library should help to meet the following objectives:

- Verifying the design of cooling systems for the MFFCS and determining the key performance indicators mass, power consumption and aerodynamic drag,
- simulating and evaluating the thermal behavior for the MFFCS for whole flight missions and the interaction with neighboring systems,
- determine the dynamic system behavior,
- support in specific system engineering activities like control unit design and investigations of system behavior under failure condition.

As this list shows, the *FuelCellCooling* library must contain various models of components, fluids and subsystems with different modeling approaches. This section presents the used modeling language and the library structure.

4.1 Dymola/Modelica

The free available modeling language Modelica is an object-orientated language created for simulation of large and complex systems. The used Modelica version is Version 3.2 [5], Dymola 7.4 FD01 is the simulation environment .

A significant advantage of Modelica is the possibility of acausal modeling. This allows to use component models in flexible ways. The differences to a causal model (for example in MATLAB/Simulink) can be demonstrated with a pipe model for one-dimensional, incompressible fluid flow. The pipe is discretized with one volume. It is assumed that the fluid properties (like the fluid density ρ_{fluid}), the loss factor ζ and the pipe parameters are constant. Additionally, the momentum balance is reduced to the following pressure drop correlation:

$$p_{out} - p_{in} = \zeta \cdot \frac{\dot{m}_{in}^{\ 2} \cdot \rho_{fluid}}{2}, \qquad (1)$$

with the input pressure p_{in}, the output pressure p_{out} and the mass flow rate $\dot{m}_{in}$ at the input port. Due to the constant fluid properties the mass flow rate $\dot{m}_{out}$ at the output port can be calculated with the mass balance:

$$0 = \dot{m}_{out} + \dot{m}_{in}, \qquad (2)$$

and the dynamic energy balance can be written as:

$$\frac{dU}{dt} = \dot{m}_{in} h_{in} + \dot{m}_{out} h_{out}. \qquad (3)$$

Here U is the internal energy, h_{in} and h_{out} are the enthalpies at the input and output interfaces. With Modelica these equations can simply be implemented in the original mathematical form. In Dymola a system of differential-algebraic equations (DAE) is created and solved. The interfaces with other models are realized with connectors.

The corresponding Simulink block diagram is shown in Figure 3. For a uniform flow direction this does not lead to problems, but a changing flow direction cannot be simulated with this model. It is also not possible to use this model with changed inputs and outputs. To determine the mass flow rate depending on the input and output pressure the model has to be recreated.

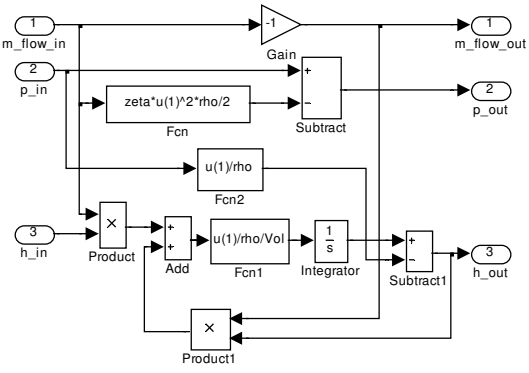


Fig. 3 Simple pipe model in Simulink

Connecting two models in Modelica is reached by the use of connectors. They contain all physical quantities with transfer across the component boundary. Normally, the same connector class is used for the same kind of interface, like mechanical, electrical or fluid flow interfaces. This leads to a system modeling which is close to the physical system and can be presented in a clear way in the Dymola diagram layer as shown in Figure 6.

4.2 Modelica libraries

A strength of Modelica is the availability of free libraries that deliver basic model structures, a variety of general used models and correlations for heat transfer and pressure drop. The following free available libraries supplement the *FuelCellCooling* library:

- *Modelica Standard Library* 3.2
- *ExternalMedia Library*
- *FluidDissipation*

The Modelica Standard Library is developed by the Modelica Association and delivers a huge number of models and interfaces from various engineering domains. Especially the *Modelica.Fluid* and *Modelica.Media* [8] packages are used by the *FuelCellCooling* library. The interfaces for one dimensional fluid flow in the *FuelCellCooling* library are taken from *Modelica.Fluid*. This ensures that models from different libraries are widely compatible. Additionally, conservation laws and general used models like fluid sinks and sources are taken from this package. The *Modelica.Media* package is used for fluid modelling. It contains basic fluid models and elements for defining fluid models by the user.

Due to a limitation of ready-to-use fluid models, especially models which are used to simulate fluid properties in the two-phase region, the *ExternalMedia* [6] library is used. More details about fluid modeling are described later. *FluidDissipation* [7] is a library which contains correlations for heat transfer and pressure drop. The library is able to interoperate with other libraries and the functions are numerically optimized.

To extend the *FuelCellCooling* library two additional commercial/scientific libraries are used:

- *TILMedia*
- *DENECS*

TILMedia is an additional media library, which delivers a wide range of fluid models and is compatible to the *ExternalMedia* library. The *DENECS* [9] library is an Airbus/DLR development to evaluate ECS architectures at aircraft level.

4.3 Library content

The *FuelCellCooling* library is developed to evaluate cooling systems for fuel cells. That is the reason why the library has to provide models of proposed components and media. Additionally, the library contains elements to simplify complex simulations. The library structure was designed to meet the users needs and to benefit from the object-orientated elements of Modelica. This starts with the top-level package arrangement. The *FuelCellCooling* library contains the following top-level packages:

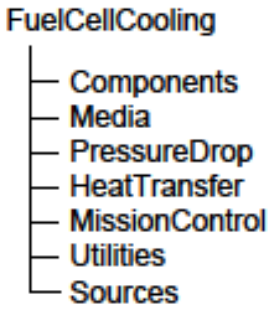


Fig. 4 FuelCellCooling top-level packages

The `Components` package contains all relevant models for the cooling system, for example fuel cells, pipes and pumps. In the `Media` package fluids and solid properties are modeled which aren't part of the *Modelica.Media* package, for example special propylene-glycol water mixtures or aluminum. The `PressureDrop` and `HeatTransfer` packages include friction and heat transfer characteristics for diverse application. They are summarized in extra packages, because they can be

used by different models. For describing aircraft typical boundary conditions the `MissionControl` package contains for example electrical load analysis and typical flight missions. The `Utilities` package aids the user to build up system models, in the `Source` package generic start and endpoints are deposited.

4.4 Library Structure and Object-Orientated Elements

In order to create and to handle the multitude of different components, mediums and physical correlations a clearly arranged library structure has to be established. Modelica enables the uses of object-orientated elements. An appropriate use of these elements allows the reusability of developed programming code and a simplification of models. Whereas a too extensive use of object-orientated design leads to problems of understanding and utilization. [10] proposes basic structures for Modelica libraries which are adapted to the *FuelCellCooling* library. The essentially used object-orientated elements are:

- **Composition**: This element is usually used to build a system model out of the individual component models. The component models are part-of the system model. This method leads to system model structure that is near to the real physical system.
- **Inheritance**: To get more specialized models from a base class inheritance can be used. For example, different fuel cell models (e.g. liquid and evaporative cooled) can reuse programming code they have got in common.
- **Polymorphism**: With this element it is possible to make a class exchangeable. For example a given characteristic diagram of a fan can be replaced by other ones. This technique is also used for declaring medium models in components. So a component is not restricted to one medium model.
- **Aggregation**: In Modelica aggregation is realized with the inner/outer concept. This allows to reference an element (declared with the prefix outer) to another element with the same name (declared with the prefix inner) in an enclosing instance hierarchy. Aggregation is not directly used in the *FuelCellCooling* library, but in the Modelica.Fluid package.
- **Multiple inheritance**: Also multiple inheritance can be realized in Modelica. This enables to inherit from more than one base class. In the *FuelCellCooling* library it is tried to avoid structures with multiple inheritance because this can lead easily to complex structures. However, since it is used in Modelica.Fluid it is also part of the *FuelCellCooling* library.

4.5 Re-used Elements

As mentioned above, using object-orientated elements makes it possible to set up large libraries, but at the same time it can also easily lead to confusing library structures. Therefore, it is essential to define which elements are suitable to be reused in a library. The following sections describe selected models and interfaces which are mainly used by different component models.

4.5.1 Interfaces

In Modelica the interfaces between models are realized with `connectors`. For one-dimensional fluid flow the *Modelica.Fluid* package provides the `FluidPort connector`. `FluidPort` defines the passing medium, the transferred mass flow rate and the absolute pressure. Additionally, the specific enthalpy, the fluid mass fraction and an extra property are defined. These dimension are defined with the prefix stream and so depend on the flow direction (see [11]).

Additional used interfaces are the `connector flange_a` interface from the *Modelica.Mechanics* package (defining the torque and the rotational angle), the `connector Pin` from *Modelica.Electrical* (defining current and voltage), `heatPort` from *Modelica.Thermal* (defining heat flow and temperature) and the `connectors RealInput/RealOutput`.

Using the connectors from the Modelica Standard library facilitates a combination of models from different libraries.

4.5.2 Conservation laws

An appropriate simulation of thermodynamic fluid flow needs a definition of the mass balance, the energy balance and the momentum balance. The `Modelica.Fluid` package defines the balance equation for dynamic and steady-state simulation. Additionally, the model initialization is treated in these classes.

It should be noted, that the conservation laws are not applied for all models. For incompressible fluid flow the momentum and the energy balance are decoupled from each other and an easier modeling structure can be chosen [12].

4.5.3 Media models

In almost every component model a media model is instantiated. By using polymorphism the media models are exchangeable. As it is not intended to use every medium in a component, the selection of medium models is restricted by the constrainedby command. For example, a pump model can only chose incompressible and table based medium models, a pipe model is adaptable to liquid, gaseous and two-phase fluids and has therefore no restrictions.

For more information about the media models themselves see section 4.7.

4.5.4 Correlations for pressure drop

The pressure drop calculation is similar for various components. For example, the pressure drop in a straight pipe and inside heat exchanger tubes can be determined using the same correlation. In addition, there are different calculation methods. You can calculate a pressure drop for laminar or turbulent flow, with different correlations or simplified with a constant loss factor. These aspects aim a use of polymorphism. In the PressureDrop package only pressure drop correlations are deposit (e. g. a friction factor), because a more general use is achieved. The pressure drop is calculated in the component models.

The *FuelCellCooling* library contains pressure drop correlations for pipes, corrugated louver fins and tube fins. Other correlations are taken from the FluidDissipation library.

4.5.5 Correlations for heat transfer

As well as correlations for pressure drop the same correlations for heat transfer can be used by different models. The library package contains heat transfer correlations for straight pipes, corrugated louver fins and tube fins. Here again, only dimensionless quantities are provided (e.g. the Nusselt number), because a more general use is possible. The heat transfer for a specific component is calculated in the related model. The *FluidDissipation* library provides additional correlations for a wide range of applications.

4.6 Modeling Components

To meet the required library objectives different modeling approaches are implemented. The library consist of quasi-static models (named simple) for a rapid evaluation of cooling architectures, dynamic models to determine the transient behavior and physical models which are able to simulate not nominal behavior. The different modeling approaches are presented in the following subsections for different examples.

4.6.1 SimplePump model

This model design is used to determine the cooling system performance. The models can predict the thermodynamic behavior, but do not include own dynamic behavior.

A typical fluid network component is a pump. To characterize the pump, the pressure increase $p_{out} - p_{in}$, the mass flow rate $\dot{m}$ and the power consumption P_{pump} are of major interest. The SimplePump model is designed to determine these parameters. A mass flow rate and an output pressure are directly set. The necessary pump input power is then determined with:

$$P_{pump} = \frac{1}{\eta_{pump}} \cdot \frac{\dot{m}}{\rho_{fluid}} \cdot (p_{out} - p_{in}), \qquad (4)$$

whereby η_{pump} is the pump efficiency, ρ_{fluid} the fluid density and p_{in} the pump input pressure which depends on the pressure losses in the cooling network. The model consists of a minimum numbers of equations. One simplification

is the assumption of an incompressible fluid flow. This decouples the momentum and the energy balance.

4.6.2 DynamicFan model

To characterize the transient system behavior an implementation of conservation laws with a time dependency is necessary. For example, the mass balance from equation (2) is extended by the mass m variation over time:

$$\frac{dm}{dt} = \dot{m}_{out} + \dot{m}_{in}. \qquad (5)$$

Also the momentum and energy balance are time dependent and coupled. So, compressible fluid flows can be simulated correctly.

A simplification that remains is the setting of the mass flow rate. The fan efficiency is determined by the volume flow rate and the pressure ratio. Knowledge of the fan speed is not necessary.

4.6.3 PhysicalCompressor model

A third level of model is implemented in the *FuelCellCooling* library to enable advanced simulation capabilities. This includes inter alia system behavior under failure conditions or the handling of reverse flow. Furthermore care is taken that the model structure and interfaces are similar to the real component. Instead of a prescribed mass flow rate the connector `flange_a` is used as the mechanical interface. This connector transfers the flange angle φ and torque τ. The use of this mechanical connector implies the availability of adequate characteristic diagrams, which have a dependency on the rotational speed.

4.7 Modeling Fluids

As with the component models different modeling approaches are chosen for fluid models. Modelica.Media delivers models for defining media computational code and data for essential medium models as moist air or water.

Gaseous fluids are modeled as an ideal gas, for moist-air the contained water is added as steam. These fluid models are sufficient, since no higher pressures (> 10 bar) are reached. They are used for the fuel cell hydrogen and air supply, for example.

Liquid fluids are implemented as incompressible fluids with a temperature dependency. The temperature influence of the density, the viscosity and other fluid characteristics are stored in tables. Fluid properties for specific applications can easily be created.

For simulating systems with two-phase fluids the available models are insufficient and the creation of own fluid models is a complex process. For this an external media library is expending the *FuelCellCooling* library. The ExternalMedia library is providing an interface to the software FluidProp [13]. This enables a use of existing fluid property code like it is available in the software REFPROP [14]. The ExternalMedia is compatible to the Modelica.Media library and so to the models of the *FuelCellCooling* library.

4.8 Component Mass Estimation

For system assessment the mass is an essential parameter. Therefore the mass of each component has to be determined. The following approaches are applied:

- For components already used in commercial aircraft a regression analysis can be applied. In [15] the determination of a pump mass m_{pump} is based on the maximum required pump volume rate Q. The pump mass is determined with:

  $$m_{pump} = a \cdot Q^b + c. \qquad (6)$$

  The parameters a, b and c are identified with built aircraft components.
- The mass of geometrically modeled components can be directly determined in the model. For this, the component volume (the material and the contained fluids) are multiplied with the corresponding density. Typical examples are pipe models or ε - NTU based heat-exchanger models (see [12]).
- If it is not possible to specify the component mass on the model parameters, the estimation will be based on technical predictions. For example, for fuel cell stacks

a power density of 2000 W/kg is the U.S. Department of Energy technical target for automotive applications [16].

4.9 Example

The application of a fuel cell as an independent energy source is demonstrated by the following example. A fuel cell system is considered, which consist of two fuel cell stacks. They deliver a required output power of 180 kW. Figure 6 shows the corresponding Modelica model in Dymola. The model includes the fuel cell stacks, a DC/DC converter and the cooling circuit. To cool the fuel cells and the DC/DC converter a propylene-glycol water (PGW) mixture is used. The cooling circuit consists of a coolant pump, a plate-fin heat exchanger, a fan and the piping.

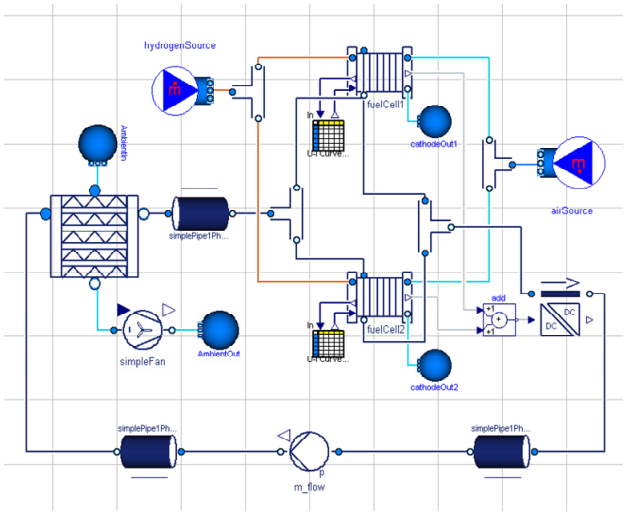


Fig. 5 Fuel cell system model in Dymola

For system assessment a typical short range flight mission is investigated. Figure 7 shows the flight altitude and the demanded electrical output power over mission time. On ground the fuel cells are responsible for energy supply and so the highest output power is requested. In flight the engine-driven generators are active and the fuel cells just have to provide inert gas. During descent a higher inert gas flow is necessary which leads to higher fuel cell activity.

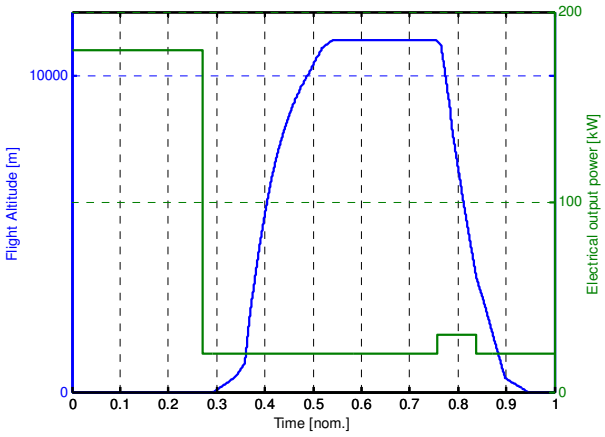


Fig. 6 Flight altitude and requested output power

The basic cooling requirements for the considered fuel cells are a limitation of maximum temperature to $T_{FC,\max} = 80\,°C$. The coolant temperature difference over one fuel cell should not exceed $\Delta T_{\max} = 10\,°C$. For determining the necessary coolant and cooling air mass flow rate a set of PI-controllers are used. They are included in the model to adjust the appropriate values, but not to represent a control system.

The cooling system demands are one simulation result (see Figure 7). For a hot day on ground (ambient temperature $T_{amb} = 38\,°C$) a cooling air mass flow rate of $\dot{m}_{air} = 5.8$ kg/s is required. The associated fan power and the cooling pumps power sums up to 7.5 kW which has to be provided by the fuel cells additionally to the required output power. Whereas in flight, the cooling air flow demand is reduced significantly due to lower waste heat and higher temperatures differences to ambient air.

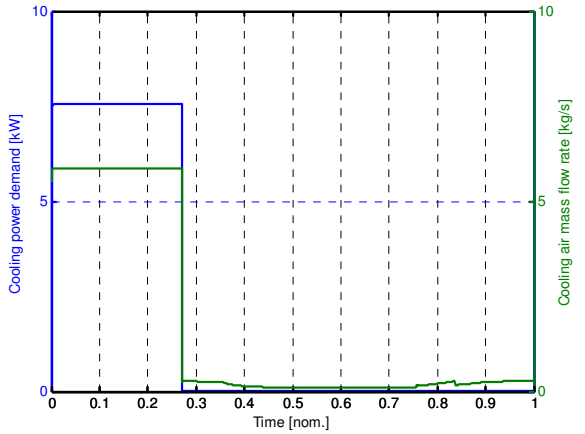


Fig. 7 Cooling system performance

Beside the cooling system demands the simulation also demonstrates fuel cell system advantages. The fuel cell efficiency and the used hydrogen are plotted in Figure 8. The consumed hydrogen for the whole flight mission is less than 8kg. This can be reached because of a fuel cell efficiency higher than 50%.

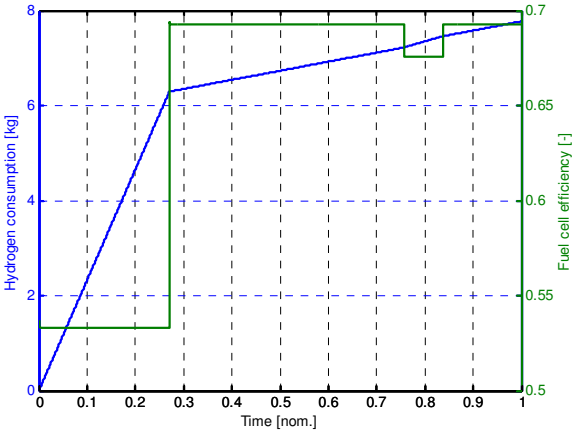


Fig. 8 Hydrogen consumption and fuel cell efficiency

5 Conclusion

Fuel cells are a promising technology to reduce the environmental impact of civil aircraft. The investigated PEM fuel cell converts chemical energy from hydrogen and atmospheric oxygen into electrical energy. To achieve an optimized integration a multifunctional approach is examined and presented in this article. This concept is based on the use of the reaction product water for onboard use and the oxygen reduced exhaust gas as an inert gas with fire-fighting properties. This leads to an additional complex system with interfaces to neighboring systems like the environmental control system.

The presented *FuelCellCooling* library is created for cooling architecture investigations and to examine the interdependencies with adjacent systems. The library is created in Modelica. Modelica provides free available libraries which enhance the *FuelCellCooling* library. The *ExternalMedia* library enables the simulation of various two-phase fluids.

Due to the numerous models object-orientated elements of Modelica are used. This allows the reuse of programming code, like pressure drop correlations or conservation laws.

This article concludes with a presentation of a fuel cell application in a civil aircraft.

References

[1] Lüdders H., Claußen S., Grymlas J., Vredenborg E., Thielecke F., "Challenges Faced by the Integration of a Fuel Cell System in Aircraft System Architectures", *H2Expo*, Hamburg, 2011.

[2] O'Hayre R., Cha S., Colella W. and Prinz F., *Fuel cell fundamentals*, 2nd ed., John Wiley and Sons, New-York, 2009.

[3] Siemer, M., "Lokale Entropieproduktionsraten in der Polymerelektrolyt-Membran-Brennstoffzelle," Ph.D. Dissertation, Institut für Thermodynamik, Helmut Schmidt Universität, Germany, 2007.

[4] Enzinger, M., "Technology Programs – Multifunctional Fuel Cell Application," *Deutscher Luft- und Raumfahrtkongress*, Hamburg, 2010.

[5] "Modelica® - A Unified Object-Oriented Language for Physical Systems Modeling - Language Specification Version 3.2," Modelica Association, 2010.

[6] Casella, F. and Richter C., "ExternalMedia: A Library for Easy Re-Use of External Fluid Property Code in Modelica," *Proceedings 6th Modelica Conference*, Bielefeld, 2008, pp. 157-161.

[7] Vahlenkamp T. and Wischhusen S., "FluidDissipation for Applications - A Library for Modelling of Heat Transfer and Pressure Loss in Energy Systems," *Proceedings 7th Modelica Conference*, Como, 2009, pp. 122-131.

[8] Franke R., Casella, F., Sielemann M., Prölß, K., Otter M. and Wetter M., "Standardization of Thermo-Fluid Modeling in Modelica.Fluid," *Proceedings 7th Modelica Conference*, Como, 2009, pp. 122-131.

[9] Sielemann M., Giese T., Oehler B. and Otter M., "A Flexible Toolkit for the design of environmental control system architectures," *Proceedings of the first CEAS European Air and Space Conference*, Berlin, 2007.

[10] Richter C., "Proposal of New Object-Oriented Equation-Based Model Libraries," Ph.D. Dissertation, Institut für Thermodynamik, Technische Universität Braunschweig, Germany, 2008.

[11] Franke R., Casella, F., Otter M., Sielemann M., Elmqvist H., Mattson S. and Olsson H., "Stream Connectors – An Extension of Modelica for

Device-Oriented Modeling of Convective Transport Phenomena," *Proceedings 7th Modelica Conference*, Como, 2009, pp. 122-131.

[12] Vredenborg, E., Lüdders H. and Thielecke F., "Methodik zur Auslegung und Simulation komplexer Brennstoffzellensysteme," *Deutscher Luft- und Raumfahrtkongress*, Hamburg, 2010.

[13] Colonna P., van der Stelt T.P., "FluidProp: A software for the calculation of thermopysical properties of fluids," Energy Technology Section, Delft University of Technology, 2004.

[14] Lemmon, E. W., Huber, M., McLinden, M., "NIST Standard Reference Database23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 8." National Institute of Standards and Technology, Standard Reference Data Program, Gaithersburg, 2007.

[15] Koeppen, C., Carl U., "Erfassung und Bewertung von Systemen im Flugzeugentwurf," *Deutscher Luft- und Raumfahrtkongress*, Stuttgart, 2002.

[16] U. S. Departement of Energy, "Hydrogen, Fuel Cells & Infrastructure Programm" Planned program activities for 2005-2015, 2009.

Hybrid Fuel Cell-based emergency power architecture for commercial aircrafts

M. M. Prats*, C. Lavoie**

** M. M. Prats, Maria Ángeles Martín Prats, Senior IEEE*
Electronics Engineering Department. Escuela Técnica Superior de Ingeniería.
Avda/ Camino de los descubrimientos s/n
41092-Sevilla. Spain
Telf: 34 954486177
Fax: 34 954487373

***C. Lavoie, Claude Lavoie*
Principal Engineering Specialist at Bombardier Aerospace
400, ch de la cote-Vertu, Dorval, QC H4S 1Y9, Canada
+1 514-855-5000

Keywords: *fuel cell power plant, emergency power systems, hybrid systems*

Abstract

In this work, an emergency electrical system has been designed and analyzed by using low temperature fuel cells, Proton Exchange Membrane Fuel Cells (PEMFC), batteries and ultracapacitors as emergency power system. In the complete design of the power plant, the power electronics converters for conditioning the electrical power are designed in order to increase the efficiency and the performance of the system. The technical requirement has been specified by Bombardier Aerospace (Montreal,) where the required power in each phase of flight is fixed during an emergency situation.
This work presents the design of the power emergency plant on board. The components and the technical requirements of the new system and the electronic control designed to assure the good performance of the management power system are analyzed.

1 Introduction

Nowadays, new trends in supply power architectures for on-board aircraft applications are focused on the integration of various energy sources. Indeed, the performance range for aerospace applications is pretty wide, going from several W up to tens of KW [1][2]. Fuel Cells (FC) energy sources have the potential to meet these supply power requirements. At this point, aeronautic industry is one driver of industrialization and early application of FC-based systems. [3][4][5]

Considerable efforts have been spent in order to optimize battery cells as well as to reduce maintenance and increment their lifetime. Commercial and military aircraft use a variety of techniques for providing backup electrical power to critical subsystems during emergency situations. Depending on the aircraft, these may include dedicated battery power, in-flight operation of the Auxiliary Power Unit (APU) [6], a Ram Air Turbine (RAT), Air Driven

Generators (ADG) or other technologies. However, hybrid FC-battery-ultracapacitor (UC) energy storage systems are promising candidates for stand-alone systems. This kind of systems combines the advantages of batteries, ultracapacitors (UC) and fuel cells (FC) systems. The main idea is to use a fuel cell for providing the required nominal power during an emergency system and include complementary energy storage system involving batteries and UC for supplying the needed power in case of transient power peaks. Batteries have higher specific energy than UC. As a result, they can provide extra power for a longer period of time. On the other hand, UC generally have higher specific power than batteries during a short period of time. Moreover, UC are more efficient and have a longer lifetime in terms of number of charge/discharge cycles. The using of UC extend the battery lifetime thanks to their longer lifetime in terms of charge and discharge cycles. Hence, it is clear that UC can be used more often to provide fast and large power transients. This UC skill has been largely employed in the automotive field.

This research is focus on the use of PEMFC and hydrogen for assuring a safety landing of the aircraft when the Ram Air Turbine of a generic aircraft of Bombardier is replaced by a hybrid system which involves fuel cell, batteries and ultracapacitors. The scenario is evaluated for studying the viability of replacing the current system by a new emergency power system based on fuel cells.

2. Study and description of the power requirements

This work considers an electrical emergency condition without main generation system; i.e. one engine failure plus an opposite generator failure plus no APU. In this scenario, ADG supplies power for all AC and DC (through Transformer Rectifier Unit, TRU) essential loads in order to continue a safe flight and to perform a safe emergency landing. This power system evaluates the viability of replacing ADG and RAT by a new hybrid FC-based system considering future demands of More Electrical and All Electrical Aircraft.

2.1. Definition of requirements

In order to obtain the complete design of the emergency power plant, the following considerations must be analysed:

- Definition of the power requirements. General power demand profile for the different phases of flight during an emergency situation when all normal electrical power has been lost.

- Weight, size, volume and cost of a new ADG (Air Driven Generator) system to comply with the requirements.

- PEMFC as the power source shall be evaluated:
    - Electrical demand for different emergency conditions (30, 60, 90, 180 minutes).
    - Quantity of fuel and oxidant, weight of different storage and supply system.
    - Improve the system performance at altitude.
    - Size and weight of PEMFC.
    - Design of the power electronics converters and electronics control.
    - Lithium-Ions battery and/or ultracapacitors for high electrical transient demands.
    - Weight.

- Proposal for complete stand-alone system integrating. Weight, volume and cost estimated of the proposed system.

Military and commercial aircrafts use different kind of techniques for providing backup electrical power to critical subsystems

during emergency situations. Depending on the aircraft, these may include dedicated battery power, in-flight operation of the auxiliary power unit, a ram air turbine (RAT), air driven generators (ADG) or other technologies.

Aeronautic industry is one driver of industrialization and early application of fuel cell systems. Although the performance range for aerospace applications is very wide (from several W up to tens of KW), fuel cells have the potential to meet these requirements [4][5]. For this reason, fuel cell technology is used in aeronautics fuel cell Projects in the field of energy conversion systems on board aircrafts. In addition, the use of fuel cells in this kind of applications is justified since they are a very important alternative power source due to their well-known specific advantages such as very low toxic emissions, low noise and vibrations, modular design, high efficiency (especially with partial load), easy installation, compatibility with a lot of types of fuels and low maintenance cost.
On the other hand, the fuel cell power converter must fulfil the requirements of the fuel cell to obtaining high efficiency and a robust performance [7][8]. Although power converters have been studied for many years, however, fuel cells power converters must be designed for each application. Fuel cells have a particular operation, due mainly to their low output voltage with high output current and restrictive specification in current ripple [8]. These are the reasons why a fuel cell power converter must be designed fulfilling each specific technical characteristics.

Hydrogen-powered aircraft are being sought to help reduce greenhouse gases from aviation. Fuel sources are compressed hydrogen gas and cryogenic liquid hydrogen [9]. Both of these fuel sources provide pure, contaminant-free hydrogen for the PEM cells.

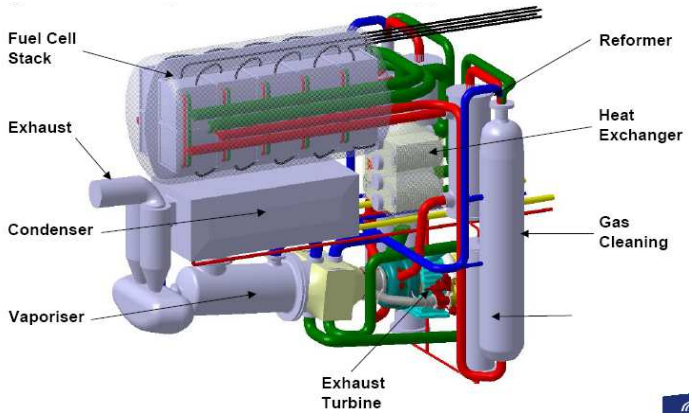


Fig. 1 System concept based on fuel cell for aeronautical application

For the next aircraft generation the pneumatic and hydraulic systems will be replaced step-by-step by electrical systems. The goal for the future is to find a highly efficient primary electric power source. Fuel cell systems have the potential to be a highly efficient system with a significantly reduction of fuel consumption and noise and gaseous emissions. In this work, the possibility of the application of fuel cell as emergency power supply is investigated.

Fuel cells also have a great advantage over battery because they eliminate charging time, allow a wide range of speeds, and operate as long as fuel is supplied.

2.2. Electric requirements

The electric system of the aircraft consists on generate, conditioning and distributing the power energy [10]. Nowadays is the area with higher potential of technology development due to mainly the increasing electric power, the decreasing of the fuel and of the weight of the on-board systems, power electronics with new solid state semiconductors, new magnetic materials and new technologies. The aircraft electric system is composed of generators, converters, transformers, DC-links, system control units, tests and ground systems, protections of the system, illuminating systems, heaters...

The electric system has specific requirements (operative, aero navigation, certification, ambient, vibrations, weight, volume, maintenance,…). It depends on the kind of the platform (military or civil aircraft, UAV, test bench).

The electric generation depends of:

- The voltage: AC (115/400Hz), AC (230/400 Hz), AC (115/WF), DC (28 V_{DC}), DC (270 V_{DC}),…
- Electric machine: IDG (CSD), VSCF, VFVV, VFCV, RAT, HVDC, S/G, …
- The function: primary generator, essential, emergency, auxiliary, of ground,…
- Platform: number of generators, power, refrigeration, reliability, etc,…

The kind of the electric conversion depends of the arquitecture: transformer (115 V_{AC} to 26 V_{AC}), TRU (115 V_{AC} to 28 V_{DC}), inverters (28V_{DC} to 115 V_{AC}), DC/DC converters for voltage regulation, new DC/DC converters for new technologies, batteries chargers.

The solid state technology is applied to the protection and control loads (SSPCs). It permits the controlled switching, the current limitation, remote control, reduction of weight and volume, the increase of reliability and modular and programmable systems.

The design of a electric device has into account the volume of the system, the dissipation capacity and the reliability that is required.

In general, the sources are designed considering the continuous consume and the maximum peak of consume for each phase of flight of the aircraft. The loads are defined taking criticality and operability into account.

The future trends are focus on the More Electric Aircraft (MEA) to All Electric Aircraft (AEA). This mainly consists on:

- The replacement of traditional mechanical and hydraulic systems by electrical, electromechanical or electro hydraulics systems.
- High voltage generation (270 V_{DC}, 350 V_{AC}, 600 V_{DC}) and consequently, the reduction of weight and volume with fuel saving.
- New magnetic materials integrated into the exe of the motor.
- Solid state technology for conditioning and distributing the power energy.

2.3. Technical specifications of the PEMFC, batteries and ultracapacitors for electrical demand profile for different emergency conditions

Once the load profile specifications of the aircraft are known, the design of the complete power system can be carry out.

The main idea is to use a fuel cell for providing the nominal power (12-14 kW aprox.) of the emergency system and include complementary energy storage system like as batteries and ultra capacitors of several kW for supplying the needed power in case of transient power peaks. In this case, the specifications fix 20-25 kW power peak.
It is necessary to develop a precise electronic control to ensure the good performance of the power management system.

The selected model of fuel cell for this application could be: Mark9 SSL of Ballard, due to the technical specifications which fulfil the power requirements.
PEMFC provides the nominal power needed by the critical system when the total aircraft power

is loss. The PEMFC electrical specifications are the following:

- Rated power is 14 kW
- DC voltage (at 300 A): 48 V
- Maximum current: 300 A
- Mass (with no coolant): 13 Kg
- Volume stack core: 228 mm lengh x 760 mm width x 60 mm height

The rated power of the selected batteries depends on the load profile of the aircraft. Batteries are used for providing power in transient peaks superior than a specified quantity. They operate when the critical loads need a continuous power, while the ultracapacitors only are used in the case of noticeable peaks of power in a reduced time period. The rating of the batteries of Global Express in Amperes hours (Ah) is: one avionics battery at 25 Ah to support the avionics load, including equipment that requires power direct from the battery, during APU starts, and another one, APU battery. This is rated at 42 Ah to start APU. The current rated power of the battery of a A330 aircraft is of 3 kW using a RAT of 25 kW and a APU of 550 kW (ground).

In this research work, we can select the Li-Ion 24 V CA244X Series battery which provides 44 Ah with a nominal voltage of 26 V. This battery uses patented nanophosphate technology. It is designed for minimal maintenance and presents half the weight of NiCd batteries.

For this application, for example: MC BMOD Energy Series 16v BOOSTCAP® Ultracapacitor Modules can be used.

The cells used in the modules operate at 2.7 volts, enabling them to store more energy and deliver more power per unit volume than any other commercially available ultracapacitor. In addition, to meeting or exceeding demanding industrial application requirements for both watt-hours of energy storage and watts of power delivery per kilogram, all of these products will perform reliably for more than one million discharge-recharge cycles.

2.4 Design of the control system

Each component of the system needs to be controlled for assuring its suitable performance when it is integrated into the complete emergency power system. In addition, once each individual control is designed and implemented, a management control of the system, which involves all the components, is designed in order to maintain the operation conditions in the range of the technical requirements and applicable normative.

The power converters will need a system for controlling the energy flow that they must provide to the receptor system according to its needs.
The data acquisition system will measure the main control variables used in the design and the study of the system (temperature, voltage, current,).
The control and acquisition system is based on a Digital Signal Processor (DSP) [Carrasco, 1994][Todorovic, 2004]. This device is programmed according to the established control in the circuit of the converter and to the study of its variables.
The control system is in charge of controlling the switching of the state solid power semiconductors. This system also protects the circuit. DSP is programmed for the protections of under/over voltage, short circuit, over current, over temperature and other critical situations that can damage to the system.

For controlling the system we will have into account that there are variables which must be always constantly controlled.

In the proposal system, the main electrical variables to be controlled are the following:

- The output current of the fuel cell, due to the current ripple must be into a

established range for avoiding an irreversible damage to the fuel cell.
- The output voltage of the fuel cell.
- The input/output current and input/output voltage of the battery charger.
- The output current and output voltage of the ultracapacitors.
- The output current and output voltage of the ultracapacitor DC/DC converter.
- The voltage and the current in the DC-link.
- The input/output voltage and the input/output current of the 270 V_{DC} DC/DC converter, if the converter was necessary.
- The input/output voltage and the input/output current of the 230/400 3 phase V_{ac} DC/AC converter.

The fuel cell DC/DC converter has as the main objective to increase/decrease the output voltage of the fuel cell for connecting it to the DC-link voltage. If we consider the selected 14 kW PEMFC with 48 output voltage (300 A), we will decrease this voltage for permitting the connection to a 28 V_{DC} bus. However, if we increase the DC-link voltage to 42 VDC, we will regulate the output voltage with an very efficient DC/DC converter. It will be possible to design a DC/DC converter for the fuel cell which adapts its performance to the output voltage of the fuel cell and to the DC-link voltage.
The PWM modulation technique permits the on/off state of the switches controlling the current flow thought the system increasing the voltage of the capacitors. We regulate the voltage of the system controlling the current.

The current measurements of the loads and of each component permit to carry out the management of energy of the complete system, deciding when and for how long each sub-system is providing energy to the emergency system.

3. Proposal for complete stand-alone system

In case of total loss of primary and auxiliary AC power, non-time limited emergency AC power shall be provided by a conventional Ram Air Turbine. The main objective of the proposed topology is replacing the Ram Air Turbine (RAT) for a fuel cell power plant which could be used as an emergency power source for a general aircraft application. The proposal stand-alone system must fulfil the technical requirements explained in section 2.

The simplified scheme of the fuel cell power unit is showed in Figure 2:

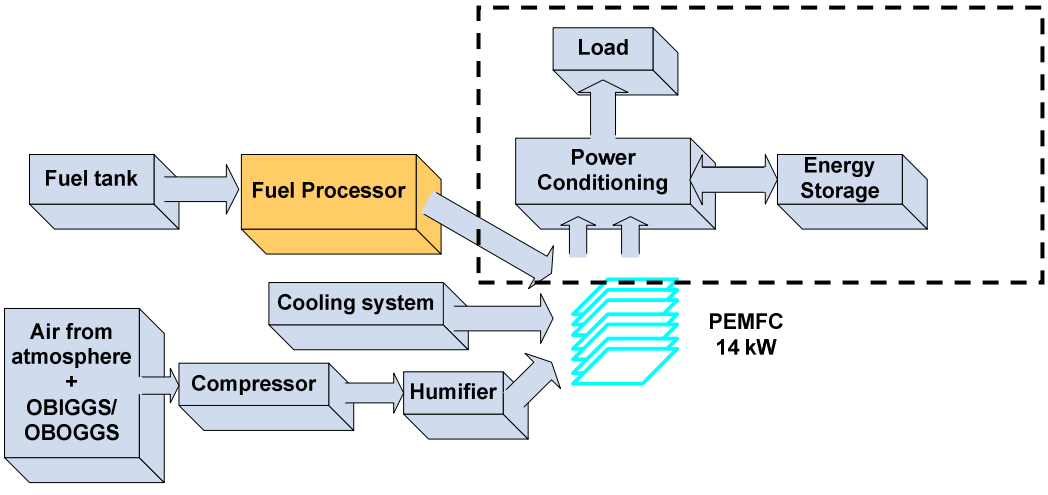


Fig. 2 Scheme of the fuel cell power unit

Where the sub-system into the dashed line is showed into detail in the following block diagram:

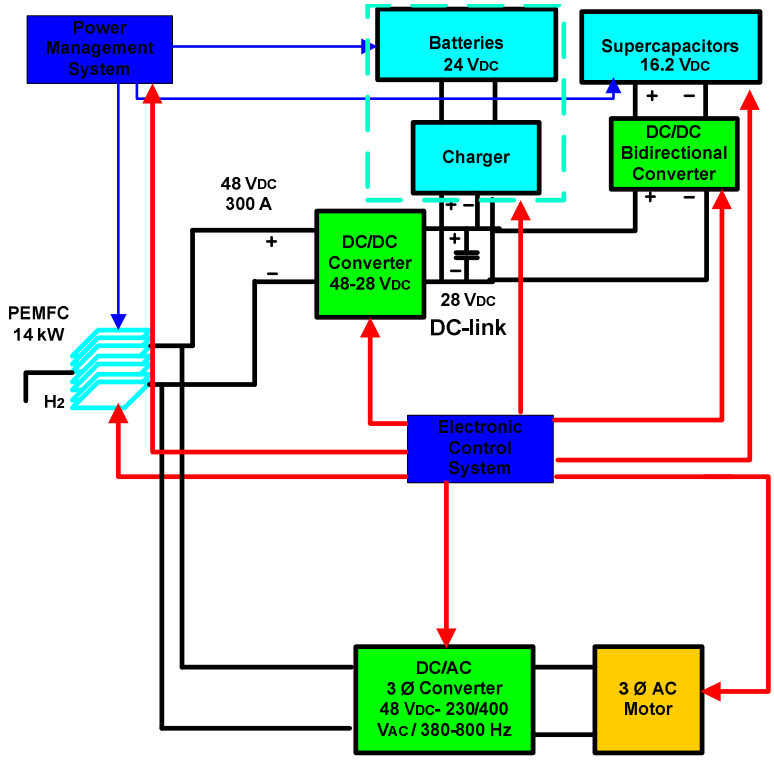


Fig. 3 General scheme of the proposed power topology

Due to the significant number of transient power peaks, the charge and the discharge and the efficiency of the batteries must be studied.

The ultracapacitors are used as secondary power sources of very fast response to transient situations, where the primary power sources (like batteries or fuel cells) are too slow.

3.1. System Integration

The activities relating to the system integration involve to:

- Motor and Drive
- Fuel Cell Systems:
  - Heat exchanger
  - Pumps
  - Compressor
  - Controller
- Li-Ion Battery
- Ultracapacitors
- OBIGGS (OBOGSS)
- Controllers and Electronics Power Converters.
- Power management and distribution.
- Storage, delivery and regulation system of hydrogen.

The general scheme of the components integration of this work is shown in figure 4:

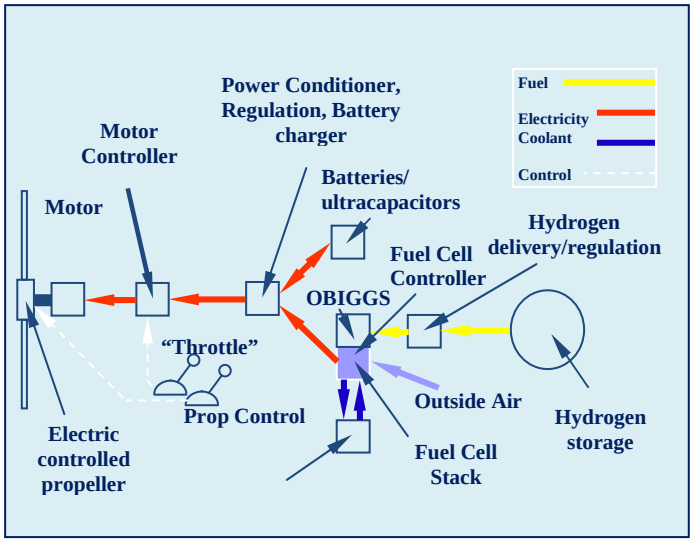


Fig. 4 System integration

A general management control system must be implemented in order to control the power balance in the complete emergency power system. We will consider the power demand profile of the critical loads for the different phases of flight, during an emergency situation when all normal electrical power has been lost. Also, motor, controllers, and power electronics are considered. The power generation system, composed by PEMFC, Li-Ion battery and/or ultracapacitors, will be regulated for providing satisfactorily the power demand in each moment.

Fuel cells provide high efficiency to produce electrical power compared to APU operation ground (75% fuel saving) and compared to main engine generators in cruise (40% fuel saving).

4. Conclusion

In this work, the design of a complete hybrid emergency electrical power system is analysed in order to replace conventional emergency power system in aircrafts. The current research working go towards development of fuel cell, high efficiency power converters, new Li-Ion batteries, new ultracapacitors, fuel processing, desulphurization, OBIGGS system for improving the performance at altitude and hydrogen generation solutions in order to provide more environmentally friendly power sources.

The advanced technology is oriented to reach the development of mature, safe and reliable fuel cell systems for future aircraft application. Fuel cells may dramatically cut emissions, noise and fuel use for aircraft. However, the weight and volume of hydrogen storage systems are presently too high, resulting in inadequate aircraft range. Materials and components are needed that allow compact, lightweight, hydrogen storage systems. Due to the current state of the technology, the use of fuel cells is only justified as emergency electrical power system, auxiliary power unit or for propulsion of light aircraft and small commuter aircraft.

The research work carried out is the starting point to join the efforts of companies and Universities interested for researching and developing a PEMFC emergency electrical power system as alternative to the conventional auxiliary power sources of the aircraft. This permits to improve the performance and the efficiency of the current emergency power system.

The new trends in More Electric Aircraft or All Electric Aircraft give special relevance to the design of the electronics power converters for assuring the good performance and high efficiency of the total system. An efficient power management will contribute to the increase of the system efficient and the increasing of the time of life of the involved components in the emergency system, including the fuel cell. The future MEA airplanes will be ideally suited for fuel cells.

References

[1] George A. "Fuel Cell Inverters for Utility Applications". *Power Electonics Specialist Conference*, Vol. 3, 2000, pp.1191-1194.

[2] Lipman T., "Fuel cell system economics: comparing the Costs of generating power with

stationary and motor vehicle PEM fuel cell systems", University of California, Berkeley, April 2004.

[3] Barbir and Gómez T., "Efficiency and economics of proton exchange membrane (PEM) FUEL CELLS". *Int. J. Hydrogen Energy* Vol. 21. No. 10, 2006, pp 891-901.

[4] Arco M.C., Prats M.M., Sanchez J., Galván E., Carrasco J.M., Moreno E. "2kW fuel cell power conditioning designs stack for isolated system", *2th European Hydrogen Energy Conference*, 2005, pp. 474-475.

[5] Prats M. M., Arco M.C., Sanchez J., Galván E.,.Carrasco J.M., Moreno E., "Power conditioning system for maritime fuel cell applications using a 10kW PEMFC as auxiliary system", *2th European Hydrogen Energy Conference, 2005*, pp. 471-472.

[6] Aicher T., Lenz a B., Gschnell F., Groos U., Federici F., Caprile L., Parodi L. "Fuel processors for fuel cell APU applications". *Journal of Power Sources*, 2006, pp. 503–508.

[7] Carrasco J. M., Franquelo L. G., Bialasiewicz J.T., Galván E., Portillo R., Prats M. M., León J.I., and Moreno N., "Power-electronic systems for the grid integration of renewable energy sources: a survey", *IEEE Transactions on Industrial Electronics*, Vol. 53, No. 4, 2006, pp. 1002 – 1016.

[8] Bates J. "Ripple Reduction for High Power Fuel Cell Converter Application". *Power Electronics Conference,PESC*, Vol. 34, 2001.

[9] Biris A. S., Biris A. R., Lupu D., Buzatu D., Darsey J., and Muzumder M. K., "Use of carbon nanostructures for hydrogen storage for environmentally safe automotive applications", *Proc. Industry Applications Conference*, Vol. 2, 2004, pp. 953 – 956.

[10] Masson P.J., Soban D. S., Upton E., Pienkos J. E., and Luongo C. A., "HTS Motors in Aircraft Propulsion: Design Considerations". *IEEE Transactions On Applied Superconductivity*, Vol. 15, No. 2, 2005.

CEAS 2011

Methods for Preliminary Design

Development of a Software for Aircraft Preliminary Design and Analysis

Fabrizio Nicolosi and Giuseppe Paduano
Department of Aerospace Engineering(DIAS) - University of Naples "Federico II"
Via Claudio 21, 80125 Napoli - ITALY
E-mail: fabrnico@unina.it
Ph: +39 0817683583 Fax : +39 081 624609

Keywords: *Aircraft preliminary design, Educational software, Aircraft Performances*

Abstract

The present paper deals with the description of a software called ADAS (Aircraft Design and Analysis Software) that has been developed at DIAS (Dep. of Aerospace Eng.) by the authors during the last 5 years.

The paper gives an overview of all main modules and used procedures and approaches that have been built and implemented in the software. ADAS code has been originally intended as an educational tool, but after many improvements carried out during the past years, can be used as a research tool and it can be of some relevance also for some industrial applications.

All the aerodynamic characteristics used for stability, control and aircraft performances calculations are estimated mainly through semi-empirical laws integrated with more sophisticated approaches like the use of aerodynamic tools like Multhopp or Weissinger to estimate wing lift spanwise distribution. The software is extremely user-friendly and can represent a valid teaching tool for aircraft design courses.

1 Introduction

In the last years many attempt have been made in order to develop and deliver to the aerospace research and teaching community some software or tool that can be used with the aim to perform preliminary design of airplanes. One of the most known software, developed by J. Roskam(one of the main scientist and researcher in this field) and his collaborators, is the Advanced Aircraft Analysis (AAA) software (see http://www.darcorp.com/Software/AAA/), that is the industry standard aircraft design, stability, and control analysis software. AAA is installed in over 55 countries and is used by major aeronautical engineering universities, aircraft manufacturers, and military organizations worldwide. The AAA software is based on the well known aircraft design book written by Roskam, i.e. [1],[2], [3], that presents several semi-empirical methodologies and fast approaches for aircraft preliminary sizing and aerodynamics analysis, mainly derived and copied by USAF Datcom [4]. During the last ten years many other software have been written and also some very well known books have been introduced in the attempt to allow to many people, often without the deep and solid experience of an old aircraft designer, the

possibility to carry out the preliminary design of an aircraft, mainly at teaching level, but sometimes also with some features typical of industrial tools. Among them it is worth to point out the book proposed by D. Raymer [5], famous aircraft designer, and the software[6] "RDS, Integrated Aircraft Design and Analysis"(http://www.aircraftdesign.com/rds.shtml) that is also widely used in the aerospace academics for aircraft design teaching. Another interesting effort in this direction has been made by Prof. I. Kroo from Stanford University. The digital textbook and the associated examples of calculations made by several java applets (see http://adg.stanford.edu/aa241/AircraftDesign.html) represents a good example of a very effective teaching tool that allow students the comprehension of aircraft design theory and the possibility to "play" and see the results of parametric analysis. The ADW software (see http://www.desktop.aero/adw.php) distributed by Desktop Aeronautics is an example of extremely user-friendly application for aircraft design activities, performed at teaching level. The ADW program was developed as an interactive museum exhibit that allows the user to investigate the trade-offs involved in the design of a commercial transport aircraft.
Another software, also developed by I. Kroo and Desktop aeronautics is PASS(see http://www.desktop.aero/pass.php). The software PASS is a conceptual design tool that is capable of evaluating all aspects of mission performance. Rapid analyses are coupled with optimization tools to quickly achieve realistic aircraft designs. Configuration design is not limited to meeting a singular performance goal. PASS ensures that designs satisfy a wide range of appropriate, real-world constraints such as FAA field length, climb gradient, and community noise requirements. PASS was developed for general applicability to conventional aircraft, but has been customized to do analysis of non-conventional aircraft such as supersonic business jets, oblique flying wings and blended wing bodies. The software is an advanced tool and maybe represents a complex tool to be used as teaching application in aircraft design courses. Other examples of aircraft design software often used at industrial level are ACSYNT software (http://www.phoenix-int.com), PIANO software (http://www.piano.aero/) developed in UK by Lissys Limited, the ADS software (http://www.pca2000.com) developed by OAD SPrl in Belgium and used by several general aviation companies. All the last examples represents some complex software that becomes extremely hard to use for non-expert students, especially if the aircraft design course does not allow (due to short period of lectures, i.e. 80 lecture hours in 2-3 months) the students to go deeply in some specific aerodynamic problems and to have a deep comprehension of the problems.
The development of the software by the authors at the Department of Aerospace Engineering of University of Naples started in 2005 as an effort to give some support to the students in the development of aircraft design exercises carried out by students in the aircraft design course. Other authors like Martinez-Val [7] and T. Young [8] have recently emphasized the importance of aircraft design course as a fundamental milestone in an aerospace student curriculum (to develop synthesis attitude and to build multidisciplinary links). It is evident that such a course can be of relevant effectiveness only with some practical exercises or better with the development by the students of the complete preliminary design of an aircraft.

2 ADAS Software structure

ADAS software has been originally intended as an educational tool, but after many improvements carried out during the past years, can be absolutely used as a research tool and it can be of some relevance also for some industrial applications. The approach to estimate all aircraft characteristics has been the use of classical semi-empirical laws (like those one proposed by Roskam [1-3] or ESDU or USAF Datcom[4]) integrated with more sophisticated approaches like the use of aerodynamic tools like Multhopp or Weissinger to estimate wing lift distribution along span and wing stall path.

One of the main goal that was fixed in the development of ADAS software was to obtain a very user friendly platform that should be easily used also by students, with many graphical helps and with many output graphs and figures, including 2D and 3D drawing of aircraft parts under development.

The software is split in modules. In fig. 1 the main software dialog window (main menu) is shown. The classical approach and first part of the software starts with the preliminary estimation of aircraft weight (both empty and maximum take-off) and required wing area and engine thrust/power under specified design requirements [1]. In this phase all FAR/CS 23/25 performance requirements have been automatically implemented based on aircraft weight. After the preliminary phase the user starts designing all aircraft parts. The main dialog window allow the user to enter and carry out the design of each airplane part (fuselage, , nacelle, wing, aileron and flap, taiplanes). Each module (fig. 1) can be used as independent module, but is linked with the other modules if an airplane design activity is developed. Figure 2 shows the classical flowchart of design/analysis steps that the designer can carry out in sequence. The drag polar and main flight/ground performance estimation can be accomplished also with a first preliminary sizing of tailplanes (fig. 2). The tailplane design and analysis modules allow a detailed and accurate design and analysis of horizontal and vertical taiplanes, performing also the aircraft stability and control analysis up to non-linear conditions. The wing module allows to design the wing and to obtain all wing aerodynamics characteristics, including lift, moment and drag. The estimation of spanwise lift coefficient distribution and wing stall path are also included and gives to the user the possibility to check the assumed wing twist and to make a preliminary estimation of wing bending moment. The wing module is linked with the fuselage module to estimate the body longitudinal instability contribution which depends on wing induced upwash and downwash. The wing analysis is also coupled with the modules for high-lift and aileron analysis. The aileron analysis and design module allow also the estimation of aircraft roll performances with the prescribed propulsive system. The software has one module dedicate to the accurate estimation of aircraft drag polar and a module for aircraft flight and ground performances. Other very important modules deal with aircraft tail sizing and aircraft longitudinal and lateral-directional characteristics. The modules are really user-friendly and allow a fast calculation of all required information. Finally the user can have the drawing of the designed aircraft. The software allow the preliminary design and analysis of a transport aircraft in less than 1 hour. The software has been recently also presented in previous conferences[9]. Other similar software are under development in other European Universities, like that one developed by D. Scholz [10] and this demonstrate the relevance and the importance of such applications, both for teaching and for some applied industrial case. Jenkinson has written an educational book [11] that put in evidence the importance for students of performing aircraft design activities and examples.

3 Description of software modules

As already outlined the software is organized in separate modules. Each module can be run separately, but the necessary link between modules are built to allow the user to perform the preliminary design of an aircraft. In this paragraph the paper present a general description of each module and each part of the software, trying to put in evidence the main characteristics and capabilities of the software.

3.1 Weight estimation

The first part of the software deals with the preliminary estimation of aircraft weight (both empty and maximum take-off). The classical approach suggested by Roskam[1] and Nicolai [12] is used. The design requirements must be fixed by the user. In particular the aircraft weight is estimated on the base of some assumed parameters (cruise aerodynamic efficiency, fuel consumption, propeller

efficiency) and especially on the base of the imposed and expected mission fundamental to calculate the required fuel fraction. The mission includes all the flight phases like take-off, climb, cruise and descent. In fig. 3a is shown the table that reminds to the user the assumed value for the imposed mission and the result of the two-equation system for weight calculations. Figure 3b shows the output graph of the software with all the aircraft weights (empty weight, gross weight, fuel weight, etc.) reported in different unit (kg and lb).

3.2 Aircraft preliminary sizing

One of the first modules is the general aircraft sizing, or the estimation of required wing area and engine thrust/power under specified design requirements [1]. In this phase all FAR/CS 23/25 performance requirements have been automatically implemented based on aircraft weight. This module allow the preliminary sizing of jet and propeller driven aircraft. The preliminary choice of wing loading and engine thrust to weight ratio (for a jet) is made on the base of design and regulation requirements and in particular take-off distance, landing distance, climb requirements and cruise requirements. The module is very fast and with user-friendly interface. An example of the final graph for preliminary sizing of a 150 seats jet transport aircraft is shown in fig. 4. The graph reports all the limitations imposed both by the design specifications and by the regulation requirements. The user can also add some point representative of existing aircraft (i.e. the Airbus A320) and can make its choice.

3.3 Fuselage design

After the preliminary phase the user starts designing all aircraft parts.

One of the first component that need to be defined is the aircraft fuselage. The fuselage is designed with the classical approach from the inside to outside. The fuselage module allows the design of the fuselage, starting with cabin interior required space up to the final shape in which the sections can be modified to account for classical squared-like sections. The section shape can be modified directly by the user with the mouse and using classical curves like nurbs or splines as weel described in [5]. The internal dimensions are calculated on the base of the assumed seat pitch and seat width. Some important guidelines for this module have been taken from some representative aircraft design books [2, 3, 13, 14]. The user can play with this dimensions (that defines the internal comfort) and look at the effect on fuselage dimensions (length, width, fuselage wetted area) and fuselage aerodynamic characteristics(fuselage drag, body aerodynamic longitudinal instability). The fuselage module allows calculation of all fuselage geometrical characteristics (i.e. wetted area) and all aerodynamic characteristics, like fuselage moments and fuselage drag. The aerodynamic fuselage moments (both longitudinal and directional) are estimated with classical strip method proposed by Munk and with the implementation suggested by Multhopp[15]. This last approach and other semi-empirical alternative analysis [3, 16] have been also implemented. The software allows an easy and fast link between components, with very fast choice of wing-fuselage relative position with the real-time calculations of the effect of this parameter on fuselage negative stability and on aircraft cg position in percentage of mean aerodynamic chord. An example of some output and interface windows from fuselage design module are shown in fig. 5.

3.4 Wing Analysis and Design

The wing module allows to design the wing and to obtain all wing aerodynamics characteristics, including lift, moment and drag. The semi-empirical approach is based on formula and suggestions reported in [3,4,17,18]. The effect of fuselage on wing lift [17] and wing induced drag (see http://adg.stanford.edu/aa241/AircraftDesign.html) has been also taken into account. The aerodynamic calculations have been extended in non-linear condition and the wing maximum lift coefficient is estimated through both semi-

empirical and wing stall path approaches. In the wing module Multhopp (non swept wing) and Weissinger (swept wing) methods have been both implemented to estimate wing span load, wing stall path and to estimate also wing shear and bending moment due to aerodynamic loads. Many comparisons with available results and literature have been done and the code seems to give reasonable and reliable results. Figure 6 shows the wing design dialog window. All main geometrical characteristics (mean aerodynamic chord, aerodynamic centre) and aerodynamic parameters are estimated. Figure 7 shows some results concerning the aerodynamic analysis and some output graphs showing relevant aerodynamic data. The wing module is coupled with the high-lift module where all the aerodynamics (mainly lift and pitching moment) of wing with high-lift device are estimated. The module is very easy and allows the design (in terms of choice of required flap chord ratio and flap span extension for required lift increment).

In figure 8 the high-lift system design dialog window is shown. The user can define all main flap and slat geometrical characteristics (flap chord, flap span extension, flap type and flap deflection) and can immediately estimate the aerodynamic effect on the wing lift curve. The semi-empirical approach that has been used is based on effect of high-lift system on [14].

Figure 9 shows the aileron design dialog window. The user can define aileron geometrical data and the software can estimate in real time the aircraft roll capabilities, also according to the chosen propulsion system characteristics.

3.5 Aircraft Performances

A very important software module is the performance analysis module. This module allow the calculation of aircraft flight and ground performances based on the aircraft drag polar. The wave drag (compressibility effects) is taken into account and the user can chose between different airfoil type (from peaky up to aggressive supercritical airfoil) and estimate different wave drag aerodynamic effects on aircraft drag. The user can chose different propulsion system(piston engine, turboprop engine, turbofan with low by-pass-ratio and high by-pass-ratio). A very deep analysis has been made in order to built these different engine models. Many indications have been taken from [19, 20] and [21] and from the Aircraft Design digital textbook available from DesktopAeronautics(http://adg.stanford.edu/aa241/AircraftDesign.html). Some examples of aircraft flight performances calculations of a typical jet transport are shown in fig. 10.

3.6 Tailplane design. Stability and control

As well known in all aircraft design textbook the tailplane design and associated stability and control(both longitudinal and directional) characteristics of an aircraft are crucial elements for the quality of the designed aircraft. ADAS software has some modules dedicated to the preliminary design of tailplanes and also for a deep stability and control analysis which can take into account also non-linear effects (like non-linearity of lift curve slope, non linearity of wing downwash, pendular stability). Figure 11 shows the main horizontal stabilizer design dialog window. The user through the definition (made very fast by the use of interface boxes and sliders) of some assumed characteristics (tailplane position, tailplane sweep angle, taper ratio) and with the calculated characteristics of the wing-body combination, can estimate some relevant limit curves in a carpet plot showing stability and longitudinal control limit. Accurate estimation of wing downwash angle[22] extended in non-linear conditions has been implemented. After a preliminary design, the user can also investigate the aircraft longitudinal characteristics including also non-linear effects. The vertical stabilizer is designed mainly with the constraint of minimum control speed for multi-engine aircraft. The vertical tailplane aerodynamic lift can be estimated both with the approach suggested by Roskam [3] and USAF Datcom [4] and with the alternative approach suggested by ESDU [23]. The two methodologies often lead to different results and the user, based on the acquired experience should use the most suitable method. Figure 12

shows the vertical tailplane design dialog window for a medium-range transport aircraft. The user has to assume some geometrical characteristics and the module helps the user in finding the best vertical tail area.

4 Drawings and export capabilities

In the software some drawing features have been implemented. The final designed aircraft shape can be also exported in general CAD format. Figure 13 shows the final aircraft layout and 3-view drawings as produced by a dedicated module in the software.

5 Conclusion

The paper has presented some capabilities and features of ADAS software. The importance of such tool for teaching and for fast estimations has been highlighted. The importance of a very user-friendly interface for such tools is crucial. An important extension and improvements for the code is the inclusion of some automatic optimization procedure to help the designer with some computer-guided decision strategies.

References

[1] Roskam J., *Airplane Design: Part I, Preliminary Sizing of Airplanes*, 1st Ed., Roskam Aviation and Engineering Corporation Ottawa, Kansas, 1985.

[2] Roskam J., *Airplane Design: Part III, Layout Design of Cockpit, Fusolage, Wing and Empennage:Cutaways and Imboard Profiles*, 1st Ed., Roskam Aviation and Engineering Corporation Ottawa Kansas, 1985.

[3] Roskam J., *Airplane Design: Part VI, Preliminary Calculation of Aerodynamic, Thrust and Power Characteristics*, 1st Ed., Roskam Aviation and Engineering Corporation Ottawa Kansas, 1985

[4] Hoak, D.E., et al, USAF Stability and Control Datcom, Flight Control Division, Air Force Flight Dynamics Laboratory, WPAFB, Ohio, 45433-0000, 1978, revised.

[5] Raymer D.P., *Aircraft Design: A Conceptual Approach*, 1st Ed., American Institute of Aeronautics and Astronautics Washington D.C., 1999, Chaps. 12, 13.

[6] Raymer D P., *RDS-STUDENT: Software for aircraft design, sizing, and performance*. American Institute of Aeronautics and Astronautics Inc., Reston, 1999.

[7] Martinez-Val R., Perez E., "Airplane Design: A must in Aeronautical Engineering Education," ICAS 2008 – 26th ICAS Congress, Anchorage, (Alaska, US), September 2008.

[8] Young T M. Aircraft design education at universities: benefits and difficulties. *Aircraft Design*, Vol. 3, No. 3, pp. 207-215, 2000.

[9] Nicolosi F., Paduano G., "Development of a Software for Aircraft Preliminary Design and Analysis," 10th European Workshop on Aircraft Design Education (EWADE), Dipartimento di Ingegneria Aerospaziale – University of Naples, Naples, May 2011

[10] Scholz D., "PreSTo The Aircraft Preliminary Sizing Tool," 10th European Workshop on Aircraft Design Education (EWADE), Dipartimento di Ingegneria Aerospaziale – University of Naples, Naples, May 2011

[11] Jenkinson L.R. and Marchman J.F. III, *Aircraft design projects: for engineering students*. American Institute of Aeronautics and Astronautics Inc., Reston, 2003.

[12] Nicolai, L.M., *Fundamentals of Aircraft Design*, METS inc., 6520 Kingsland Court, San Jose, CA, 95120.

[13] Jenkinson L.R., Simpkin P., Rhodes D., *Civil Jet Aircraft Design*, 1st Ed., Arnold, a member of the Hodder Headline Group London, 1999

[14] Torenbeek E., *Synthesis of Subsonic Airplane Design*, 1st Ed., Delft University Press, 1982

[15] Multhopp H., "Aerodynamics of the fuselage," NACA TM 1036, Washington, Dec. 1942.

[16] Perkins C.D., Hage R.E., *Synthesis of Airplane Performance, Stability and Control*, 1st Ed., John Wiley & Sons Inc., New York,1976

[17] Anderson Jr J.D., *Aircraft Performance and Design*, 1st Ed., McGraw-Hil, New York, 1999.

[18] Abbott I.H., Von Doenhoff A.E., *Theory of Wing Sections*, Dover Pubblications New York, 1959..

[19] Roskam J., Lan C.T.E., *Airplan Aerodynamics and Performance*, 1st Ed., DARcorporation Lawrence Kansas, 1997.

[20] McCormick B.W., *Aerodynamics, Aeronautics and Flight Mechanics*, 1st Ed., John Wiley & Sons Inc., New York New York, 1979.

[21] Hartman E.P., Bierman D., "Report No 640, The Aerodynamic Characteristics of Full Scale Propeller Having 2, 3 and 4 Blades of Clark Y and A.F. 6 Airfoil Sections" NACA, 1938.

[22] Hartman E.P., Bierman D., "Report No 648, Design Chart for Predicting Downwash Angles and Wake Characteristics Behind Plain and Flapped Wings." NACA, 1939.

[23] ESDU 82010, "Contribution of Fin to Sideforce, Yawing Moment and Rolling Moment Derivatives Due to Sideslip, in the Presence of Body, Wing and Tailplane.".

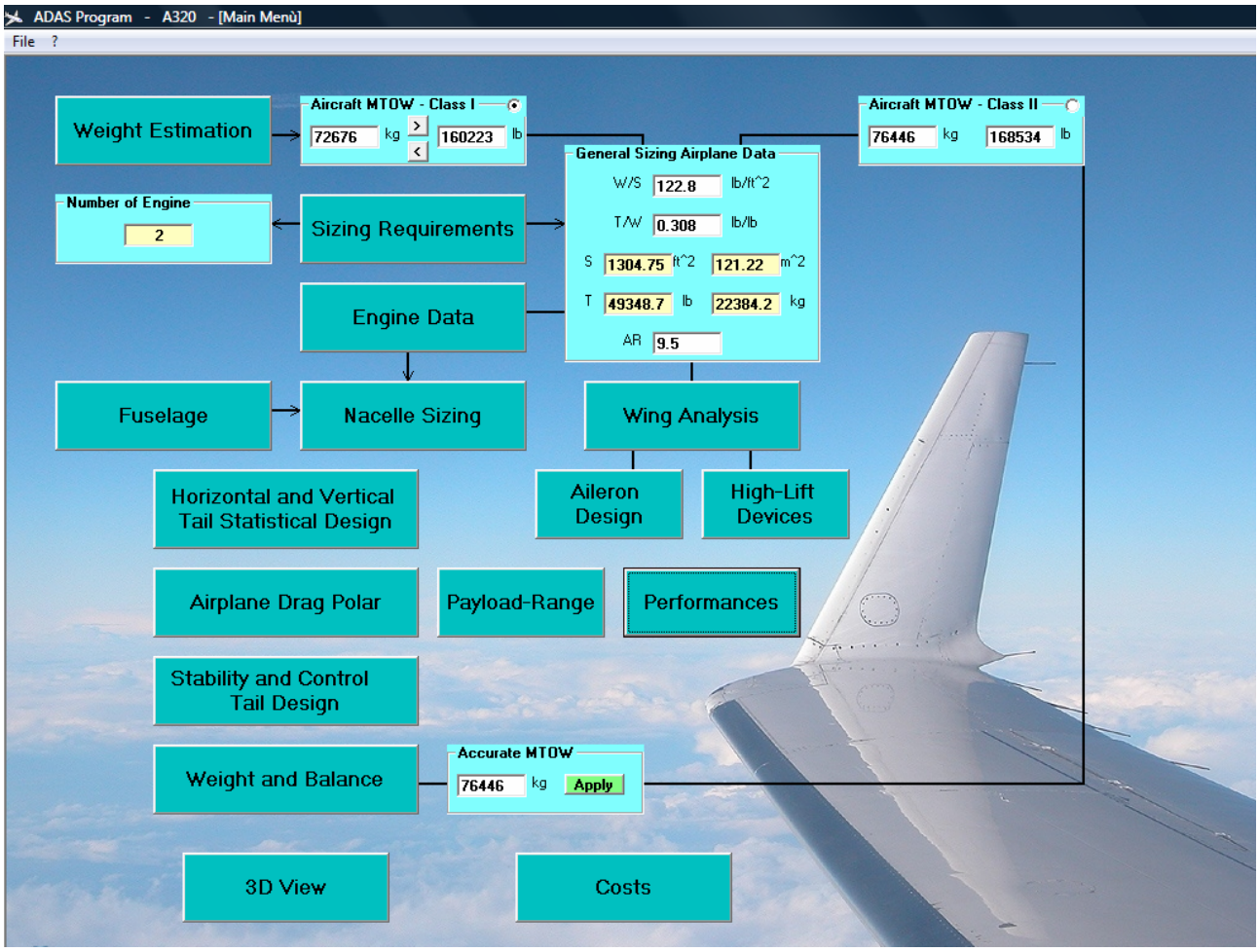


Fig. 1 ADAS Software Main Dialog Window (main menu) and displayed design/analysis modules

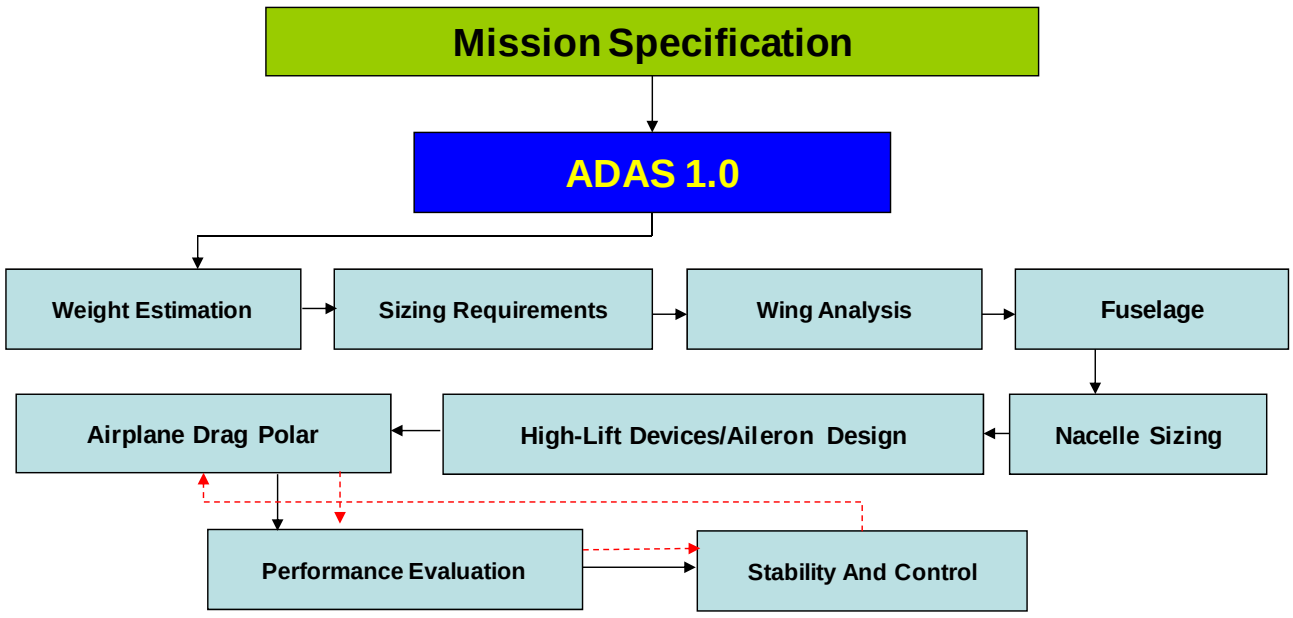


Fig. 2 ADAS Software Analysis and Design modules and flowchart of possible design sequence

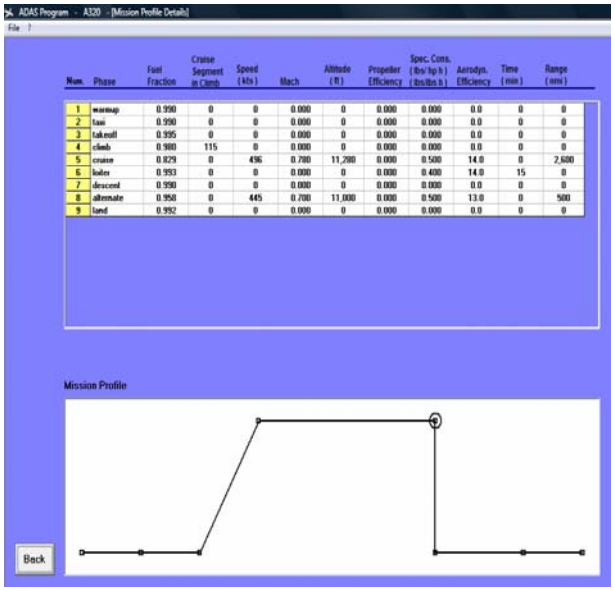


Fig. 3a Weight estimation module. Flight mission data table

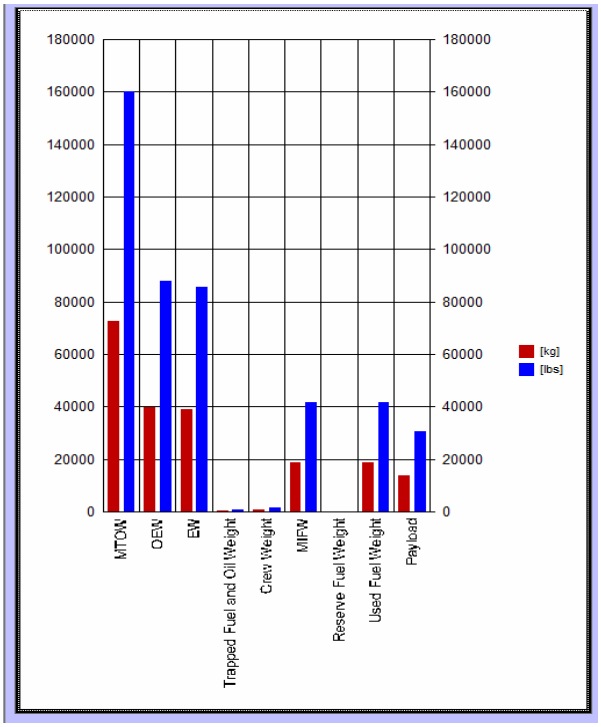


Fig. 3b Weight estimation results. Weight breakdown in different unit.

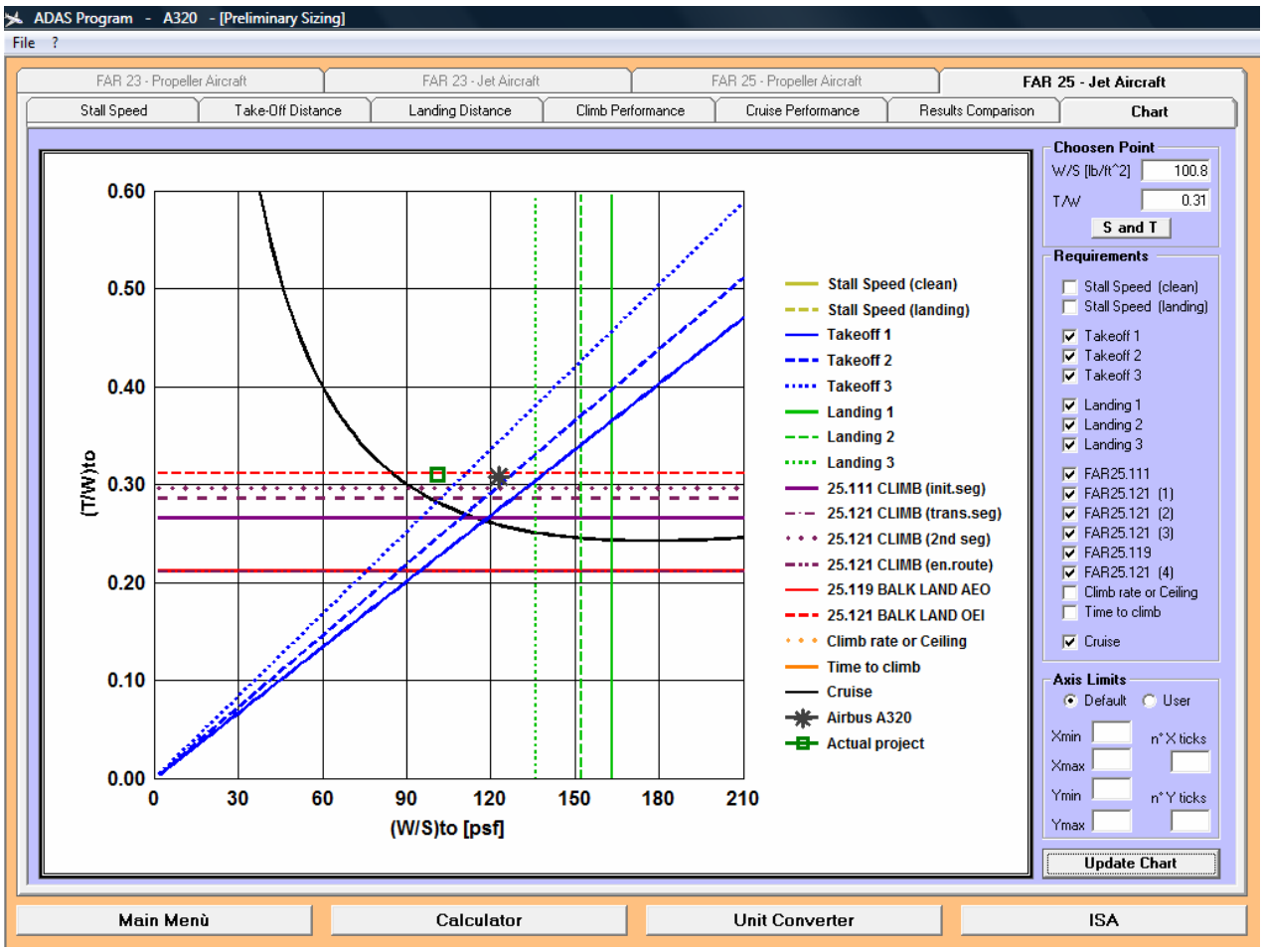


Fig. 4 Example of final graph of Preliminary Sizing Module (150 seats medium range transport jet)

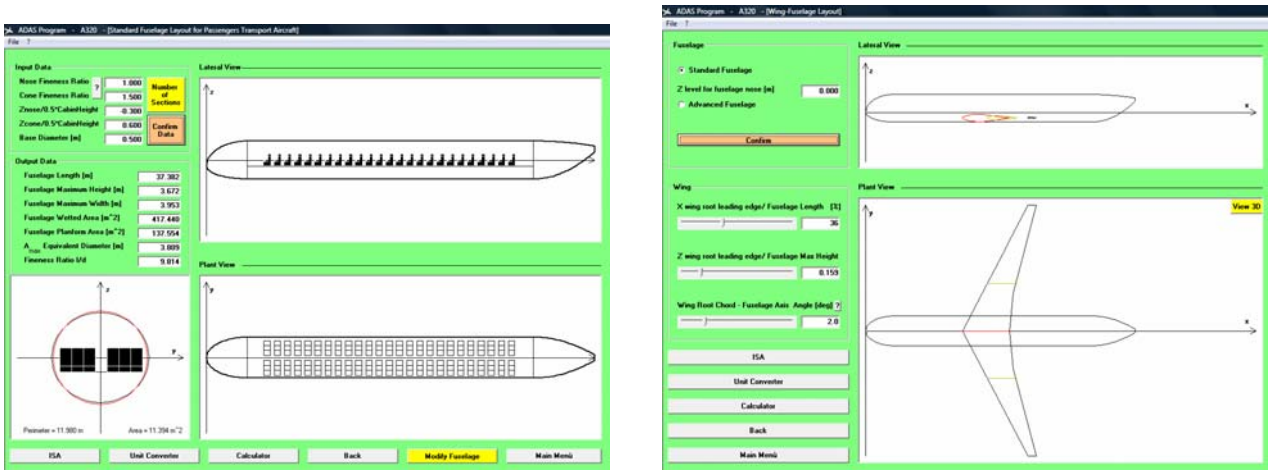


Fig. 5 Example of interface windows of fuselage design module. External dimension and general fuselage characteristics (left) and wing-fuselage integration (example of a 150 seats medium range transport jet)

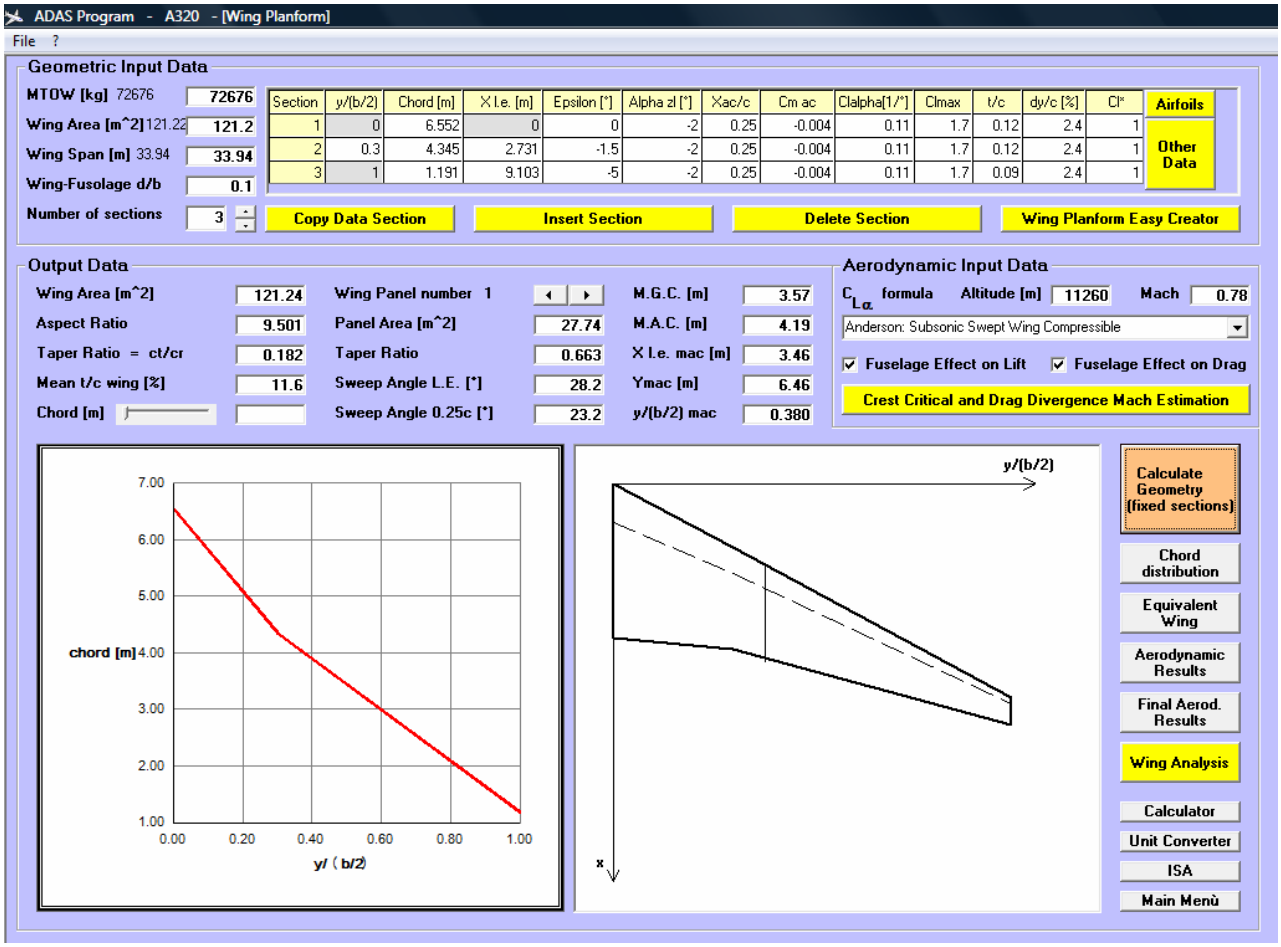


Fig. 6 Wing design dialog window. Definition of wing geometrical and main airfoil characteristics.

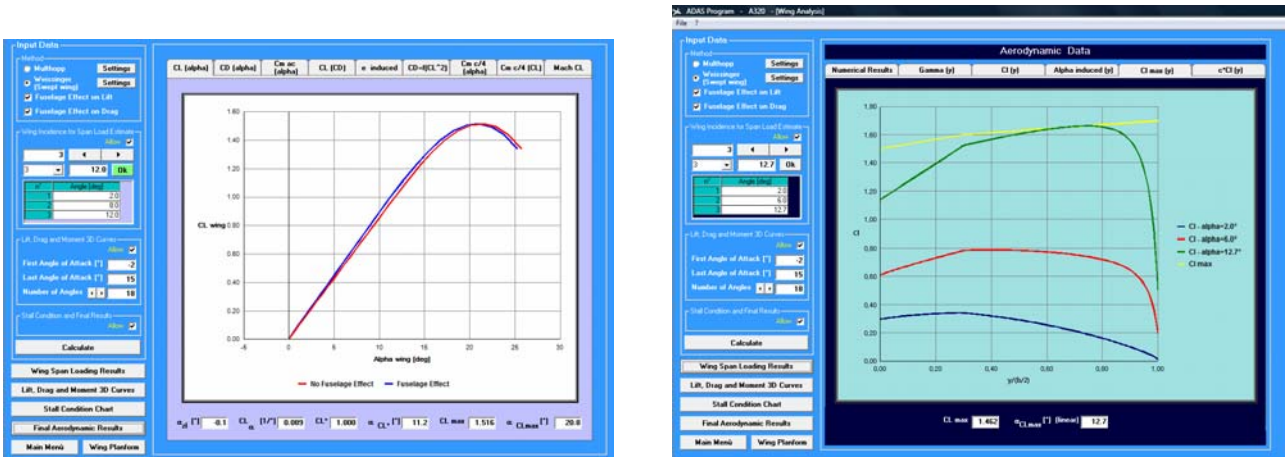


Fig. 7 Wing aerodynamic analysis results.
Lift curve (left) and lift coefficient spanwise distribution and stall path (right)

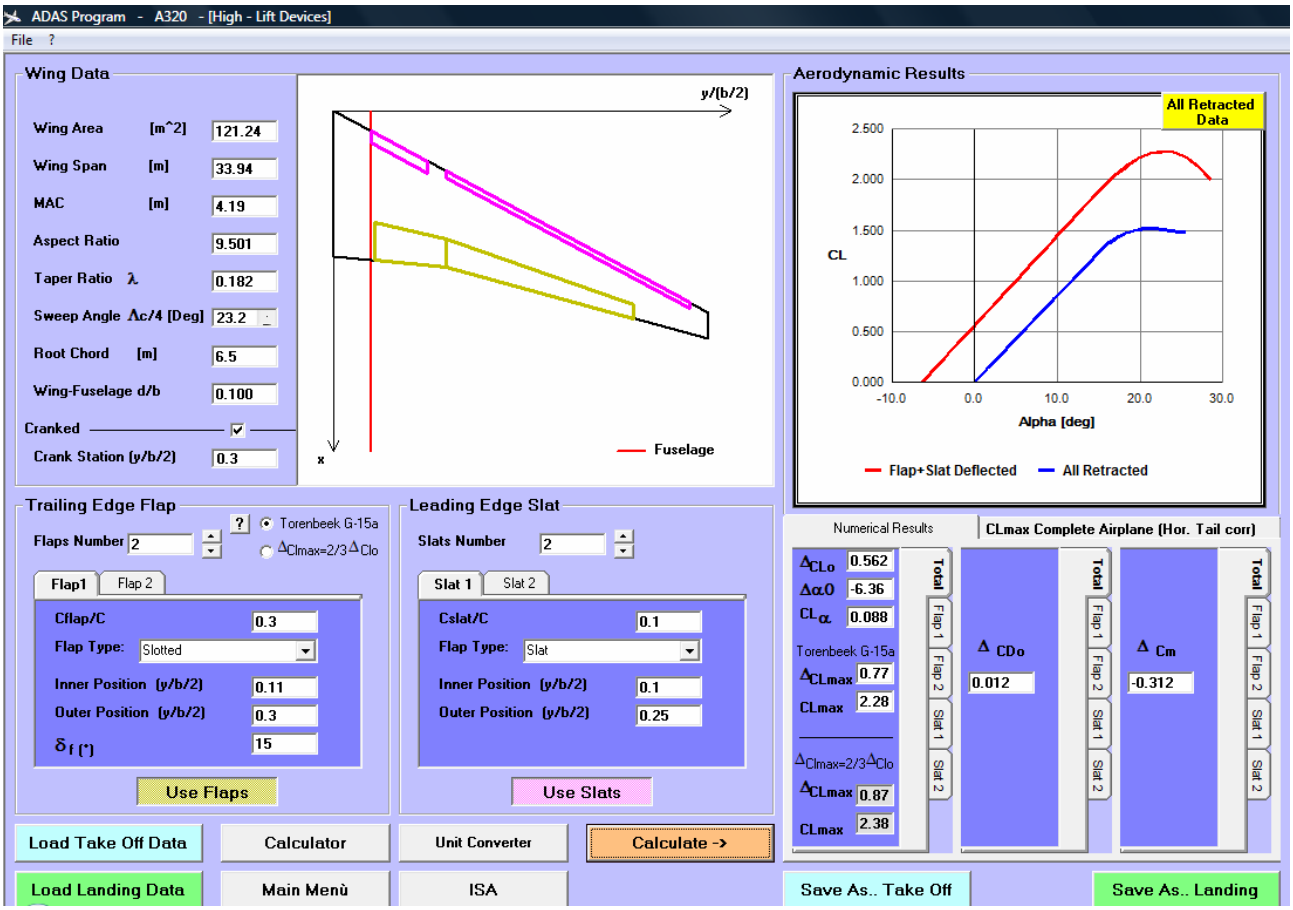


Fig. 8 High-lift system design dialog window. Definition of flap/slat characteristics and estimation of effects on lift.

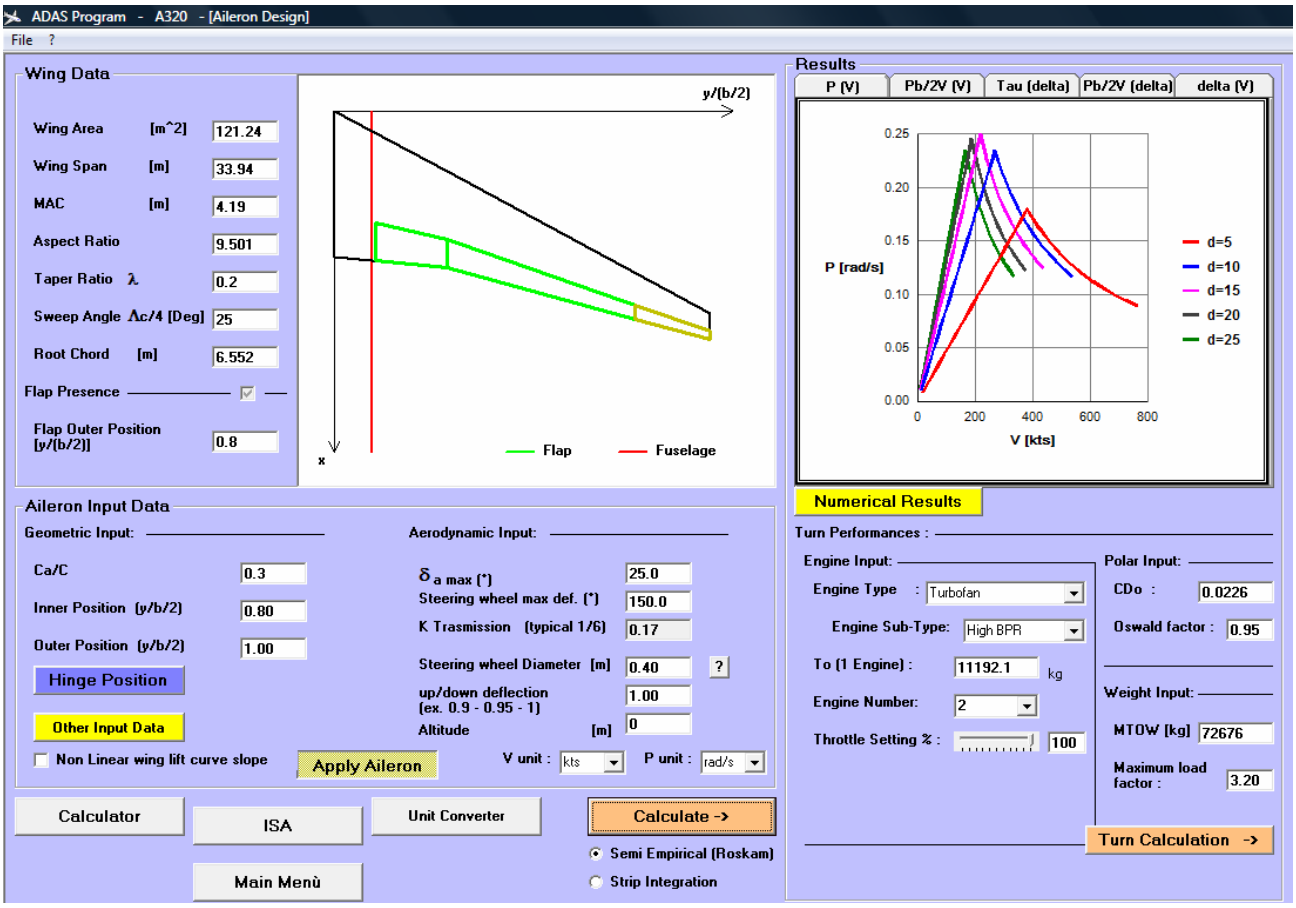


Fig. 9 Aileron Design dialog window and aircraft roll performances calculations.

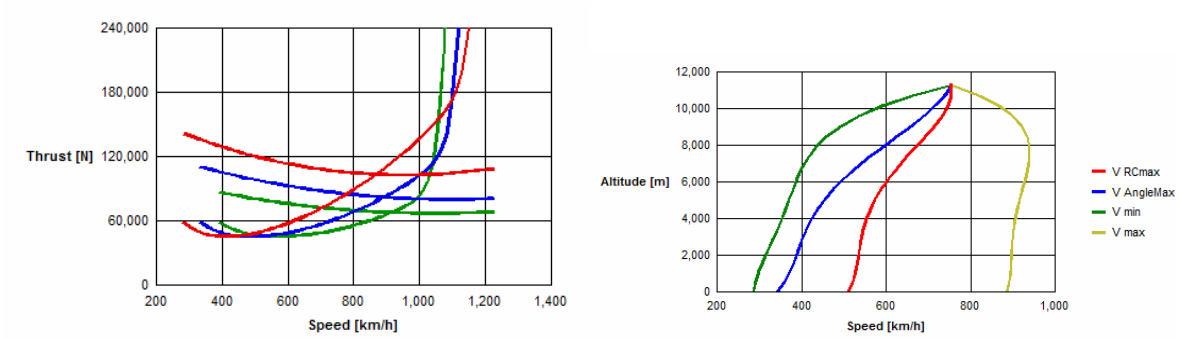


Fig. 10 Flight Performances calculations. Necessary and available thrust (left) and flight envelope (right).

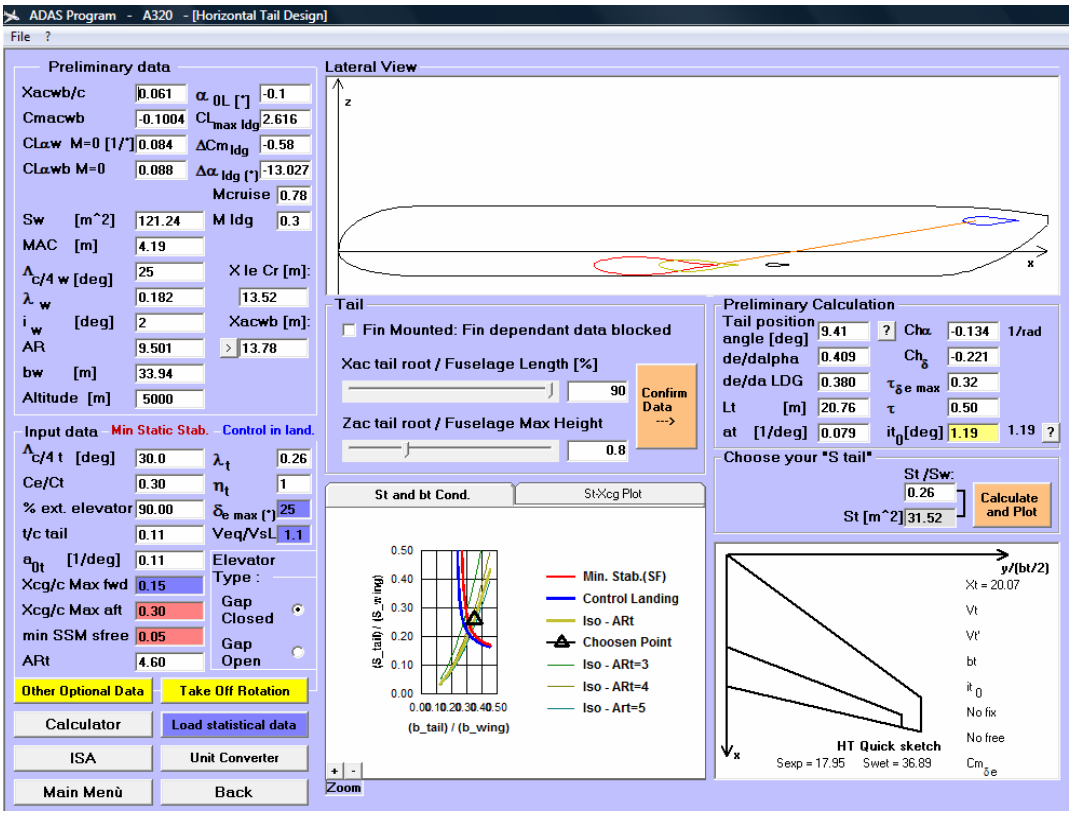


Fig. 11 Horizontal tailplane design. Dialog and interface window and output results.

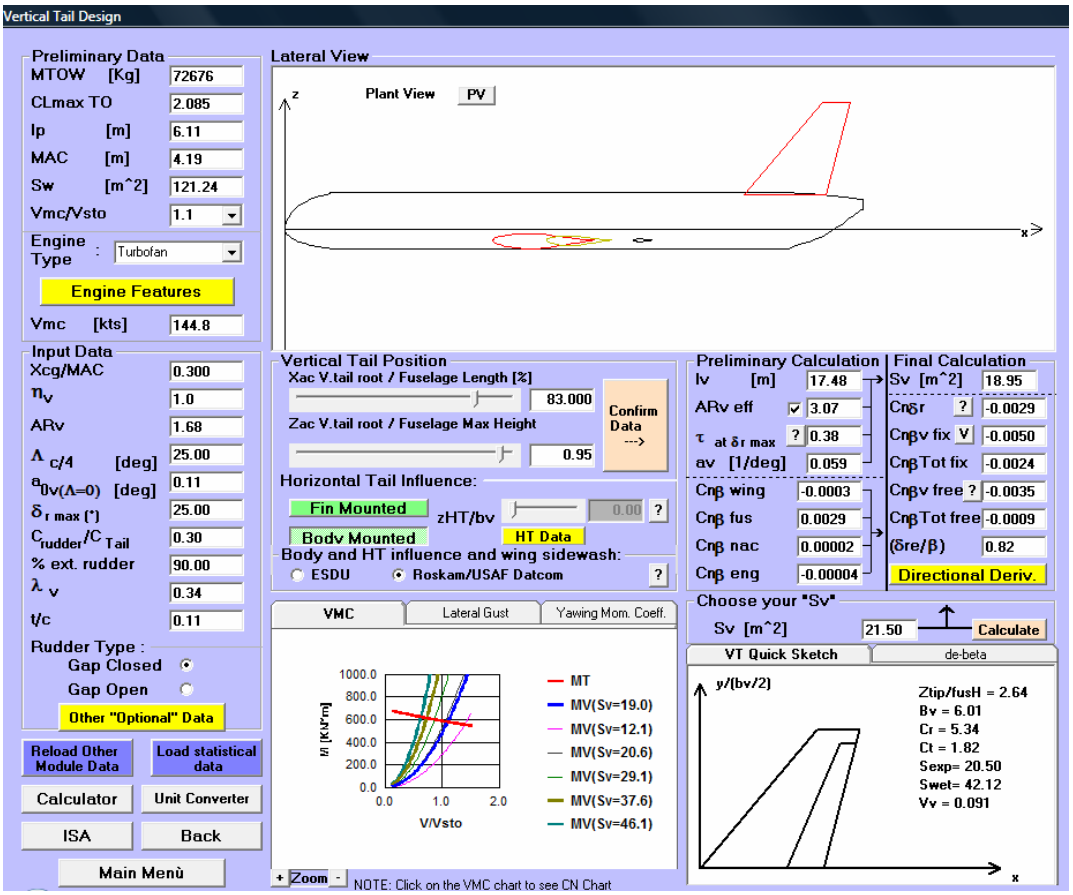


Fig. 12 Vertical tailplane design. Dialog and interface window and output results.

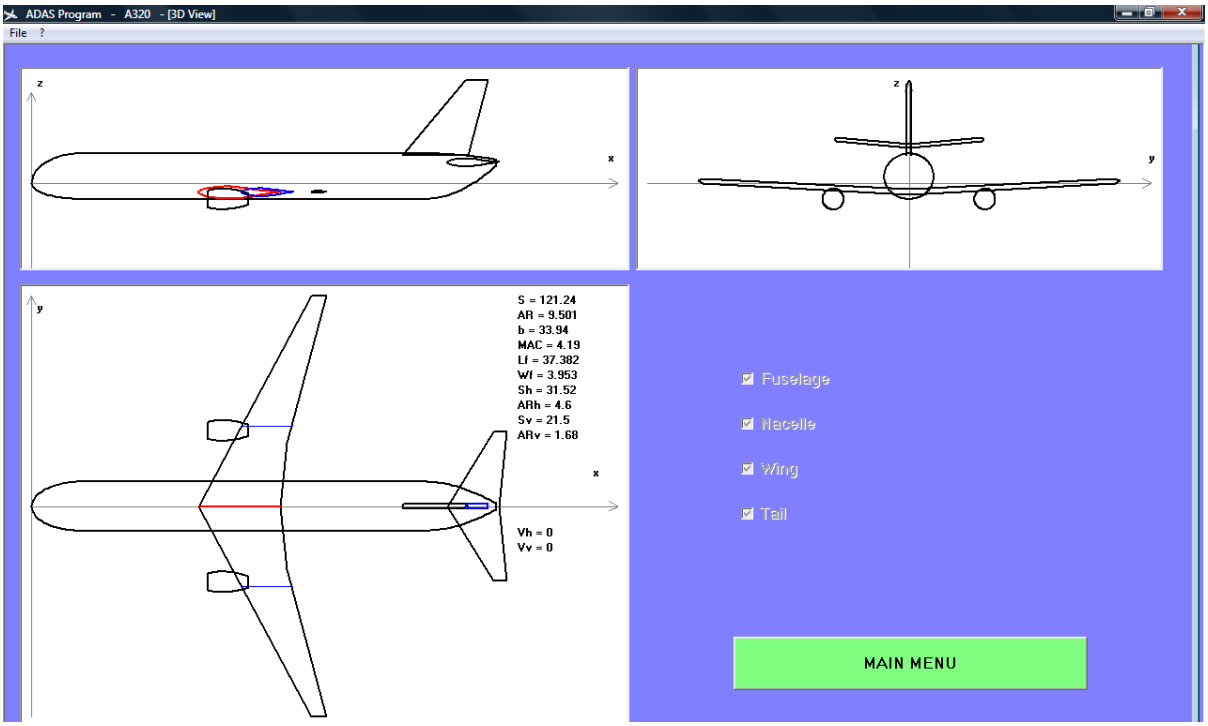


Fig. 13 Final designed aircraft 3-view drawings.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Whole Aircraft Simulation for System Design and Optimsation in Preliminary Design

Petter Krus
Division of Fluid and Mechatronic Sytems, Linköping University, Sweden

Keywords: *System Simulation, Flight Dynamics, Aircraft Systems*

Abstract

This paper address some key methods needed to realize whole aircraft simulation already at the preliminary design stage. E.g. the system has widely different time scales from fractions of seconds to hours, and although little is known about the aircraft, a control system is needed for flying the mission. Methods used here are; system simulation using distributed solvers for robust large scale simulation, time scale compression to deal with the widely different time scales involved, e.g. the transient behaviour of control surface actuators, to the tracking of fuel consumption.

1 General Introduction

Traditionally aircraft conceptual design involves very few aspects of system design. There are, however, a great advantage if system subsystem design also can be involved early on. There are several reasons for that. Modern aircraft are very compact so system installation is important to take into consideration, and this can not properly be done unless there is a preliminary design of the systems. Furthermore, energy efficiency and the impact of subsystem on the energy efficiency of the whole aircraft are becoming important in order to select the proper subsystem concept. The aim here is to allow the whole aircraft to be simulated with its subsystem as early in preliminary design as possible, and then use this model to develop the design further. In this way the simulation model becomes a point of convergence for the aircraft conceptual design as well as the conceptual design of the subsystem as well as a tool for further development in preliminary design.

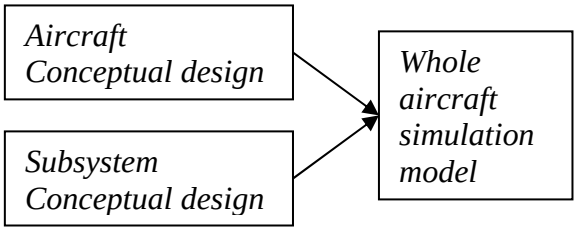


Figure 1. The simulation model as a convergence point in aircraft conceptual design.

Using simulation performance characteristics at the whole aircraft level can be evaluated very straightforwardly, and things like trim drag are accounted for as a byproduct. Furthermore, as the design progress into sub system design, systems such as actuation systems, fuel systems etc, can also be included and verified together in the actual flight profile. Using simulation models of whole aircraft with subsystems, it is then possible to do design analysis, e.g. sensitivity analysis and trade of analysis, as well as design optimisation. In this paper this is demonstrated on a flight simulation model with subsystems, such as control surface actuators.

In recent years there has been a lot of development in methods and tools suitable for simulation of systems. One example is the Hopsan-NG simulation package developed at Linköping University. This means for example that it is possible to model basic aircraft systems, such as hydraulic system, air system

and fuel system, much more efficiently than before and that a lot of systems can even be simulated in real time or faster than real time. In this paper it is shown how subsystem models also can be coupled to models of flight dynamics, propulsion, and flight control, to produce a more complete aircraft system model. Such a model can be used already in preliminary design, thus allowing the preliminary subsystem designs to be designed concurrently with the aircraft layout.

One problem when dealing with large complex systems, however, is that most simulation packages rely on centralised integration algorithms that scale rather poorly with respect to system size. For large-scale systems it is an advantage if the system can be partitioned in such a way that the parts can be evaluated with only a minimum of interaction. The reason that centralised solver dominates is most likely, that until recently, the typical system simulation has been of a moderate size. In this paper, distributed solvers with linear scaling properties is used, which means that simulation speeds, orders of magnitude higher than real time, can be achieved for system simulation.

Furthermore, development in single processor performance is leveling out. On the other hand, multi-core architectures have become the norm. However, using conventional simulation software, the simulation can only use one core for the simulation, thus not exploiting that potential. The distributed modeling approach used in here, however, makes it intrinsically suitable for multi-core simulation, and Hopsan already has that capability implemented.

2 Distributed Modelling

A very suitable method for modelling and simulation of large complex dynamic systems is represented by distributed modelling using transmission line elements. The origin of this concept goes back at least to Auslander 1968 [1] who first introduced transmission lines (or bi-lateral delay lines). This method evolves naturally for calculation of pressures when pipelines are modelled with distributed parameters. This approach was adopted for simulation of fluid power systems with long lines in the HYTRAN program [2] already in the seventies. The method can be generalised to both mechanical and electrical systems.

A related method is the transmission line modelling method (TLM) presented by Johns and O'Brien (Ref. [2]) for simulation of electrical networks.

Johns and O'Brien pointed out that an important aspect of modelling using transmission line elements is that most of the numerical errors introduced by an ordinary solver are avoided. The errors made due to the introduction of transmission line elements, are better described as modelling errors.

An attractive feature with this is that laws of conservation of mass and energy still hold for the solution, since there always exist a plausible physical system for the model although the line lengths may vary compared to the original system. This also implies that the user may tolerate a larger numerical error since, generally, quite large modelling errors are present anyway (errors of the order of 10% are generally considered acceptable from an engineering point of view).

A key feature of the transmission line is the finite signal propagation speed (speed of sound) of the signal travelling through the line. This means that events at one component do not affect another immediately. Using this approach the subsystems can be solved independently in each time step in contrast to using conventional solver. The implementation of the numerical solver of differential algebraic equations can therefore be implemented in the subsystem rather than at a central level. This approach is sometimes referred to as distributed modelling for several reasons. The simulation of waves in transmission lines means that distributed parameters are used in the lines, secondly it allows for distributed solver of differential algebraic equations in component and subsystems. There is also the possibility of using distributed processing by allocating subsystems to different processors.

In this way the system is divided into subsystems that generally are of limited size and

the very robust method for DAE:s described earlier, can be used.

2.1 The Unit Transmission Line Element

In transmission line modelling the basic dynamic element is the unit transmission line. In the Hopsan package this is used to connect different components to each other. In the general case it can be used to model both capacitances and inductances. In the Hopsan-package, however, it is used only to represent capacitances (oil volumes and mechanical springs).

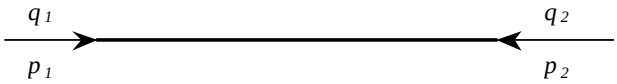


Figure 2. Transmission line

The complete set of equations that describes a lossless transmission line are:

$$p_1(t) = p_2(t-T) + Z_c q_1(t) + Z_c q_2(t-T) \quad (1)$$

$$p_2(t) = p_1(t-T) + Z_c q_2(t) + Z_c q_1(t-T) \quad (2)$$

Here Z_c is the characteristic impedance of the line, p and q are pressures and flows respectively. T is the time delay in the line. Note that the main property of these equations is the time delay they introduce in the communication between the ends. Introducing

$$c_1(t) = p_2(t-T) + Z_c q_2(t-T) \quad (3)$$

$$c_2(t) = p_1(t-T) + Z_c q_1(t-T) \quad (4)$$

Here c is the wave variables that represent information that has been transmitted from the other side of the transmission line. With these, the following set of equations is obtained.

$$p_1(t) = c_2(t) + Z_c q_1(t) \quad (5)$$

$$p_2(t) = c_1(t) + Z_c q_2(t) \quad (6)$$

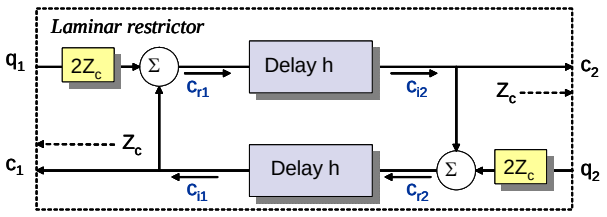


Figure 3. Block diagram of transmission line.

An interesting observation is found if c_2 in Eq. (18) is substituted with Eq. (19) and the outlet at 2 is blocked.

$$p_2(t) = p_2(t-2T) + Z_c[q_1(t-2T) + q_1(T)] \quad (7)$$

Compared to the trapezoidal method for integration (h is the time step).

$$y_{h+t} = y_t + \frac{1}{2} h(f(u_t, t) + f(u_{h+t}, h+t)) \quad (8)$$

where

$$\dot{y} = f(u, t) \quad (9)$$

These equations are the same if $T = h/2$

The relationship between flow entering a volume and the pressure can be written as:

$$p = \frac{q}{C} \quad (10)$$

where C is the capacitance. Identification yields

$$Z_c = \frac{h}{C} \quad (11)$$

The implication of this is that if we use the trapezoidal method to integrate pressure in a volume (capacitance) between two components, this corresponds to introducing a short pipe instead of a pure capacitance. This can also be extended to other domains such as mechanical and electric.

3 Differential Algebraic Systems

In order to solve the dynamics of the individual components and subsystems any type of solver can be used. A general approach to represent a system is to represent it as a differential algebraic system. This also allows for algebraic loops.

$$F(x, \dot{x}, t) = 0 \quad (12)$$

where x is the variable vector.

However, Eq. (1) implies that the system essentially has to be written in state space form, something that may be considered as too limited. Many relationships are usually given in transfer function form, which makes it more natural to allow for higher derivatives. The system can then instead be expressed as

$$F\left(y, \frac{dy}{dt}, \frac{d^2y}{dt^2}, \ldots, \frac{d^ny}{dt^n}, t\right) = 0 \quad (13)$$

This also has the advantage that the variable vector is reduced, since y is shorter than x, y contains a subset of the states in x. It should, however, be pointed out that it is only possible to impose strong non-linearities (such as limitations on the state variables) represented in the y vector. Also all variables that are of any interest must be included in the y vector, otherwise they will not be computed explicitly.

In order to solve the dynamic part of the system in a numerically stable way, the trapezoidal rule can be used. Using the trapezoidal rule (or bilinear transform) the time differential is solved as:

$$x(h+t) = x(t) + \frac{1}{2}h(x(t) + x(h+t)) \quad (14)$$

A more effective way of using the trapezoidal rule is to reformulate it in the form known as the bilinear transform.

$$\frac{d}{dt} = \frac{2(1-q^{-1})}{h(1+q^{-1})} \quad (15)$$

where q in this context represents the time displacement operator such that:

$$qy = y(h+t) \quad (16)$$

Using the bilinear transform in Eq. (4) means that it can be rewritten as a function G of y and old states.

$$G(y(t), y(t-h), \ldots, y(t-nh)) = 0 \quad (17)$$

When solving the system all the old values y(t-h)... y(t-n h) can be regarded as constants since they have already been established in previous time steps. It is therefore rewritten as:

$$G(y(t), t) = 0 \quad (18)$$

In order to solve this system of equations in a numerically stable way, the Jacobian matrix is needed, which is defined as:

$$J_{ijk} = \frac{\partial G_i(y_k(t))}{\partial y_j} \quad (19)$$

The equation can then be solved numerically using Newton-Raphson iteration.

$$y_{k+1}(t) = y_k(t) - J_k(t)^{-1} G(y_k(t)) \quad (20)$$

Since an iterative procedure is used, there is a potential for performance loss due to the number of iteration needed to solve the system. However, since the values from the previous time step can be used as start values.

$$y_0(t) = y(t-h) \quad (21)$$
$$J_0(t) = J(t-h)$$

If the system is linear, the system can be solved in only one iteration, and it is usually sufficient with only one iteration even for non-linear systems, especially if a small time step is used. There are, however, situations when input signals changes suddenly, e.g. a valve is changed step wise during one time step that requires more than one iteration. In practice, however, it has been found that two iterations increase the tolerance against non-linearities dramatically, while a further increase to three iterations gives only minor improvement. Two iterations have therefore been found to be something near to an optimum for almost all situations. For implementation it is better to use LU-decomposition rather than using the matrix inverse of the Jacobian.

Provided the system is reasonably linear (slow variation of J, Eq. (9) is an A-stable method. However, in reality, rather large variations of J can be tolerated. Even pure discontinuities can also be handled satisfactory using the above approach, when fixed-time step is used (as in real-time simulation).

Eq. (9) also illustrates a dilemma associated with all numerically stable methods. They need knowledge of the Jacobian, and if the system is stiff and highly non-linear this must be updated (and inverted) very often. We also realize that the computational burden is much more than linearly dependent on the size of the system. This makes these methods unsuitable for large problems. Eq. (9) is, however, very effective for solving small systems which makes it very suitable for solving subsystems.

Instead of using equation (9) directly it is wise to replace the inverse of the Jacobian by using LU-decomposition instead, there are also

a few other actions that can be done in order to further enhance the efficiency of the solver.

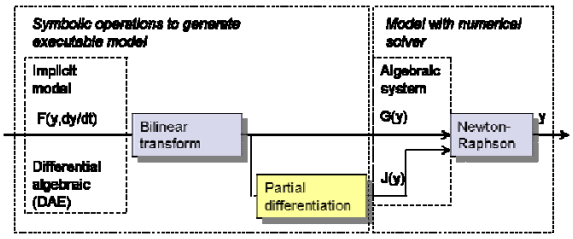


Figure 4. Scheme of transformations

Figure 4. shows a scheme of the transformations involved. The differential algebraic system DAE is transformed into time discrete form using bilinear transform. This method needs the Jacobian of the system. This can, however, be obtained analytically using symbolic math packages such as Mathematica, and a model generator has been written in the Mathematica symbolic manipulating language called Compgen, which takes the acausal component description similar to Modelica, [5], and transforms it into a component model implementation in C++ that can be used for simulation. This approach is described in more detail in Ref. [5] (although for generating Fortran models for the old Hopsan package).

Packages for symbolic math can also be used to facilitate the modelling work in other ways, so that the user can concentrate on formulating the equations rather than solving them. The 6DOF-flight dynamics model is an ideal example since it involves a lot of co-ordinate transforms that can be set up entirely in the Mathematica environment. Using the Compgen tool a model for the Hopsan-NG simulation package can then be generated.

As an example a small unmanned jet power aircraft is simulated. The control surfaces are powered by conventional hydraulic servo actuators. There is an attitude control unit that is fed by an external altitude controller. There is a fully non-linear six degree of freedom flight dynamics model with a non-linear aerodynamic model. Data for the aerodynamic model can be obtained in the early design phases from panel codes and/or handbook formulas, and is here based on a static version of the model presented in [6], although the unsteady effects can of course also be included. .

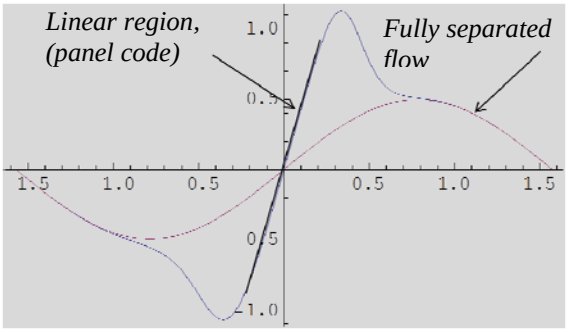


Figure 5. Non-linear aerodynamic model.

automatically included, and the effect of reduced weight as fuel is consumed.

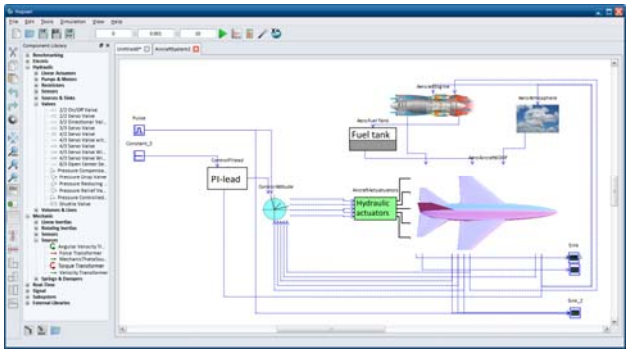


Figure 6. Screen shot of Aircraft system simulation model.

The connection between the actuators and the aircraft control surfaces are through transmission line elements that models the aerodynamic stiffness of the control surfaces. The moment at the actuator link for the elevator, can be calculated as:

$$M_1(t) = c_\delta(t-h) + Z_{c\delta}\dot{\delta}(t) \quad (22)$$

$$c_\delta(t) = M_2(\delta, \theta, \alpha, q) \quad (23)$$

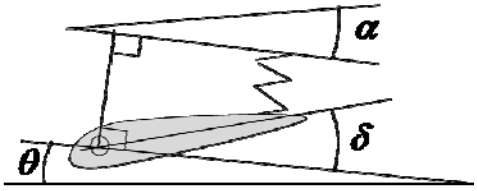


Figure 7. Model of aero load on rudder as a (non-linear) spring.

Here M_2 is the moment from the aerodynamic forces, which is a function of surface angle δ, aircraft pitch angle θ, and angle of attack α, and dynamic pressure q. M_1 is the corresponding moment at the actuator hinge, Ideally they are the same, but through the separation with the transmission line element they becomes slightly different. M_2 is a non-linear function and hence not a purely linear spring. The characteristic impedance is calculated as

$$Z_{c\delta} = hk_{\delta} \tag{24}$$

where the linearized aerodynamic stiffness and is calculated as

$$k_{\delta} = \frac{\partial M_2}{\partial \delta} \tag{25}$$

Note that this linearization is only used for the numerical solver, and does not affect the steady state value of the moment.

3.1 Subsystem modeling

In order to deal with complexity it is necessary to be able to model in a hierarchical way. This is done by introducing subsystems. In this example the hydraulic actuation system is one subsystem representing the control surface actuators and the hydraulic supply unit. The hydraulic supply system is represented by the pressure controlled pump, a pressurized tank (represented by an asymmetric piston in the model), and an accumulator.

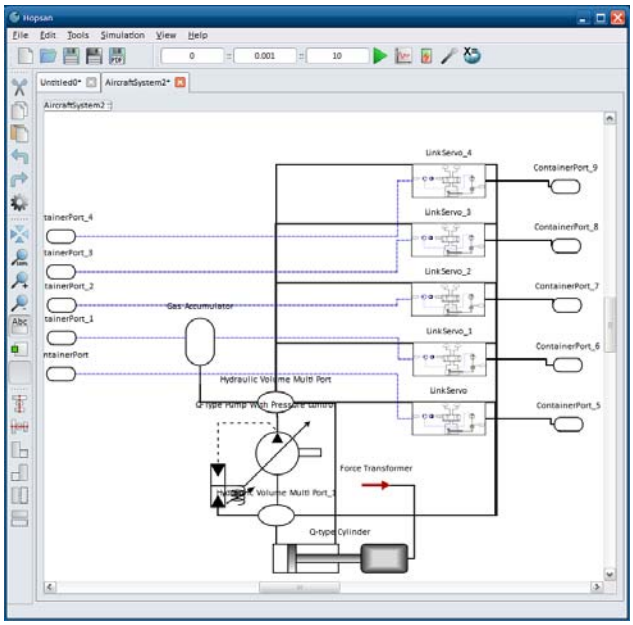


Figure 8. Model of the hydraulic system.

The servo actuators are also modeled as subsystems with servo valve, piston and linkage.

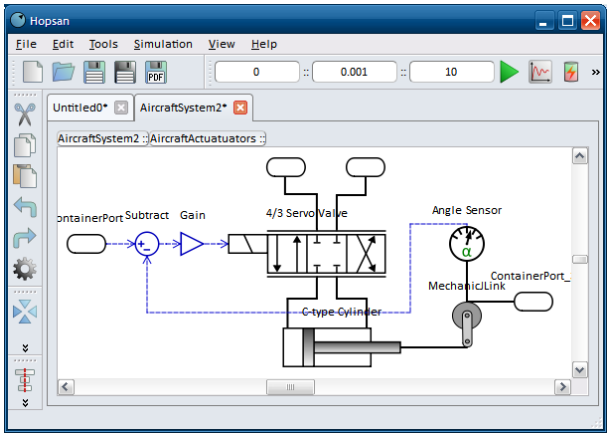


Figure 9. Model of the servo actuator subsystem.

The redundancy has not been modeled here, but it is very straightforward to include.

3.2 Simulation results

Using the simulation model a simulation of a mission can be performed several times faster than real time PC. This means that is possible to use the model for design analysis such as sensitivity analysis. Furthermore, it is even possible to use it for optimization of some design parameters.

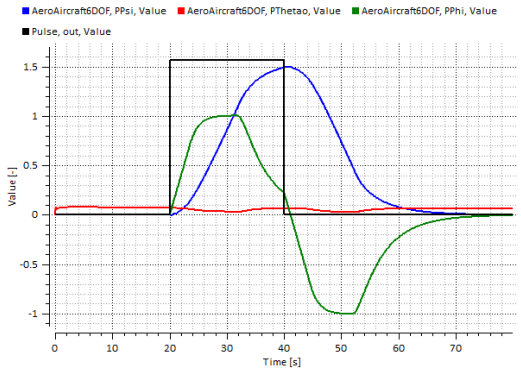


Figure 10. The aircraft attitudes during an S-maneuver.

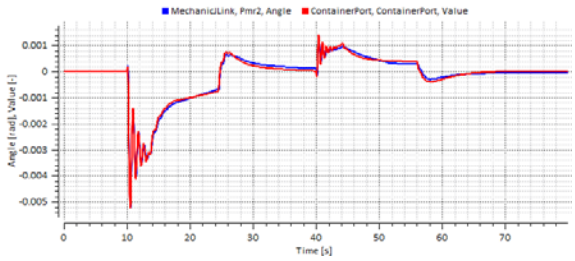


Figure 11. Angular position and reference position of the rudder actuator.

This simulation model is part of an integrated framework for aircraft design developed at Linköping University. In order to be part of a work flow, the system simulation needs to be able to integrate with other software. Hopsan-NG models can be exported for simulation within Matlab Simulink. Furthermore, the flight dynamics model has a parameter set that is very general, which means that coefficients can be obtained from other sources. Data is stored in XML for simple conversion and import of data.

4 Whole Mission Simulation

This model can also be simulated throughout a mission. In addition mission simulation also need a mission control unit where the flight is controlled using waypoints. In this way also performance measures such as fuel consumption can be assessed. Furthermore, with this kind of simulation full mission system analysis can be performed, including sensitivity analysis and studying the influence of uncertainty in parameters.

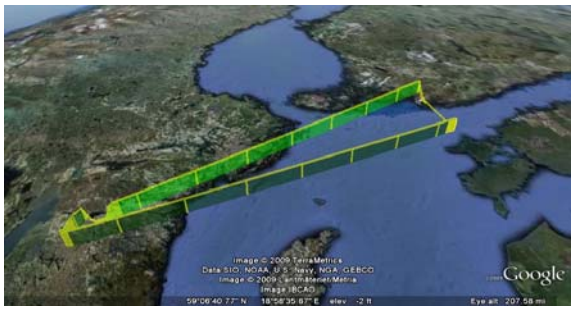


Figure 1. Simulated flight path.

4.1 Influence of uncertainty in coefficient values

One way to manage the required accuracy of the coefficients involved in the simulation is to study the influence of removing the uncertainty of one coefficient (uncertainty variable), as in [16] and [17]. This is done through sensitivity analysis and is presented in a matrix where the relative influence on the functional characteristics, such as fuel consumption and mission time, is indicated.

Functional characteristics	Units	Normalized deviation	ap1Lp	Cd01Lp	Cd02Lp	Cd0bLp	CLalpha1Lp	CLde2Lp	e1Lp	SFC	MeLp	xbcgeLp
			0.09	0.15	0.15	0.50	0.10	0.20	0.04	0.11	0.03	0.05
fuel_consumed	[kg]	0.53	0.00	0.00	0.01	0.09	0.00	0.00	0.00	0.07	0.00	0.11
fuel_reserve	[kg]	1.04	0.00	0.01	0.01	0.17	0.00	0.00	0.00	0.15	0.00	0.23
mission_time	[s]	0.71	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.14	0.00	0.13
maneuvering_ti	[s]	0.24	0.00	0.00	0.02	0.02	0.02	0.00	0.00	0.02	0.00	0.02

Figure 12. Influence matrix of uncertainties in coefficients on some selected functional characteristics derived from the simulation results.

5 Conclusion

In this paper the use of flight simulation, including subsystems, for design evaluation of the whole aircraft in the preliminary design phase is studied. Since efficient, hi-speed simulation of a complex system is desirable, robust simulation is a requirement. A key element is the method of using bi-lateral delay lines for partitioning complex system models which has been described. In this way it is possible to use highly robust distributed solvers on subsystems that are then connected to each other using the bi-lateral delay lines, for system simulation. This means that highly robust models are achieved where very large time steps can be used. As a consequence fairly detailed

simulation models can be used for whole mission simulation already at the preliminary design stage, for system optimisation and design analysis.

References

[1] Auslander D M , 'Distributed System Simulation with Bilateral Delay-Line Models' Journal of Basic Engineering, Trans. ASME p195-p200, June 1968.

[2] Amies G., R Levek, D Struessel, 'Aircraft Hydraulic Systems Dynamic Analysis. Volume II. Transient Analysis (HYTRAN) Computer Program Technical Description'. A730930. 1977.

[3] Burton, J.D., Edge, K.A. and Burrows, C.R., 'Modelling Requirements for the Parallel Simulation of Hydraulic Systems', ASME Journal of Dynamic Systems and Control, March 1994

[4] Duquette. 'Effects-Level Models for UAV Simulation', AIAA Modeling and Simulation Technologies Conference, Chicago, Illinois, USA, 2009

[5] Fritzson P., Principles of Object Oriented modelling and Simulation with Modelica 2.1, Wiley-IEEE Press. 2003.

[6] Jouannet C., Krus P. 'Unsteady aerodynamic modelling: a simple state-space approach'. AIAAA Aerospace sciences meeting and exhibit,2005 Reno, USA.

[7] Johansson (now Lundén) B., Andersson J., and Krus P., "Thermal Modelling of an Electro-Hydrostatic Actuation System," in Proceedings of the International Conference on Recent Advantages in Aerospace Actuation Systems and Components, Toulouse, France, June 13-15, 2001.

[8] Johns P., B. and M.O'Brien. 'Use of the transmission line modelling (t.l.m) method to solve nonlinear lumped networks.' The Radio Electron and Engineer. 1980.

[9] Krus P., 'An Automated Approach for Creating Components and Subsystems for Simulation of Distributed Systems' Presented at the 'Ninth Bath International Fluid Power Workshop', Bath, UK 1996.

[10] Krus P, 'Complete Aircraft Simulation for Distributed System Design", in Proceedings of the International Conference on Recent Advantages in Aerospace Actuation Systems and Components', Toulouse, France, June 13-15, 2001.

[11] Krus P., A Jansson, J-O Palmberg, 'Optimization Using Simulation for Aircraft Hydraulic System Design', Proceedings of IMECH International Conference on Aircraft Hydraulics and Systems, London, UK, 1993

[12] Krus P., A Jansson, J-O Palmberg, K Weddfelt. 'Distributed Simulation of Hydromechanical Systems'. Presented at 'Third Bath International Fluid Power Workshop', Bath, UK 1990.

[13] Eriksson B., P Nordin, P Krus 'Hopsan NG, A C++ Implementation using the TLM Simulation Technique link', The 51st Conference on Simulation and Modelling (SIMS), Oulu, Finland. 2010

[14] Axin, M., R Braun, A Dell'Amico, B Eriksson, P Nordin, K Pettersson, I Staack, P Krus, 'Next Generation Simulation Software using Transmission Line Elements link', Fluid Power and Motion Control (FPMC), Bath, UK, 2010

[15] Braun B., P Nordin, B Eriksson, and P Krus, 'High Performance System Simulation Using Multiple Processor Cores', The Twelfth Scandinavian International Conference on Fluid Power (SICFP), Tampere, Finland, 2011.

[16] Krus P., 'Computational tools for aircraft system analysis and optimization'. The 26th International congress of the aeronautical sciences. ICAS, Anchorage, USA, 2008.

[17] SteinkellnerS. Licentiate thesis. 'Aircraft Vehicle Systems Modeling and Simulation under Uncertainty', LIU-TEK-LIC-2011:36, Linköping University.

Parametric wing profile description for conceptual design.

Tomas Melin, tomas.melin@liu.se, Kristian Amadori, kristian.amadori@liu.se
and Petter Krus, petter.krus@liu.se

Division of Fluid and Mechatronic Systems
Department of Management and Engineering
Linköping University
SE-581 83 Linköping, Sweden
Tel: +46 13 28 10 00 Fax: +46 13 13 04 14

Introduction

A fundamental part of aircraft design involves the wing airfoil optimization, establishing an outer shape of the wing which has good aerodynamic performance, good internal volume distribution for fuel and systems and which also serves as an efficient structural member supporting the weight of the aircraft. As for all optimization tasks, the complexity of the problem is directly coupled to the parameterization of the geometry. Of highest relevance are the number of parameters and the number of additional constraints that are required to ensure valid modeling.

This paper proposes a parameterization method for two dimensional airfoils, aimed at providing a wide design space, while at the same time keeping the number of parameters low. With 15 parameters defining the wing profile, many of the existing airfoils can be modeled with close tolerance.

Several approaches to parameterization of wing profiles can be found in the literature. Airfoils can be described by point clouds as done in most airfoil libraries [1]. The number of parameters is twice as large as the number of points used (x and y coordinates) and in the case of aerodynamic optimization this parameterization will most certainly be not well behaved, since no smoothing function is included and must therefore be added. Other problems may arise for the fact that the airfoils sometimes are defined with too few coordinate points and/or too few decimals, a problem occurring especially with old airfoils. On the other hand, the design space that this kind of parameterization allows representing is extremely large, as any and all shapes can be reproduced, even degenerate ones.

Airfoils can also be represented by mathematical functions. Among the most famous representatives of this category are indeed the NACA 4-, 5- and 6-digits formulations [2] [3]. Compared to point clouds, they could be said to represent the opposite case: they are very well behaving parameterizations, but they cannot cover a very large design space, since they only provide four or more parameters respectively to be tuned. The NACA 4 digit series is particularly interesting as the parameters are a part of the name of the airfoil. In the case of the 5- and 6 digit series, the name is instead constructed from the airfoils aerodynamic characteristic and geometry.

Another known set of theoretically defined airfoils are the Joukowski profiles [4]. Using the conformal mapping method, airfoils with a round nose and sharp trailing edge can be represented. Sadly the method is not to recommend for trying to match known airfoils and the design space it describes is quite confined to airfoils with often poor performances. More recently a very interesting and powerful representation method was presented by Kulfan [5]. The method is based on simple well behaved mathematical functions that allow for increasing the

number of parameters at will. Also the formulation is not restricted to round nose and sharp trailing edge profiles, but can describe almost any shape.

Problem

The authors identifies the need of a wing profile modeling method which shares the large design space with the point cloud method of representation, while maintaining a small number of parameters. By employing a functional parametric description with built-in smoothing and interpolation, coordinate resolution becomes infinite. Some basic assumptions regarding the shape of an airfoil are made: The airfoil is assumed to have one distinct top and bottom point, the chord line has a length of one, forming the x-axis of the profile, the trailing edge is assumed to be either sharp, i.e. with a y-coordinate of zero for both top and bottom sides, or blunt with the y coordinates of the top and bottom trailing edges being symmetric around the chord line.

Method

The airfoils top and bottom curves are both modeled by two cubic Beziér curves in parametric form being C^1 continuous at the defined top and bottom points. Figure 1 gives an example of a generic high speed profile and show the outer envelope, the camber line and the Beziér control points. Equation 1 shows the parametric Beziér curve.

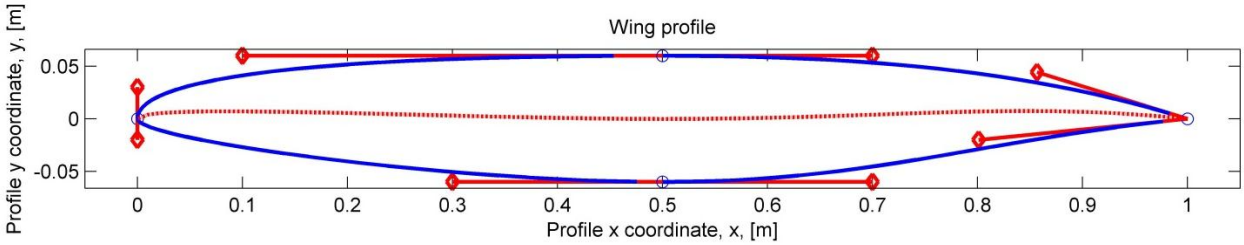


Figure 1: Generic model of a modern high speed airfoil. Diamonds are the Beziér control points. The dotted line is the camber line.

$$\mathbf{B}(t) = (1-t)^3\mathbf{P}_0 + 3(1-t)^2 t\mathbf{P}_1 + 3(1-t)t^2\mathbf{P}_2 + t^3\mathbf{P}_3 \ , \ t \in [0,1]. \quad (1)$$

The method of representing a new airfoil with the proposed method is straight forward. As a baseline, a four part cubic Beziér curve, all 13 control points are needed to define the curve, giving 26 variables when both x and y coordinates are taken into account. However when some symmetries and simplifications are taken into account the number of independent parameters are reduced to 14. The Beziér control points for the wing profile are defined clockwise around the profile from the trailing edge bottom side, according to figure 2. Control point 1 and 13 always have the same x-coordinate, one, and the same absolute value of their y-coordinate. Control points 3,4 and 5 on the bottom side have the same y-coordinate. The same goes for the top side control points 9, 10 and 11. The leading edge points 6,7 and 8 all have an x-coordinate of zero. Table one show the dependent parameters for the control points.

	Coordinate	
Point #	**X**	**Y**
1	1	a
2	X_2	Y_2
3	X_3	b
4	X_4	b
5	X_5	b
6	0	Y_6
7	0	0
8	0	Y_8
9	X_9	c
10	X_{10}	c
11	X_{11}	c
12	X_{12}	Y_{12}
13	1	-a

Table 1: Coordinate dependency for the different control points. Some always take the value one or zero, while others (a,b,c) always shares the same numerical while (X_n Y_n) are allowed any value.

.

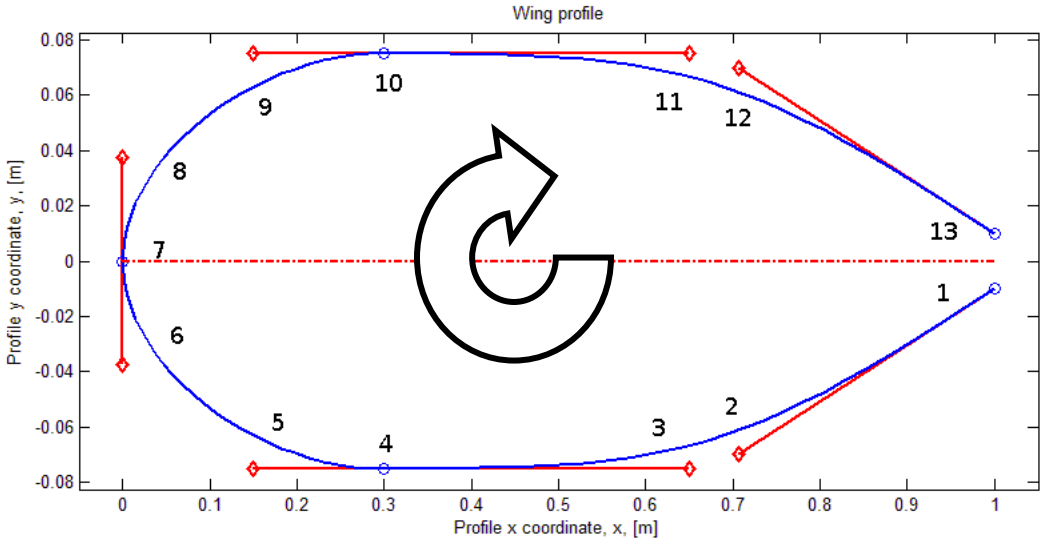


Figure 2: Control point distribution and numbering

Parameter reduction

Handling the control point coordinates directly when specifying an airfoil is not so user friendly. The coordinates themselves have no obvious relation to shape of the airfoil when reviewed alone and the possible range of the coordinates exceed the allowed range for non-degenerate airfoils. To alleviate this concern, and to reduce the number of free parameters, the following parameter transformation is done.

The lengths of the control vectors are limited to prevent degenerate cases. For example the upper leading edge control vector needs to be shorter than the distance between the top point and the chord line to prevent leading edge overshoot.

Thickness: The x and y coordinates of the top and bottom points (# 4 and #10) are kept as is and named upper thickness and upper thickness position and similarly for the lower side. The difference between the upper and lower thickness then becomes the ram thickness, which in case of the thickness positions being the same equals the thickness of the airfoil. If the upper and lower thickness positions are not equal, the ram thickness will be slightly larger than the thickness.

Control vectors:
The length of control the upper leading edge vector drawn by its start point, the leading edge which is point number 7, and point number 8, is defined as a fraction (k_1) of the upper thickness (h_u). The constant k is only allowed to be in the range [0..1], se figure 3. This ensures that the part of the profile going from the leading edge to the top point does not overshoot the top points y value.

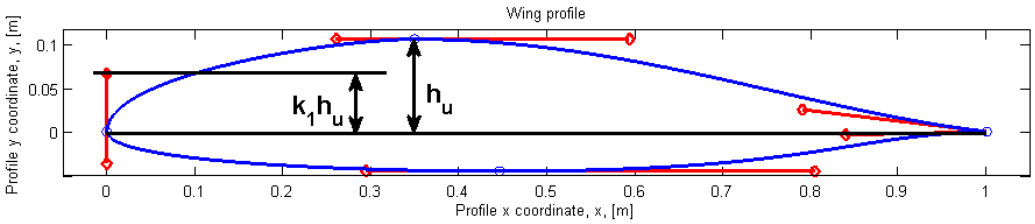


Figure 3: Definition of the leading edge control vector.

The top side of the airfoil is controlled by three parameters, shown in figure (4). The top point position, which is the x-coordinate with the largest y-value, and two fractions (k) which control the lengths of the control vectors. The control vector going from the top point towards the leading edge, i.e. between points #10->#9 has its length limited to a fraction of the top point position in the same way in order to prevent an overshoot of the profile into negative x-coordinates.

Likewise the control vector defined by points 10 to 11, the top to TE vector is a fraction (k_4) of the length between the upper thickness position (#10) and the trailing edge (x=1).

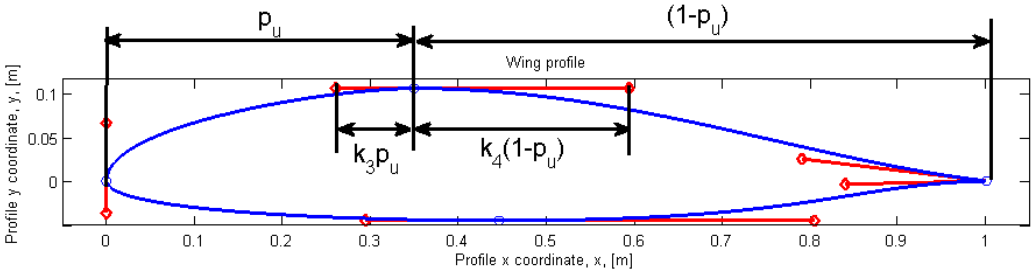


Figure 4: Parameters defining the top side of the airfoil, in this case a parameterized NACA747a415.

The trailing edge is defined by four parameters, shown in figure 5: The boat tail angle (β), which is the angle between tangents of the upper and lower sides of the airfoil at the trailing edge; The release angle (α), which is the angle between the camber line tangent and the chord line at the trailing edge; the trailing edge gap (g), which is the distance between the upper and lower side of the airfoil at the trailing edge; and finally the height of the trailing edge control vector which as the leading edge control vector lengths is defined as a fraction (k_7) of the upper thickness. To fully define the trailing edge we also need the lower side control vector height, which is defined in the same manner as the upper one.

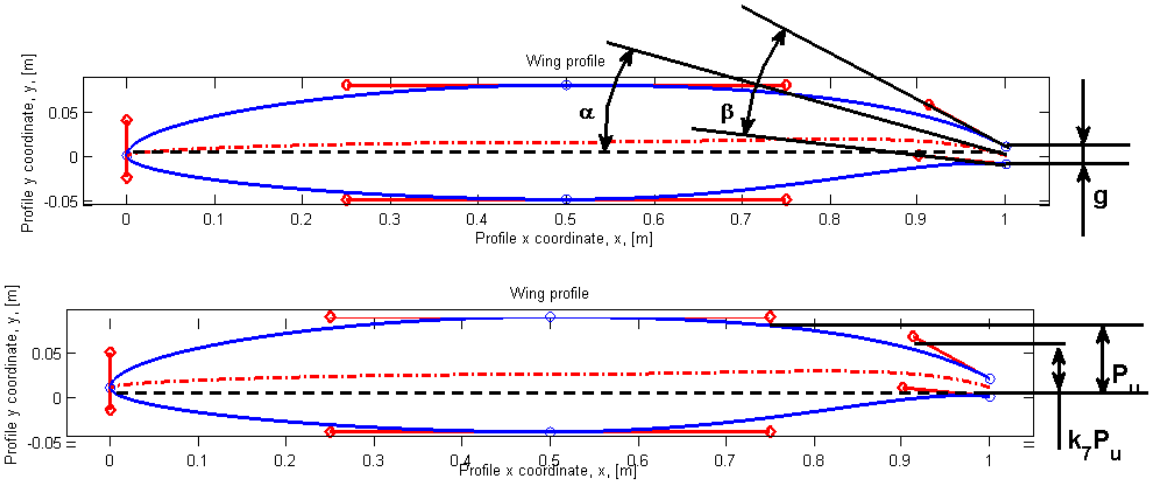


Figure 5: Definition of the parameters of the trailing edge.

The parameters of the lower side of the airfoil are defined in the same manner as the upper side parameters.

The number of parameters are now reduced to 15, see table 2, and by the limits of the fractions many degenerate airfoils are now impossible to generate. By using the chord length as normalizing agent, the parameters are also dimensionless and fully scalable.

Name	Variable	Range
Leading edge		
Upper Nose fraction	k_1	[0..1]
Upper Nose fraction	k_2	[0..1]
Upper side		
Upper thickness	H_U	[0..1]
Upper thickness position	P_U	[0..1]
Upper forward fraction	k_3	[0..1]
Upper rearward fraction	K_4	[0..1]
Lower side		
Lower thickness	H_L	[0..1]
Lower thickness position	P_L	[0..1]
Lower forward fraction	k_5	[0..1]
Upper rearward fraction	K_6	[0..1]
Trailing edge		
Trailing edge gap	g	[0..1]
Boat tail angle	β	$[-\pi..\pi]$
Release angle	α	$[-\pi..\pi]$
Upper trailing edge fraction	k_7	[0..1]
Lower trailing edge fraction	k_8	[0..1]

Table 2: Control parameters for the parameterized airfoil

Reproducing known airfoils

In many design cases it will not be appropriate to start with a generic airfoil for design optimization. Instead, by employing a known airfoil, whether it is a Clark-Y, NACA0012 or any other known airfoil there is a need of having the traditional airfoil shape expressed in these parameters. Modeling known airfoils require slightly more work than creating new airfoils.

Preprocessing

When fitting known wing profiles to the parameterization, the input data are sometimes in poor condition and in need of preprocessing. The point cloud have to have its leading edge at [0 0] and the trailing edge cord have to end at [1 0], if the profile has a trailing edge gap, this should be symmetrically distributed around the [1 0].

One example of wing profiles needing preprocessing is the Clark-Y profile, an older propeller profile today used frequently amongst model aircraft builders for its ease of manufacture rather than aerodynamic performance. The coordinate data for the Clark-Y profile, as presented in some databases [6] has the lower side flank parallel to the x-axis, rather than letting the chord line define the x-axis.

A starting guess of parameters were created by evaluating the profile radius at the leading edge, top and bottom, at the upper maximum y- position and at the lower minimum y-position. Together with the trailing edge gap and the slope of the upper and lower sides of the airfoil at the trailing edge. The radius at these points are computed with a three point circle algorithm from which the length of the control vectors can be computed according to equation 2, where a is the outgoing control vector length and h the perpendicular distance to the next control point, see figure 6.

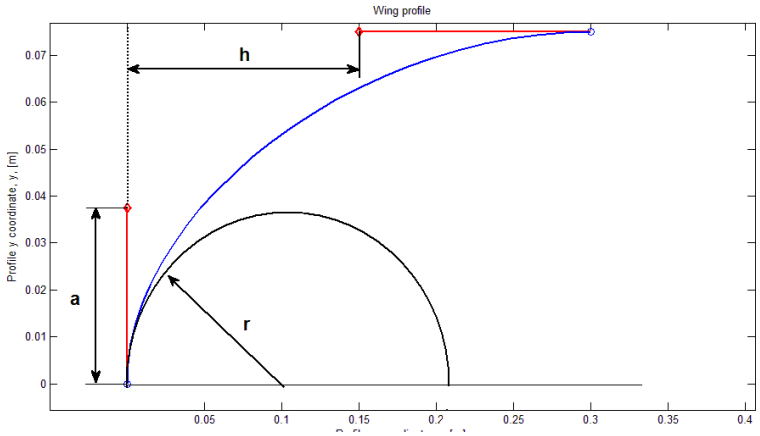


Figure 6: Radius definitions at the leading edge Beziér curve.

$$\frac{1}{r} = \frac{2}{3}\frac{h}{a^2} \qquad (2)$$

Optimization

With the obtained starting guess, a standard optimization is run in MATLAB to fit the parameter curve to the corresponding point cloud. The mathematical method used is the built-in `fsolve` method, which in this case utilizes the Levenberg-Marquardt algorithm. The run is set to minimize the root mean square error of the vertical position of the coordinate points between parameter points and point cloud.

The optimization is nested in two layers in order to treat the parameter nature of the Beziér curve. The outer loop solves for the control parameters governing the curve segment in question and an inner loop solving for the curvilinear position parameter vector t, with respect to the x coordinate. –As the original data points may have any distribution along the x-axis, it can't be assumed that the t vector in equation 1 is evenly distributed.

Comparison with existing data

The optimization results can be evaluated both from a geometrical point of view, were the error of vertical position of each point can be plotted for the profile. The average RMS error for the profile is a general quality control. The average RMS error was usually in the order of 10^{-4}, necessitating the use of the unit "error counts" (ects), where 1 ects = 10^{-4}. Similarly, the curvature and error in curvature between were evaluated – Notably many of the available airfoils had very noisy curvature distributions, making an error plot of curvature fitting interesting but not a good indicator to whether the parametric interpolation was good or bad. Instead, as long as a good scatter is found any offset bias or systematic error can be assumed to be small.

To further investigate the correlation between the original data point cloud and the parametric wing profile, an aerodynamic study was performed utilizing the panel method XFOIL [7]. The original airfoil and the parameterized airfoil were both subjected to a C_L sweep ranging from -0.3 to 1 at a Reynolds number of $6x10^6$. The drag and moment coefficients were evaluated, as well as the boundary layer transition point for both profiles.

Results

The proposed parameterization method was tested on a set of different airfoils. Results for profiles: NACA 0012, Clark-Y, NASA747A415 and a Whitcomb supercritical profile are presented in this report. For the investigated airfoils the reported curvature distribution graph is truncated at the leading edge in order to give a better scale of the results for the rest of the profile. In the Clark-Y example, the curvature of the lower flank is zero, which gives an infinite error there as the parameterization curvature is non-zero.

Figure 7 show the comparison between a point cloud representation of the NACA0012 profile, and the parameterization. The RMS error **of 0.51 ects**, which should be regarded as a good benchmark number when evaluating other airfoils. As the coordinates of the NACA0012 is a parametric curve in functional form in itself, there was a very close match between the two datasets. The geometrical similarity distribution shown in figure 8.

Maximum Thickness	0.12 @ 30 %C
Maximum Camber	0 @ 100 %C
Trailing Edge Gap	*0 %C*
Upper nose radius	1.5 %C
Lower nose radius	1.5 %C
Boat-Tail Angle	16.32 deg
Release Angle	0 deg

Table 3: NACA0012 geometrical properties.

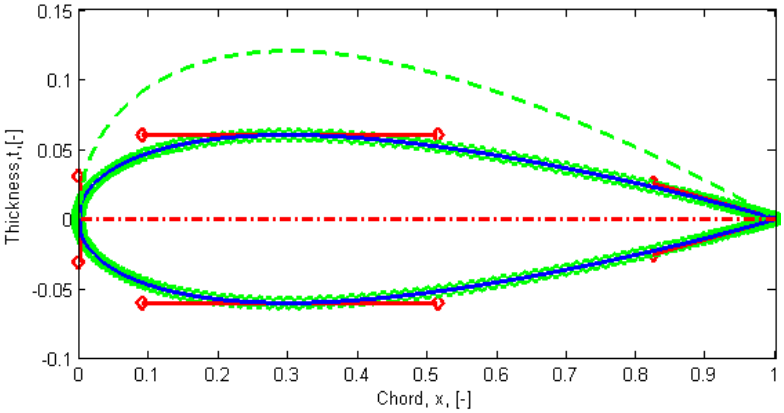


Figure 7: NACA0012 profile. Original points and parametric curve overlaid. Height and width not to scale.

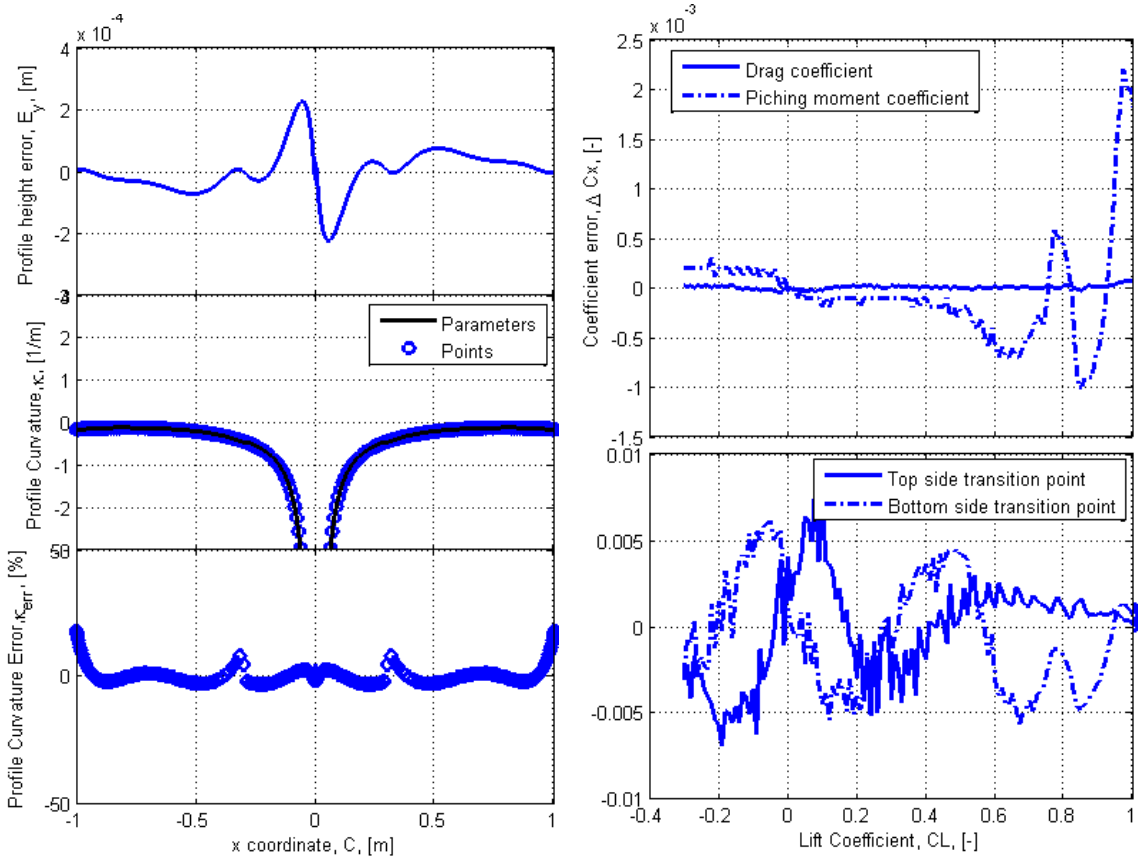


Figure 8: Geometrical and aerodynamic similarity errors between the two representations of the NACA 0012. To the left, comparisons between the vertical position of the points and the profile curvature. To the right a comparison between the aerodynamic properties of the two airfoils.

The Clark-Y airfoil shown in figure 9 rendered an error of **1.42 ects**, which should be also considered good, especially since this airfoil has a challenging straight lower side flank. The geometrical similarity distribution shown in figure 10

Maximum Thickness	*0.12 @ 30 %C*
Maximum Camber	*0.03 @ 42 %C*
Trailing Edge Gap	*0.12 %C*
Upper nose radius	*0.2 %C*
Lower nose radius	*2.1 %C*
Boat-Tail Angle	*15.74 deg*
Release Angle	*5.72 deg*

Table 4: Clark Y geometrical properties.

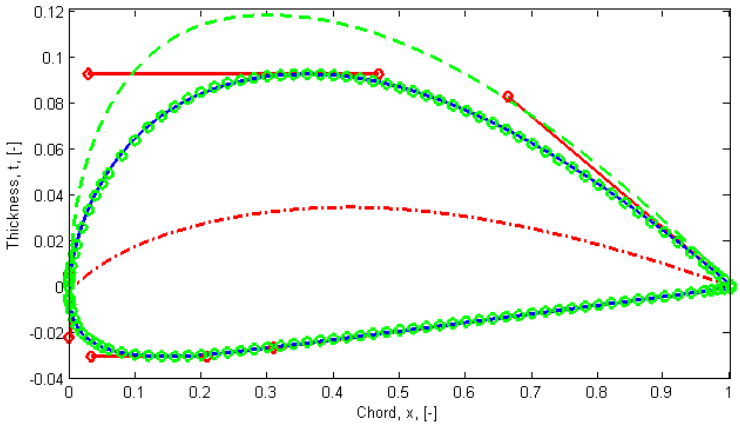


Figure 9: Clark Y profile. Original points and parametric curve overlaid. Height and width not to scale.

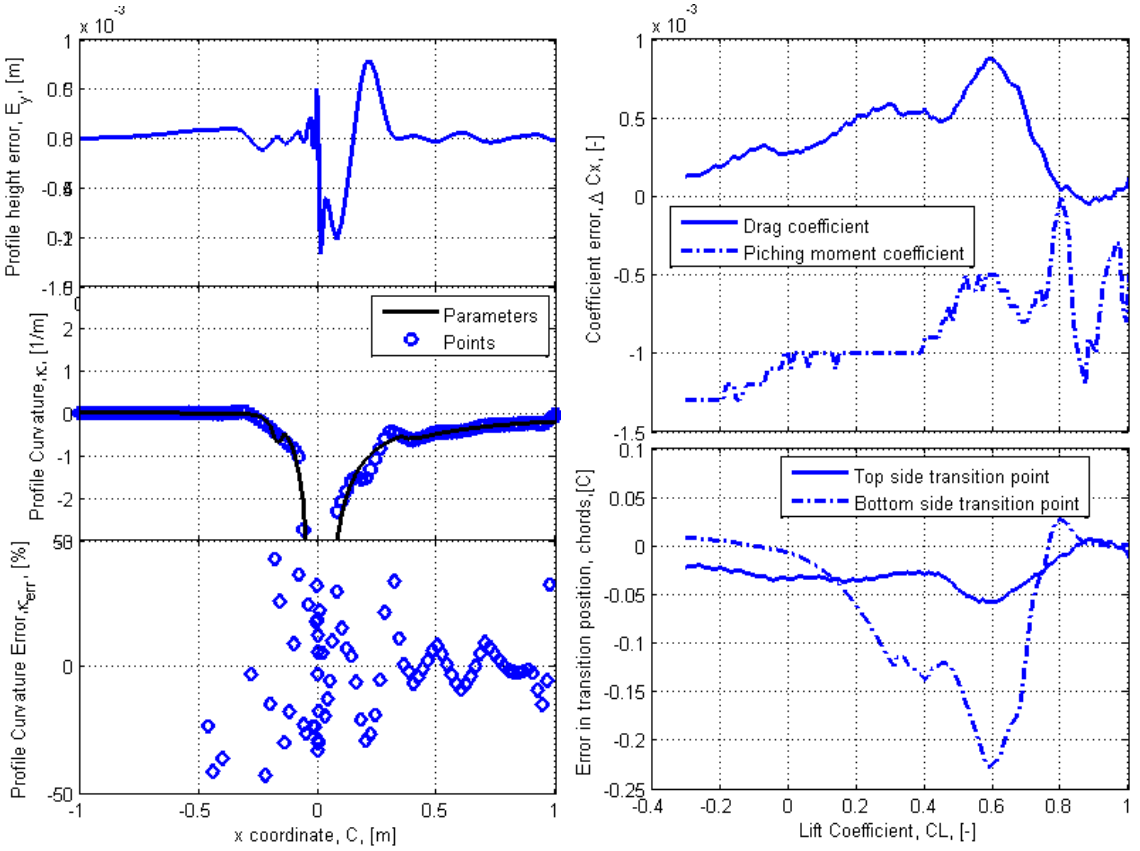


Figure 10: Geometrical and aerodynamic similarity errors between the two representations of the Clark-Y. To the left, comparisons between the vertical position of the points and the profile curvature. To the right a comparison between the aerodynamic properties of the two airfoils.

The NASA 747A415 profile shown in figure XX had fewer original datapoints than the earlier profiles, but was still possible to model with **3.98 ects**. The geometrical similarity distribution shown in figure 11 shows an error spike at the leading edge. This is where the optimization is the most sensitive to numerical errors.

Maximum Thickness	*0.15 @ 37 %C*
Maximum Camber	*0.03 @ 34 %C*
Trailing Edge Gap	*0 %C*
Upper nose radius	*2.5 %C*
Lower nose radius	*0.7 %C*
Boat-Tail Angle	*8.13 deg*
Release Angle	*2.74 deg*

Table 5: 747A415 geometrical properties.

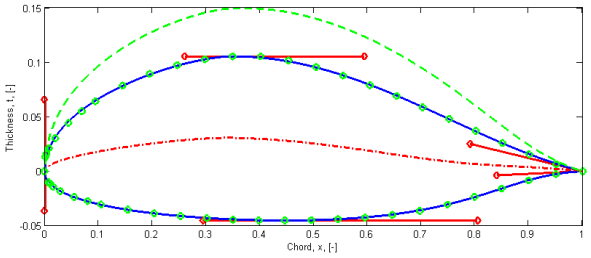


Figure 11: 747A415 profile. Original points and parametric curve overlaid. Height and width not to scale.

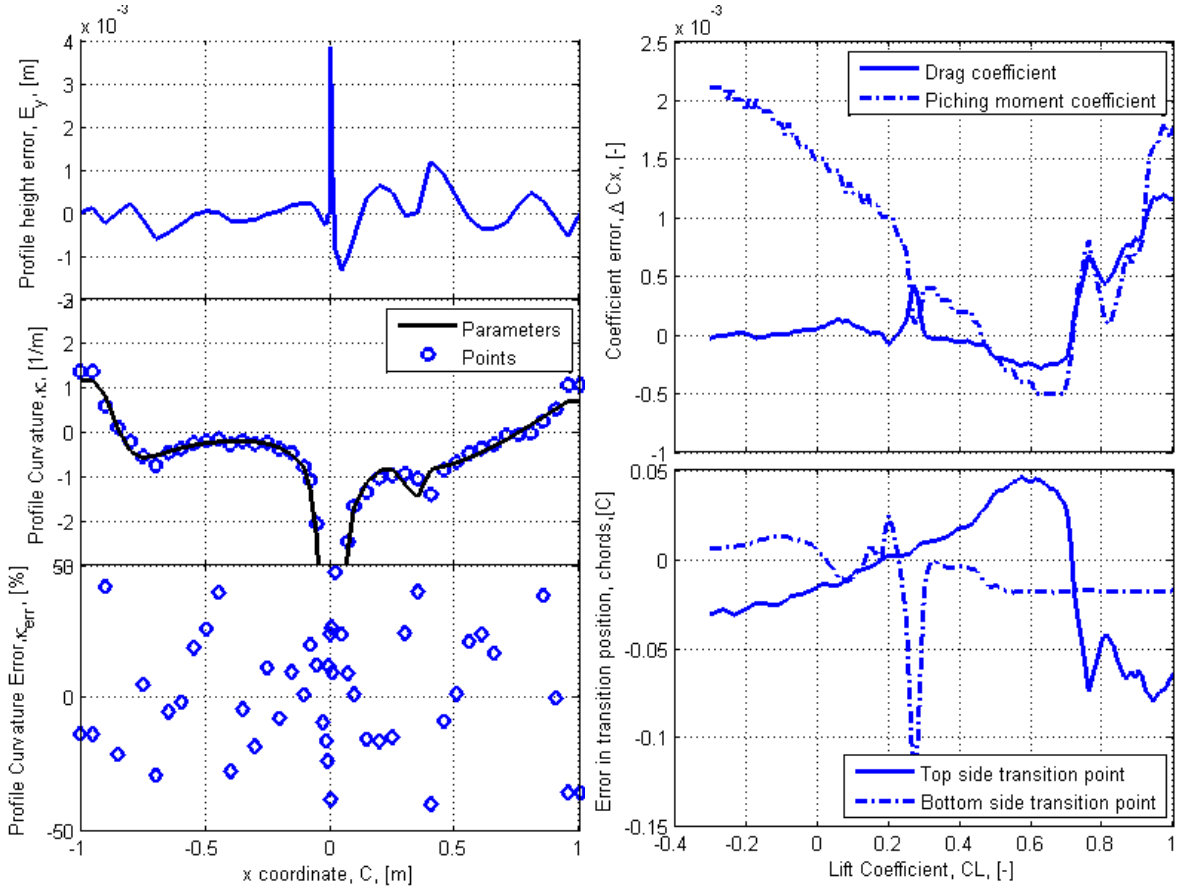


Figure 12: Geometrical and aerodynamic similarity errors between the two representations of the Nasa 747A415. To the left, comparisons between the vertical position of the points and the profile curvature. To the right a comparison between the aerodynamic properties of the two airfoils.

The supercritical Whitcomb profile was possible to emulate with the parameterization with **5.49 ects**. As shown in figure 13, the largest error was at the trailing edge lower side where the original dataset a discontinuous curvature. Still the aerodynamic similarity was good with drag being within 5 drag counts for most CL's.

Maximum Thickness	*0.11 @ 36 %C*
Maximum Camber	*0.02 @ 84 %C*
Trailing Edge Gap	*0.05 %C*
Upper nose radius	*2.4 %C*
Lower nose radius	*2.6 %C*
Boat-Tail Angle	*4.67 deg*
Release Angle	*17.95 deg*

Table 6: Whitcomb geometrical properties.

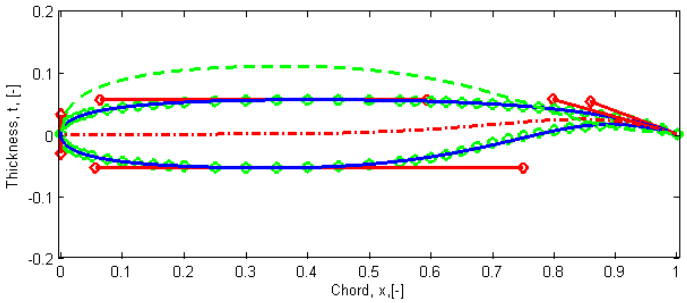


Figure 13:Whitcombe profile. Original points and parametric curve overlaid.

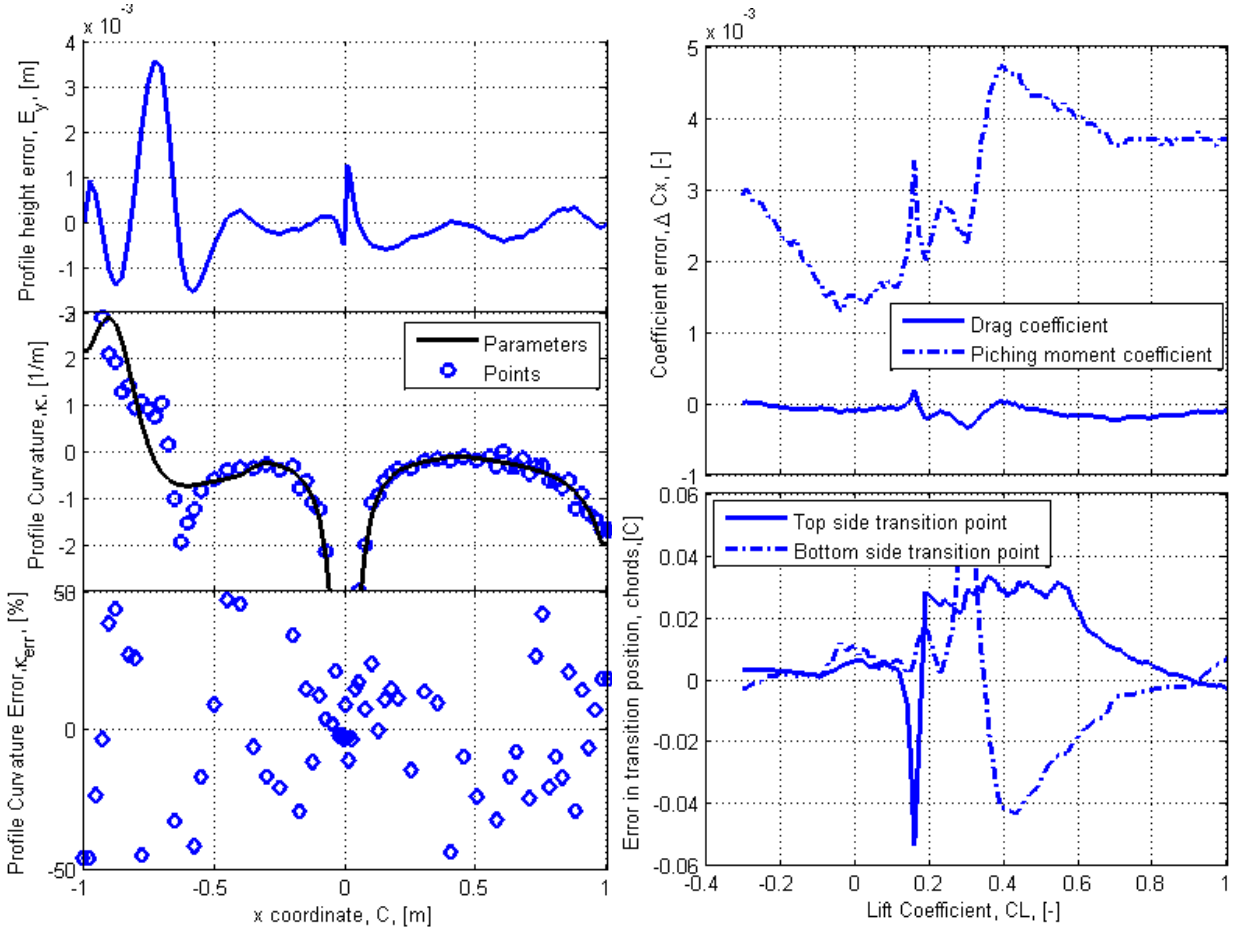


Figure 14: Geometrical and aerodynamic similarity errors between the two representations of the Whitcomb profile. To the left, comparisons between the vertical position of the points and the profile curvature. To the right a comparison between the aerodynamic properties of the two airfoils.

Figure 15 below show the pressure distribution on the top and bottom sides of the Naca747A415 profile at C_L =0.3 for both original data and parameterized airfoil .Additionally, the skin friction coefficient of the top side skin is shown for both airfoil representations. For both aerodynamic properties a good agreement between the two representations can be found.

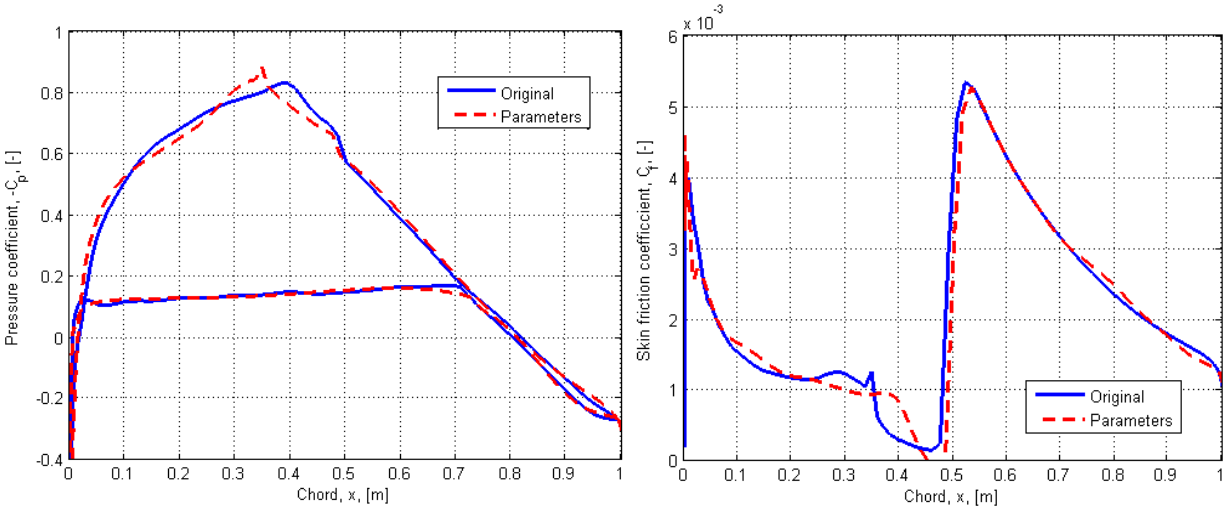


Figure 15: Pressure distribution and upper skin friction distribution of the two airfoil representations.

It's also interesting to look at a larger population of airfoils. 28 different classical airfoils have been parameterizes and analysed. The distribution of their geometrical model error is shown below in figure 16 for the profiles modeled so far. Most of the airfoils have an RMS error smaller than 0.0005, or half a millimeter to a meter chord.

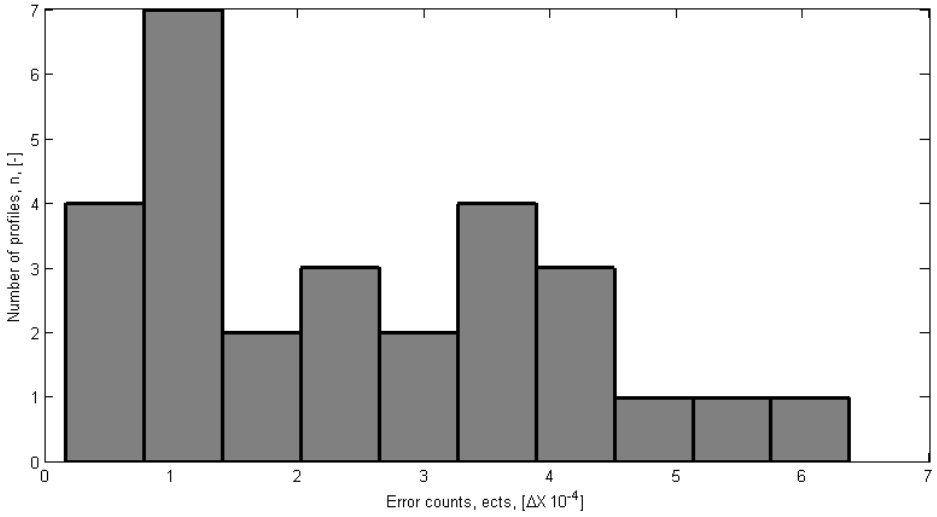


Figure 16: RMS error distribution for 28 parameterized profiles.

Conclusion

The proposed parameterization offers a stringent way of representing known airfoils without the problems of a finite number of coordinate points in a list or with the limitation of round-off and truncation errors in the number of decimals in the coordinates themselves.

Furthermore the proposed parameterization offers a way of continuously optimize an airfoil shape without the discrete changes occurring when going from one airfoil to the next.

Most known airfoils seem to be possible to parameterize with good geometrical and aerodynamic similarity with their original point-cloud representations.

Discussion

A further aerodynamic study should be performed to investigate any difference other aerodynamic phenomena. Notably stall angle of attack and post-stall behavior, which cannot be captured well with XFOIL. Likewise High speed shock formation similarities needs to be assessed before this parametric wing profile description can be adopted for wide scale airfoil interpolation.

The method as described does not require C2 continuity at the top and bottom points. Adding this constraint to the method will reduce the number of free parameters with two, and also give a more traditional curvature continuity. However, the geometrical fitting of known airfoils would suffer in having a higher geometrical error. –For most known airfoils tested, the optimized curvature is indeed discontinuous, but with a very small step. This discontinuity does not appear to have a significant impact on the aerodynamic performance of the parameterized airfoil.

References

[1] UIUC Airfoil Coordinates Database, http://www.ae.illinois.edu/m-selig/ads/coord_database.html

[2] E. N. Jacobs, K. E. Ward, & R. M. Pinkerton 1933 “The characteristics of 78 related airfoil sections from tests in the variable-density wind tunnel”, NACA Report No. 460

[3] Marzocca, Pier. "The NACA airfoil series", Clarkson University, Retrieved 07-03-2009

[4] R. Von Mises, ”Theory of Flight”, Dover Publications Inc., New York, 1959

[5] B. Kulfan, “Universal Parametric Geometry Representation Method”, AIAA Journal of Aircraft, vol.45 no.1, pp.142-158, 2008

[6] Louden, F A , “Collection of wind-tunnel data on commonly used wing sections”, NACA-TR-331, 93R20691,1930

[7] Drela, M, “XFOIL - An analysis and design system for low Reynolds number airfoils ”, Proceedings of the Conference, Notre Dame, IN; GERMANY, FEDERAL REPUBLIC OF; 5-7 June 1989. pp. 1-12. 1989.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

An Integrated Method for Determination of the Oswald Factor in a Multi-Fidelity Design Environment

Daniel Böhnke, Jonas Jepsen, Till Pfeiffer, Björn Nagel, Volker Gollnick

Air Transportation Systems, DLR e.V., Blohmstr. 18, 21079 Hamburg

Carsten Liersch

Institute of Aerodynamics and Flow Technology, DLR e.V., Lilienthalplatz 7, 38108 Braunschweig

Keywords: *Multi-Fidelity, CPACS, Distributed Design Environment*

Abstract

Aircraft conceptual design often focuses on unconventional configurations like for example forward swept wings. Assessing the characteristics of these configurations usually requires the use of physic based analysis modules. This is due to the fact that for unconventional configurations no sufficient database for historic based analysis modules is available.

Nevertheless, physic based models require a lot of input data and their computational cost can be high. Generating input values in a trade study manually is work-intensive and error-prone.

Conceptual design modules can be used to generate sufficient input data for physic based models and their results can be re-integrated into the conceptual design phase. In this study a direct link between a conceptual design module and an aerodynamic design module is presented. Geometric information is generated by the conceptual design module and the physic based results, in form of the Oswald factor, are then fed back.

Apart from the direct link, an equation for determination of the Oswald factor is derived via a Symbolic Regression Approach.

1 Introduction

At early stages of design large parts of the overall costs are fixed, since decisions have a major influence on the global configuration, [1]. It is therefore important to determine the characteristics of an aircraft as early and accurately as possible.

In a simplified treatment, analysis modules can be grouped into historic and physic based models. Historic based models are quick to evaluate due to their statistic approach and deliver reliable results for conventional configurations. The necessary inputs are few. Physic based models are, depending on the level of detail, more cost intensive to evaluate and require more input, e.g. a detailed geometry description. Nevertheless, these modules are more accurate, especially for unconventional configurations.

Hence, at the earliest stages of the design only historic based models can be applied, since the necessary input data for physic based models is not available yet. These models are however unsuitable for most new configurations since they are based on the existing design space, and can not be used for assessment in the diverging phases of design, [2]. At this point in time, unconventional configurations should already be a

part of the design space. In a simplified declaration, historic models are categorized in conceptual design and physic based models in preliminary design.

The question arises whether it is possible to close the gap between conceptual and preliminary design, i.e. historic and physic based analysis. Preferably this should be accomplished in an automated manner, to make the results of physic based analysis available at an earlier point of the design.

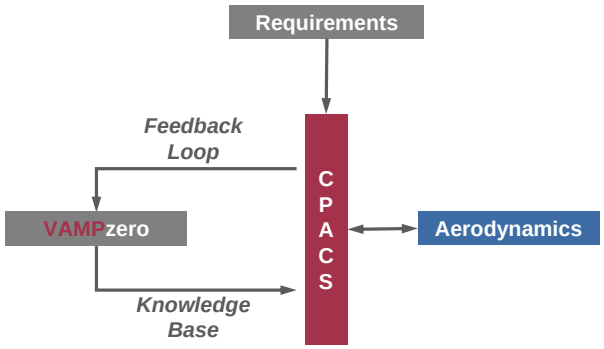


Fig. 1: Multi-Fidelity Process

Within this study a central model approach (CPACS) is proposed that combines a historic based conceptual design module (VAMPzero) with a physic based aerodynamic module (LIFTING_LINE), [3, 4]. As use-case the calculation of the Oswald factor is chosen. On a first step, a set of requirements is defined and fed into the conceptual design tool. With knowledge based engineering techniques a detailed geometry definition is derived from the conceptual design and stored in the central model data format. The aerodynamic module is linked to the central model and triggered without further user input. The resulting aerodynamic performance map is subsequently reduced to the Oswald factor and reused within the conceptual design module. Provided that a sufficient number of trades is carried out, a symbolic regression approach, [5] is applied to derive new conceptual design methods from the multi-fidelity loop, as displayed in Figure 1.

The paper is grouped into six sections: This section gives a short introduction. Section 2 elaborates on existing methods for the determination for the Oswald factor in conceptual design. The proposed design environment is described in Section 3. A trade study is then carried out and examined in Section 4. Finally, a discussion of the results is given in Section 5 and some conclusions and an outlook in Section 6.

2 State of the Art

In this section an overview is given for the calculation of the Oswald factor (e) in conceptual and preliminary design. The Oswald factor is needed for the estimation of the lift induced drag (C_{D_I}), whereas drag due to lift (C_{D_L}) in this paper is meant to take into account lift induced drag and drag due to viscous forces (C_{D_V}). The calculation of the wave drag (C_{D_W}) is not considered any further. The lift induced drag in conceptual design is usually noted as in Eq. 1:

$$e = \frac{1}{\pi AR}\frac{C_L^2}{C_{D_I}} \tag{1}$$

Additionally, the drift for the zero lift angle of attack (C_{L_0}) needs to be taken into account. Furthermore, local lift distributions may result in lift induced drag even if the overall lift sums up to zero, e.g. for twist distributions. Therefore a parameter for induced drag at zero lift ($C_{D_{I0}}$) is introduced. Equation 1 is then extended to Eq. 2:

$$e = \frac{1}{\pi AR}\frac{(C_L - C_{L0})^2}{(C_{D_I} - C_{D_{I0}})} \tag{2}$$

The first outlined approach originates from Raymer and is based on the aspect ratio (AR) and the leading edge sweep angle of the wing (Λ), [6]. Two different equations valid for unswept (Eq. 3) and swept (Eq. 4) wings are defined.

$$e = 1.78(1 - 0.045AR^{0.68}) - 0.64 \tag{3}$$

$$e = 4.61(1 - 0.045AR^{0.68})(cos(\Lambda))^{0.15} - 3.1 \tag{4}$$

$$with \quad \Lambda > 30$$

These equations can be evaluated quickly as the number of variables is low. Drawbacks exist because the design space is limited to a specific area via the limitations on the leading edge sweep angle.

An approach by Howe (Eq. 5) takes into account additional parameters like the flight Mach

number (M_N), the taper ratio (λ) and the thickness to chord ratio (t/c), [7]. The equations derived by Howe are defined for drag due to lift and it is hard to extract an exact formulation for the Oswald factor. Nevertheless, they are included in this overview to display complex symbolic equations known in conceptual design as we will derive such an equation at a later point of this study. The functions for the taper ratio and aspect ratio are given in Eq. 6 and Eq. 7:

$$C_{D_I} = \frac{(1 + 0.12M_N^6)}{\pi AR} \left[1 + g(AR) + \frac{0.142 + ARf(\lambda)(10t/c)^{0.33}}{cos(\Lambda)^2}\right] \quad (5)$$

$$f(\lambda) = 0.005\left[1 + 1.5(\lambda - 0.6)^2\right] \quad (6)$$

$$g(AR) = \frac{0.1}{(4 + AR)^{0.8}} \quad (7)$$

A semi physic based approach is described by Jenkinson for the induced drag utilizing lifting-line theory, [8, 9]. The approach leads to a function (C_1) depending on the taper ratio and aspect ratio, as shown in Figure 2.

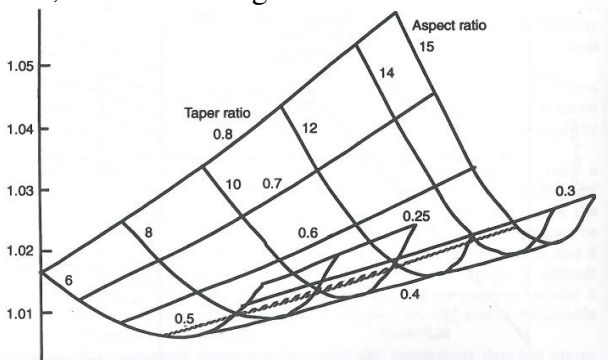


Fig. 2: C_1 depending on λ and AR, from [8]

Subsequently, the results are corrected for conventional (Eq. 8) and CFD (Eq. 9) wing design so that e results in C_2 over C_1. For the change in drag due to lift due to viscous forces a formulation is used that is based on the zero lift drag.

$$C_2 = 1.235 - 0.0245AR \quad (8)$$

$$C_2 = 1.113 - 0.0116AR \quad (9)$$

The results from Jenkinson already implement a physic based model. In this study this appraoch will be extended by a set of relvant vaiables with their corresponding formulation.

3 Design Environment

In a generalized manner, a distributed design environment consists of three components. The basis for all communication is a *common language* that is achieved via a central model approach. The second component is a set of *analysis modules* that can exchange information using the common language. The benefits of using a central model are depicted in Figure 3. Due to the fact that communication now only exists between the analysis modules and the central model the number of interfaces decreases. Additionally, interaction between the analysis modules is decreased in a way that different modules can be utilized for similar tasks. For example a fuselage weight prediction can be made using a historic based or a physic based, either beam or shell, model. Regardless from the chosen analysis strategy, the fuselage mass is included in the overall design and available for further modules.

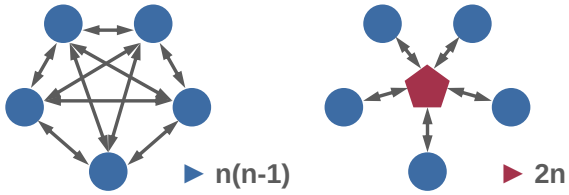


Fig. 3: Central Model Approach

This process of communication is guided by an *engineering framework*. The engineering framework also takes over general tasks like process control and optimization. As can be seen from Figure 1, in the current study the Common Parametric Aircraft Configuration Schema (CPACS) is used as the central model and VAMPzero and LIFTING_LINE as the analysis modules. For the engineering framework the Remote Component Environment (RCE) is used.

3.1 CPACS

Since 2005 CPACS is under development at DLR, [3]. It is a XML-based definition of the air transportation system. Historically, the fo-

cus is set on aircraft nevertheless the bandwidth of applications spreads from conceptual and preliminary aircraft design to climate impact modeling and engine design, [4, 10, 11]. Using a central model enables an easy communication for designers as the underlying model stays the same. Additionally, tool development becomes more efficient as core tasks can be unified. For example CPACS comes with supplementary libraries for geometry handling and visualization as well as CPACS specific XML functionalities. Furthermore, XML toolboxes are available for most programming languages both commercial and open source. One important feature of CPACS is the handling of product and process information, i. e. not only the aircraft is described but also *tool-specific* information to control parts of the analysis process, e. g. the number of panels on a wing for aerodynamic analysis.

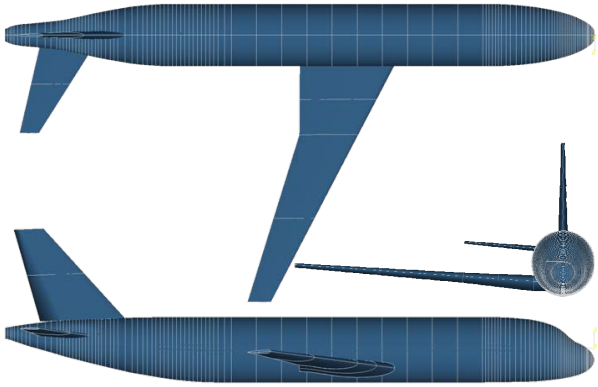


Fig. 4: Baseline Geometry

In Figure 4 a sample geometry in CPACS is displayed. This geometry is generated by VAMPzero (s. Section 3.3) and is the baseline for the trade study (s. Section 4). It shows a conventional short range configuration in a quasi-symmetrical view.

3.2 LIFTING_LINE

In the aerodynamic analysis the CPACS dataset is imported via an interface tool to trigger the multiple lifting line program LIFTING_LINE, [9]. The interface from LIFTING_LINE to CPACS and the interaction with other analysis modules is outlined in more detail by Liersch et. al. in [3]

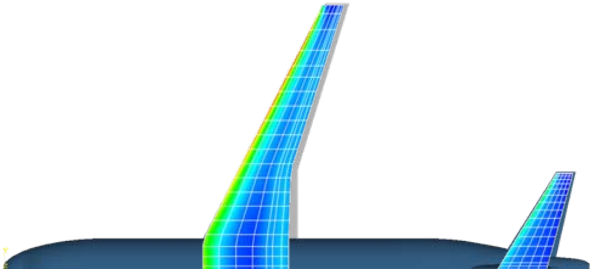


Fig. 5: Lift Distribution for Baseline Aircraft

LIFTING_LINE enables a calculation of the lift induced drag from which the Oswald factor is derived. A lift distribution for the baseline geometry is depicted in Figure 5.

Additionally, the outlined research benefits from the low computing cost of a single computation. Calculation times from LIFTING_LINE are in the order of few seconds and only minor overhead occurs due to file transfer and model transformation.

3.3 VAMPzero

For the purposes of DLR's distributed multi-disciplinary design environment a new conceptual design module is developed. VAMPzero is based on handbook methods taken from well-known design literature, [6, 7, 12]. The key modeling aspects of VAMPzero are described in more detail in [4].

In a design environment that is mostly based on preliminary level, two possible tasks remain for a conceptual design tool. The first one is the initialization of the datasets. Designers usually need to start working with only the top level aircraft requirements (TLAR) at hand. In this case the conceptual design code can close the gap between the TLAR and sufficient information to trigger preliminary analysis modules.

To overcome the gap, the results from the conceptual design are transferred into CPACS via a knowledge based engineering approach. For the conceptual design calculations and sizing VAMPzero uses a single trapez wing shape. For the export the wing shape can be mapped to a double trapez with an additional rectangular section up to the wing-fuselage intersection. Figure 6 shows the geometry. The wing planform is described by the wing span (b_W), the leading

edge sweep, the aspect ratio, the taper ratio, and the kink ratio (η_k). The trailing edge from root to kink is perpendicular to the fuselage. The twist angle (ϑ) is applied in a linear distribution from root to tip. Thickness to chord ratios at root and tip stations are a result from VAMPzero. For the empenage symmetrical airfoils are used. The wing geometry is generated with NACA-6 series airfoil. Further airfoil sets and their assortment is an ongoing topic within the development of VAMPzero.

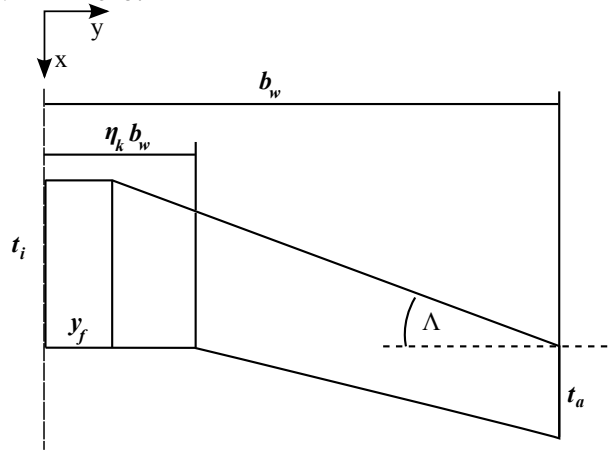


Fig. 6: Double Trapezoid Wing

Using the Π-Theorem by Buckingham dimensionless parameters for the description of the geometry can be defined as in Table 1 to describe the introduced wing geometry [13, 14]. During the mapping to CPACS Π_1 to Π_4 are constant to assure geometric similarity to the single trapez wing. Π_5 and Π_6 are then derived for a constant wing reference area ($S_{ref} = f\left(t_i, t_a, b_W, \eta_f, \eta_k, \varphi\right)$) and the fuselage diameter.

Table 1: Dimensionless Constants

$\Pi_1 = \Lambda$	$\Pi_2 = \vartheta$	$\Pi_3 = \lambda = \frac{t_a}{t_i}$
$\Pi_4 = AR = \frac{4b_W^2}{S_{ref}}$	$\Pi_5 = \eta_k = \frac{y_k}{b_w}$	$\Pi_6 = \eta_f = \frac{y_f}{b_w}$

The fuselage shape is derived from a baseline configuration that is afterwards scaled to fit to the results from the conceptual design. The empenage is exported *as is* as single trapez surfaces. In this way the conceptual design tool can generate geometries as shown in Figure 4 and all additional data to trigger a LIFTING_LINE analysis.

The second task for a conceptual design tool in a preliminary design environment is the integration of results. Results from several higher level analysis modules need to be aggregated to obtain the effects on the overall aircraft configuration. For example, the wing mass in the mass breakdown is overwritten by higher level analysis module (e.g. FEM) and VAMPzero takes these changes into account when re-calculating the aircraft. In the case of this study, VAMPzero interprets the results from LIFTING_LINE to derive the Oswald factor and calculate a new configuration.

3.4 RCE

VAMPzero and LIFTING_LINE exchange information via CPACS. To establish this connection also in a distributed environment where parallelization, optimization and design of experiments are possible an engineering framework is needed. RCE is a general engineering framework that is specialized for this task via the Chameleon tool suite, [15, 16]. Chameleon orginiates from the beginnings of CPACS development within the Technology Integration for the Virtual Aircraft (TIVA) project and is now integrated into RCE.

RCE offers a graphical user interface to connect analysis modules and run different stages of the design process. As a first step a workflow is set up and is sent to executing computers. All information created during the analysis, e.g. intermediate results, tool-specific outputs, transferred datasets, is available via an integrated data management system.

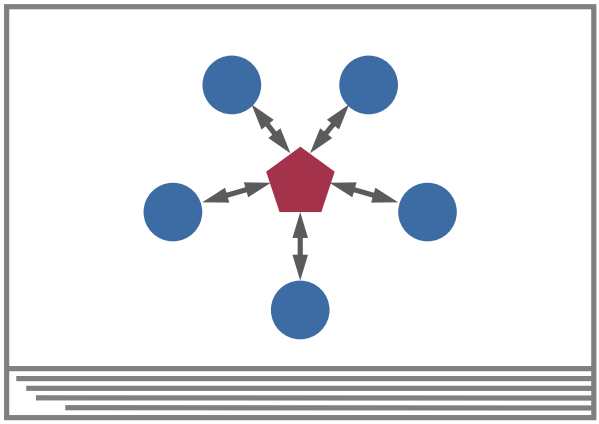


Fig. 7: Distributed Design Environment

The integrated Chameleon tool suite enables disciplinary experts to wrap their tools in the distributed design environment. Global operations like file transfer, remote execution and XML-mappings, as well as local operations like the interpretation of CPACS files are supported by Chameleon. Addtionally, a geometric library offers CPACS-specific operations to extract information from datasets.

3.5 Eureqa

The goal of this study is to derive a method for determination of the Oswald factor that is both accurate and available at low computational cost. Symbolic regression is a suitable tool to find equations from experimental data, [5].

The input for Eureqa, the symbolic regression toolbox used in this study, consists of a set of input variables along with the corresponding experimental results. Additionally, mathematical operators are specified that shall be used in the equation. Several combinations of operators and variables are then generated by a genetic algorythm to find a suitable approximation in form of a symbolic equation. Derived equations are rated for their complexity and fit.

In the current study Eureqa is used as a post processor. Future research might include it in the calculation loop. In this way Eureqa can define new design points where the current solution does not fit to the design space. Therefore the number of necessary runs could be decreased.

4 Trade Study

The baseline configuration for this study is a single aisle ($nPAX = 150$), medium range ($R = 3000km$, $Mach = 0.78$) aircraft.

To determine the effects of different parameters on the lift induced drag, design variables are varied as shown in Table 2. Due to the fact that the number of dimensions and the range of the variables are large, a full factorial design is unsuitable. Instead the design space is explored by Latin-Hypercube sampling. Apart from the specified values for the baseline configuration the reference area for the wing is kept constant.

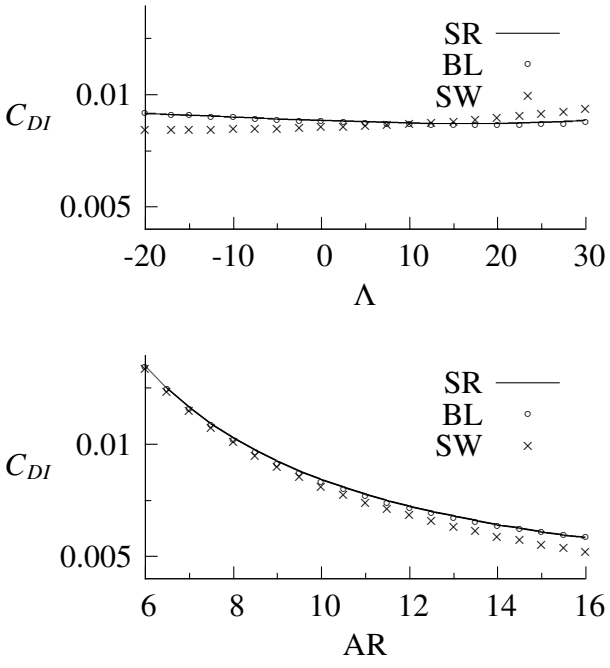
Table 2: Design Space

Variable	Symbol	Range
Sweep Angle	Λ	-20.0 - 30.0
Aspect Ratio	AR	6.0 - 16.0
Taper Ratio	λ	0.1 - 0.6
Twist Angle	ϑ	-10.0 - 0.0
Kink Ratio	η_k	0.2 - 0.4

In a distributed design environment one has to cope with the drawbacks of data transfer and interpretation. This leads to some overhead in the calculation time. Nevertheless, more time is spent on the symbolic regression and post processing as the runtimes of VAMPzero and LIFTING_LINE are low.

The trade study is grouped into two parts. In the first part design variables are varied independently from each other. These calculations are used as control data and the analysis of the effects from the different input variables benefits from the transparency of this setup.

The second part of the study consists of a Latin Hypercube sampling in the complete design space. Subsequently, a symbolic regression approach is applied to the results to derive an equation for determination of the Oswald factor. The results of the symbolic regression are then compared to the control data.



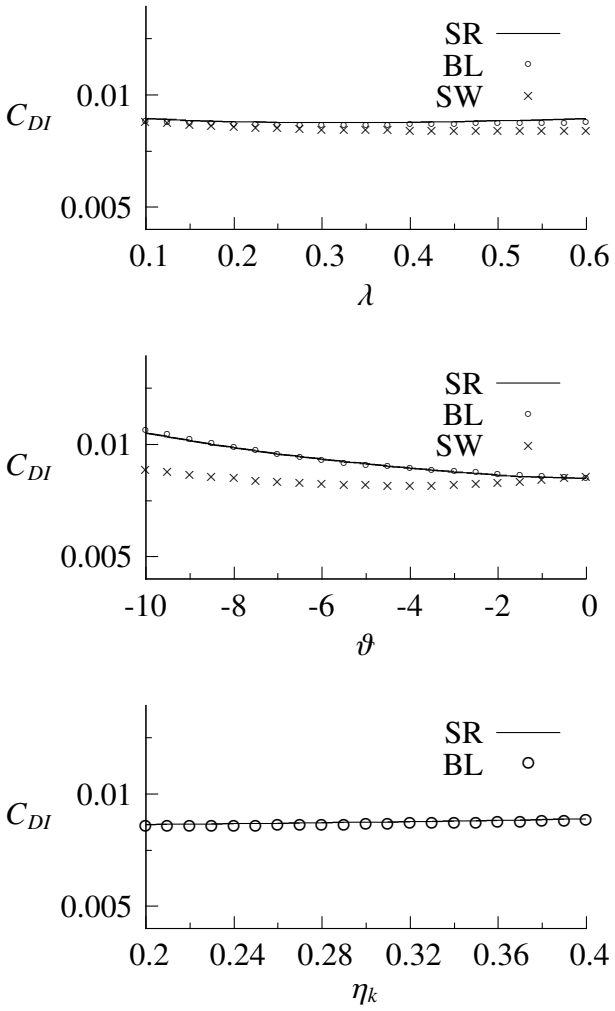


Fig. 8: Trade Study Results

In the first part the baseline parameters ($\Lambda = 28$, $AR = 9.5$, $\lambda = 0.2$, $\vartheta = -2$, $\eta_k = 0.35$) are kept constant and only one design variable is varied. The results are displayed as baseline (*BL*) in Figure 8. Additionally, a set of the trades was carried out with a single trapez geometry. The parameters are varied in the same design space but the wing is unswept and untapered, i.e. $\Lambda = 0$, $AR = 9.5$, $\lambda = 1$, $\vartheta = 0$. For the singe trapez the kink parameter is ignored. The results are marked as (*SW*) in Figure 8. All plots are made on the same scale to illustrate the impact on the lift induced drag in comparison. Displayed values are shown for a constant lift coefficient ($C_L = 0.5$). These parts of the trade study serve as the control data for the derived equation from the Latin-Hypercube sampling.

The second part consists of the previously named Latin-Hypercube sampling. A number of 500 samples throughout the complete design space is calculated. The data from this sampling is fed into a symbolic regression toolbox to derive a symbolic formulation for the lift induced drag in dependence on the design variables.

For comparison of the method these simplified graphs are shown as *SR* in Figure 8. The symbolic formulations are noted in Eq. 10. The solution derived from the Latin-Hypercube sampling is valid in the full range of the design space.

5 Discussion

This chapter contains a discussion of the results found from Figure 8 as well as more general review of the introduced design environment and calculation strategy.

For the wing sweep the minimum for the lift induced drag for the single trapez is at a forward swept position. The double trapez minimum lies at approximately $\Lambda = 23$. The difference in the results can be explained by the additional taper for the double trapez wing that already shifts the load distribution in spanwise direction.

The results for the aspect ratio both for the single and the double trapez wing show similar trends. As the results are plotted for the lift induced drag the influence of the aspect ratio is increased, see Eq. 1, where the Oswald factor is now also a function of the aspect ratio. At medium aspect ratios the derived equation overpredicts the lift induced drag.

Curves for the taper ratio align at smaller taper ratios. For the unswept wing the minimum lift induced drag is located at taper ratios larger than the considered design space. The minimum for the unswept single trapez lies within the design space but sensitivity is low compared to the aspect ratio. For larger taper ratios the derived equation produces higher values of lift induced drag.

Twist is applied linear on both wings. For the double trapez there are three twisted sections: fuselage section, kink section and tip section. The single trapez wing is only twisted at the tip. For the double trapez the lift induced drag can not be decreased by additional twist. A coupling with a structural tool would give more de-

$$e = 0.04 - 0.0007AR - 0.00019\Lambda\vartheta + \lambda^{0.03}\, cos(0.16 - 0.0007AR\Lambda - 0.0007AR\Lambda\eta - 0.55\lambda) \qquad (10)$$

tails on the benefits from twist due to load reduction at the tip. The single trapez wing can still benefit from the twist. The fitted curve is in accordance with the baseline. The twist has a high influence on $C_{D_{I0}}$. The aerodynamic (in)efficency can therefore not be described by the Oswald factor alone.

As already mentioned, the change in the kink ratio only applies to the double trapez wing. The more the kink moves outward the higher the lift induced drag. This is due to the fact that the inward area increases rapidly in comparison the outer part of the wing. The symbolic regression constantly over estimates the lift induced drag.

The plots show a good agreement between the derived equations and the single parameter variations. The root mean squared error for the Oswald factor is 0.011. This gives confidence to implement the derived equation (Eq. 10) into the conceptual design tool VAMPzero. As all parameters that are needed to generate the higher level geometry are avaiblable in VAMPzero no additional customizations of the code are necessary.

The outlined study focuses on the interaction of analysis modules on different levels of fidelity. As it was desired to derive a new calculation method in a large design space the number of design evaluations was high. In this case the study benefitted from the fact that LIFTING_LINE offers a physic based approach at low computational cost. It is questionable at which cost a similar appraoch is possible with higher detail models, e.g. CFD.

Integration of higher level analysis modules often is done using response surface models. In this way the computational time is kept low, as most of it is spent a-priori. Within the outlined method this approach is possible as seen from the derived equation. Additionally, the direct link between VAMPzero and the disciplinary analysis module LIFTING_LINE is existent and needs to be triggered as soon as the design space is enlarged

It is shown that the derived equation from the symbolic regression approach fits well to the control data. It is assumed that reliable results can be obtained within the area of the design space. A detailed assumption on the validity of the equation outside of the design space is not yet given.

Finally, it is possible to automize large parts of the analysis chain. Nevertheless the extraction of knowledge is not yet possible in a similar way. This is due to practical reasons, as the process works also well regardless of faulty data. For validation and verification the *Engineer in the Loop* is a necessity.

6 Conclusion

The current study outlines a methodology to derive the Oswald factor in a multi-fidelity environment. This methodology can be extended for different use cases targeting single higher fidelity analysis modules, e.g. for mass estimation. For future research the target is to analyze couplings of higher fidelity tools, as for example in aeroelastics, and try to reflect these couplings also on the level of a simple analysis module such as VAMPzero. These coulings are not necessarily limited to different disciplines but may also couple analysis modules of different fidelity levels.

Within the outlined approach the *curse of dimensionality* is of crucial importance, [17]. Due to the fact that the number of design variables is usually beyond two, the number of calculation runs increases rapidly. Determining the number of required runs for a good fit is currently only a smart guess. Integrating the symbolic regression toolbox into the design loop may give the chance to optimize this process. The symbolic regression algorithm is capable of deriving new experiments at locations in the design space where the current error is high. Starting from a small number of calculation, e.g. a full factorial design only on the outer limits of the design space, the design loop could be configured until the overall error reaches a certain minimum.

Acknowlegdements

The authors would like to thank Markus Litz, Markus Kunde, Arne Bachmann and Doreen Seider from DLR Simulation and Software Technology for their support on the engineering framework and Chameleon. Additionally, the authors would like to thank Markus Schmidt and Hod Lipson from Cornell University for making the Eureqa toolbox available.

References

[1] Stark, J., "Product Lifecycle Management , 21st Century Paradigm for Product Realisation," *Springer London*, 2005.

[2] La Rocca, G. and van Tooren, M., "Enabling distributed multi-disciplinary design of Complex Products : a knowledge based Engineering Approach," *Journal of Design Research*, Vol. 5, 2007, pp. 333–352.

[3] Liersch, C. and Hepperle, M., "A Unified Approach for Multidisciplinary Aircraft Design," *CEAS European Air and Space Conference*, 2009.

[4] Böhnke, D., Nagel, B., and Gollnick, V., "An Approach to Multi-Fidelity in Conceptual Aircraft Design in Distributed Design Environments," *IEEE Aerospace Conference*, 2011.

[5] Schmidt, M. and Lipson, H., "Distilling Free-Form Natural Laws from Experimental Data," *Science*, Vol. 324, 2009, pp. 81–85.

[6] Raymer, D., "Aircraft Design: A Conceptual Approach," *AIAA Education Series*, 1989.

[7] Howe, D., "Aircraft Conceptual Design Synthesis," *Professional Engineering Publishing*, 2000.

[8] Jenkinson, L., Simkin, P., and Rhodes, D., "Civil Jet Aircraft Design," *Butterworth Heinemann*, 1999.

[9] Horstmann, K. H., "Ein Mehrfach-Traglinienverfahren und seine Verwendung für Entwurf und Nachrechnung nichtplanarer Flügelanordnungen," *DFVLR-FB 87-51*, 1981.

[10] Zill, T., Böhnke, D., Nagel, B., and Gollnick, V., "Preliminary Aircraft Design in a Collaborative Multidisciplinary Design Environment," *AIAA Aviation Technology, Integration and Operations Conference*, 2011.

[11] Koch, A., Lührs, B., Dahlmann, K., Linke, F., Grewe, V., Plohr, M., Nagel, B., Gollnick, V., and Schumann, U., "Climate impact assessment of varying cruise flight altitudes applying the CATS simulation approach ," *CEAS European Air and Space Conference*, 2011.

[12] Roskam, J., "Airplane Design," *DARCorporation*, Vol. I-VII, 1989.

[13] Buckingham, E., "On Physically Similar Systems; Illustrations of the Use of Dimensional Equations," *Physical Review*, 1914, pp. 345–376.

[14] Jepsen, J., "Wissensbasierte Erstellung von Geometrien auf der Grundlage von Handbuchverfahren," *IB-328-2011-07, TU Hamburg-Harburg*, 2011.

[15] Litz, M., Bachmann, A., and Kunde, M., "TIVA XML Interface," *tixi.googlecode.com*.

[16] Litz, M., Bachmann, A., and Kunde, M., "TIVA Geometric Library," *tigl.googlecode.com*.

[17] Koch, P. N., Simpson, T., Allen, J. K., and Mistree, F., "Statistical Approximations for Multidisciplinary Design Optimization: The Problem of Size," *Journal of Aircraft*, Vol. 36, 1999, pp. 275–286.

CEAS 2011

UAS 2

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

All electric Secondary Power System for future advanced MALE UAV

S. Chiesa
Department of Aeronautics and Space Engineering, Politecnico di Torino, Italy

S. Farfaglia
Alenia Aeronautica S.p.A., Italy

N. Viola and M. Fioriti
Department of Aeronautics and Space Engineering, Politecnico di Torino, Italy

Keywords: *all electric secondary power system, MALE UAV*

Abstract

The paper deals with the design of the all electric Secondary Power System of a future advanced MALE UAV (Fig. 1 and Fig. 2), which weights less than 4000 kg, flies at 14000 m of altitude and has an endurance of more than 30 hours. Two main modes of operations have been envisaged for the aircraft: the normal mode of operation, during which it performs a typical surveillance and monitoring mission, and the emergency 1 mode of operation, that lasts up to 2 hours, thanks to the aircraft high aerodynamic efficiency. The paper presents the refined Hybrid Fuel Cells configuration of the Secondary Power System of the reference MALE UAV, named SAvE aircraft. The results obtained from the virtual tests within the Integrated Simulation Environment are shown and discussed.

1 Introduction

SAvE (Systems for UAV Alternative Energy) is a research project, that was funded in 2007 by Piemonte Regional Government, Italy, and was assigned to Politecnico di Torino and Alenia Aeronautica. The project started in October 2007 and ended in April 2011. Aim of the project has been the study of new, more efficient, more effective and more environmentally friendly on board systems for future advanced Unmanned Aerial Vehicles (UAV), particularly for future advanced Medium Altitude Long Endurance (MALE) UAVs. The Secondary Power System, applying the all-electric philosophy, has been the primary field of research of the SAvE project. However, all other on board systems have also been considered either because of their close connection with the Secondary Power System or because they represent crucial interfaces, thus being a source of requirements and constraints.

The paper deals with the design of the all electric Secondary Power System of the SAvE aircraft, a future advanced MALE UAV, which weights less than 4000 kg, flies at 14000 m of altitude and has an endurance of more than 30 hours.

The reference mission profile is illustrated in Fig. 1. The 3D CAD model of the aircraft [1], that includes the structures, the engines, the landing gears and the fuel system, is shown in Fig. 2, while Fig. 3 depicts the installation of

avionics equipment and of all electric secondary power system components of the Hybrid Fuel Cells configuration of the SAvE aircraft.

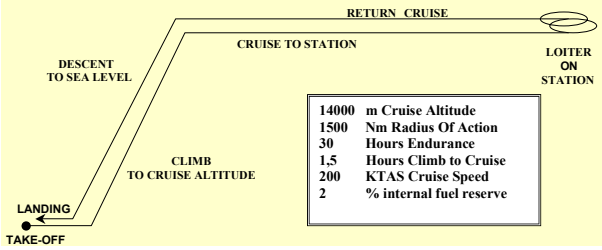


Fig. 1. SAvE reference mission profile

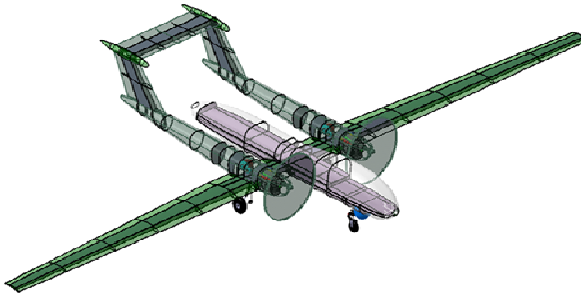


Fig. 2. 3D CAD model of SAvE reference aircraft

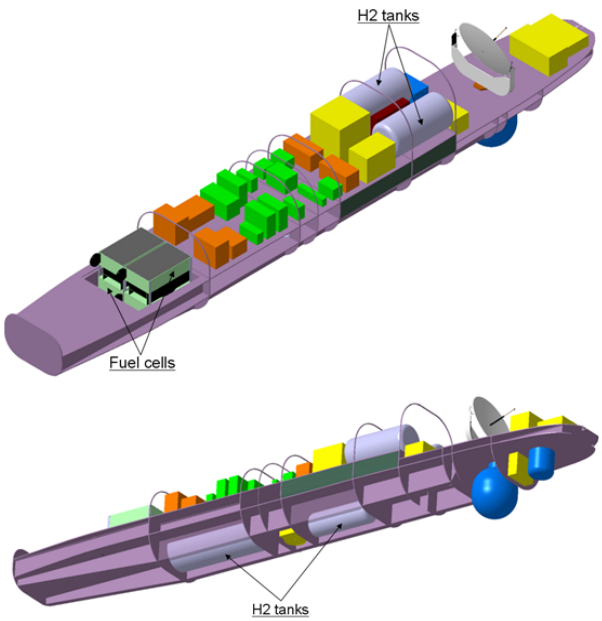


Fig. 3. Installation of the avionic equipment and all electric secondary power system components: Hybrid Fuel Cells configuration

Taking into account the different mission phases of the reference mission profile (engine start, taxing, take-off and first climb, climb 3000-7000 m, climb 7000-14000 m, cruise, descent 14000-7000 m, descent 7000-3000 m, approach and landing) and the various avionics equipment used in each mission phase, depending on the considered mode of operation, the electrical loads per each mission phase can be estimated [2].

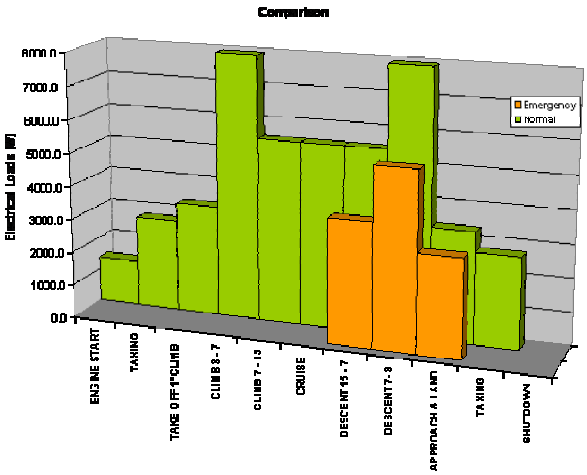


Fig. 4. Electrical loads per each mission phase for different modes of operation

Fig. 4 illustrates the electrical loads per each mission phase for different modes of operation: the normal and the emergency 1 mode of operation [3]. In normal mode the planned mission (surveillance) is regularly performed, whereas in the emergency 1 mode the planned mission is aborted and the aircraft starts immediately the descent. Apart from normal and emergency 1 mode of operation, also emergency 2 mode has been envisaged for the SAvE aircraft. Unlike emergency 1 mode of operation, which allows performing a controlled descent to a Dedicated Crash Site (DCS) or to approach an emergency landing, the emergency 2 mode allows performing only an uncontrolled descent to the Open Country (OC). In particular in emergency 1 mode of operation the glide duration increases, thus extending the time available (about two hours have been estimated from the cruise altitude) to choose the right spot and to make the aircraft land there safely, without damaging any people or infrastructures. As it can be noted from Fig. 4, peaks of electric loads are expected during climb from 3000 to 7000 m and during descent from 7000 to 3000 m. These peaks are mainly due to the possibility

of using the electric anti-ice system for the Pitot Static System and the propellers, besides the Goodrich system for the wing de-icing [4]. Eventually comparing the graphs for the normal and the emergency 1 modes, it can be highlighted that smaller electric loads are estimated for the emergency mode because some mission specific avionics equipment are turned off. Because of the high electric loads for quite a long time, the emergency is the crucial mode of operation of this aircraft and it has to be coped with innovative technologies.

Before describing the secondary power system, it is worth saying that the SAvE aircraft shall be designed with a cumulative probability to lose the system of $5*10^{-6}$ per flight hours or less, which is compliant with the standard Italian military requirement. The architecture of the secondary power system shall have an adequate level of redundancy to approach the safety design objective of $5*\ 10^{-8}$ per flight hours, as a consequence of the above mentioned safety requirement.

After a thorough trade-off analysis between alternative configurations of the all electric Secondary Power System, the Hybrid Fuel Cells configuration has been selected as the most promising [5]. The architecture of the Hybrid Fuel Cells configuration in normal mode of operation is shown in Fig. 5.

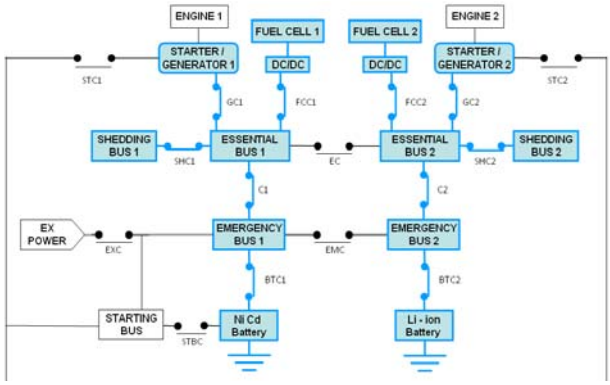


Fig. 5. Hybrid fuel cells configuration architecture – Normal mode of operation

In normal mode of operation both the fuel cells and the starter/generators supply all aircraft buses with the electric power. Since the fuel cells work in parallel with the starter/generators to provide the aircraft with the entire electric power, the electrical loads have been split into essential, shedding and emergency buses. Unlike the normal mode of operation, in emergency 1 mode of operation the fuel cells supply all users with the total amount of electric power. Batteries are used for engines start-up and to absorb power peaks.

2 Methodology

Detailed investigations to find the best way to supply the aircraft with secondary power, considering the various suppliers (starter/generators, batteries and fuel cells) in different modes of operations, have then been performed thanks to the Integrated Simulation Environment [7], where physical, functional and mission scenario simulations continuously exchange data and results. In particular the Integrated Simulation Environment integrates the SAvE Aircraft Model, the Thermal Model and the Mean Value Electrical Model [8] [9] and allows simulating a complete 30 hours mission of the aircraft, either with conventional configurations or with the Hybrid Fuel Cells configuration. The SAvE Aircraft Model simulates the dynamic behaviour of the whole aircraft and allows simulating a typical aircraft mission. The Thermal Model allows estimating dynamically the superficial temperature of the avionic equipment, which are installed on board the aircraft, and the mean temperature inside the aircraft, throughout the mission. Eventually the Mean Value Electrical Model allows estimating the electrical power consumption of the systems, which are installed on board the aircraft, and the mean power generation by means of generators and fuel cells. Fig. 6 illustrates the integration of the different models to constitute the Integrated Simulation Environment and shows all data exchanged between the models. The Integrated Simulation Environment allows simulating a complete mission of the SAvE aircraft, as well as the failure of its electrical sources during flight, thus making it possible to get detailed information on the duration and on the flight controllability of emergency modes of operation. The Integrated Simulation

Environment takes into account the variation with time of weights and electrical loads and it allows to investigate how these variations affect the aircraft performance and, particularly, to estimate the impact of the various Secondary Power System configurations in terms of environment (fuel consumption) and performance. Moreover, thanks to the Integrated Simulation Environment, it is possible to size more accurately the Secondary Power System components (electric machines, fuel cells, batteries, etc.) and the consumables (H_2, diesel), that have to be stored on board.

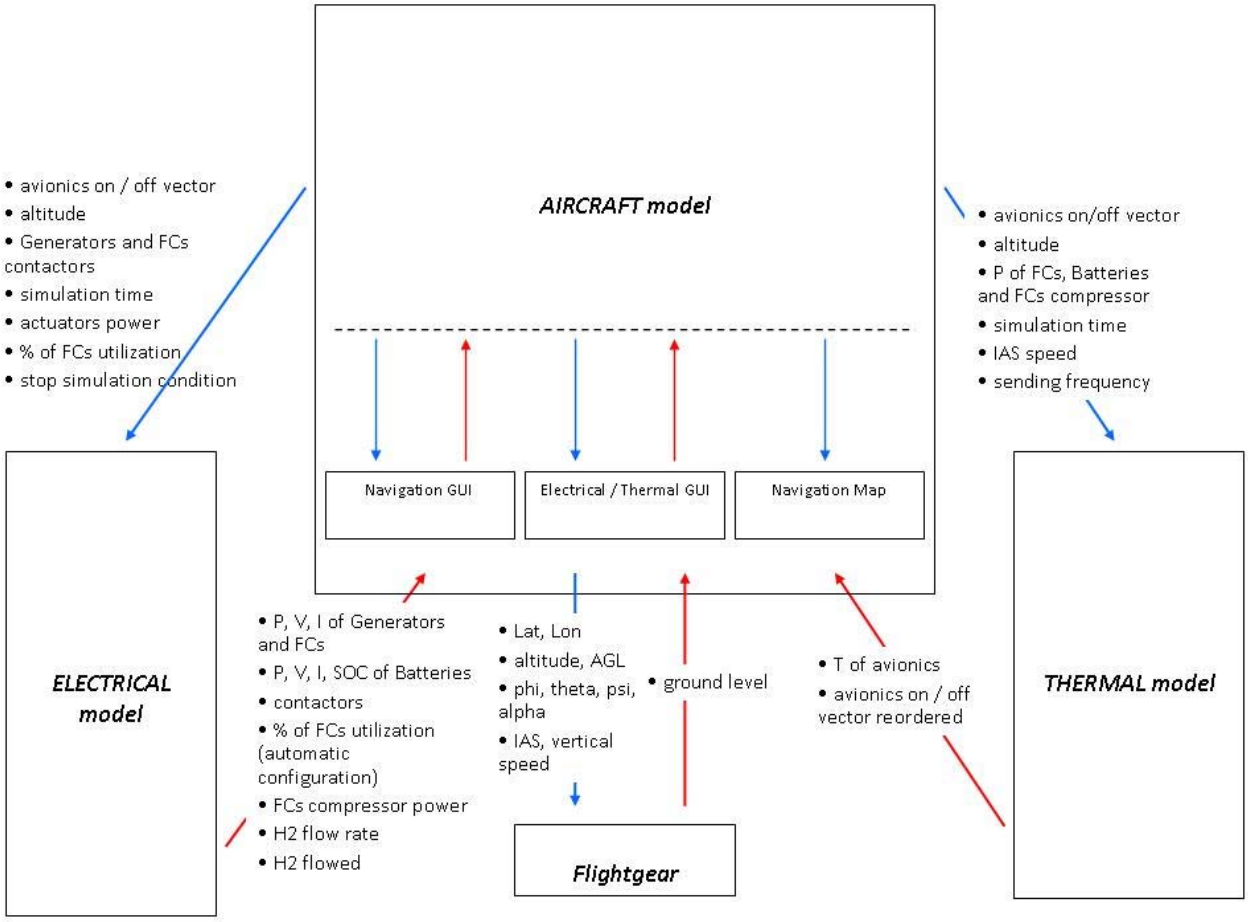


Fig. 6. Integration of the Electrical, Aircraft and Thermal model to constitute the Integrated Simulation Environment

The detailed investigations, to find the best way to supply the aircraft with the secondary power, have been carried out through various tests of SAvE aircraft in different working conditions. All tests have been performed through the Integrated Simulation Environment and have led to the refinement of the Hybrid Fuel Cells configuration. Main purpose of the tests has thus been the optimization of the Fuel Cells System parameters, in order to maximize the Fuel Cells System efficiency, $\eta_{SystemFC}$ (the ratio of the net energy supplied by the fuel cells, i.e. the energy supplied by the fuel cells, E_{FC}, minus the energy absorbed by the compressor, E_{COMPR}, and the calorific power due to H_2 consumption, P_{H2}, see Eq. 1), according to different flight phases constraints, mainly in terms of Fuel Cells System working conditions. The Fuel Cells efficiency, η_{FC} (i.e. the ratio of the energy supplied by the fuel cells, E_{FC}, and the calorific power due to the H_2 consumption, P_{H2}, see Eq. 2) has been also calculated. It is

worth remembering here that the compressor is used to give the fuel cells the correct air flow rate at high altitude where the air density decreases [10] [11].

$$\eta_{SystemFC} = \frac{E_{FC} - E_{COMPR}}{P_{H2}} \qquad (1)$$

$$\eta_{FC} = \frac{E_{FC}}{P_{H2}} \qquad (2)$$

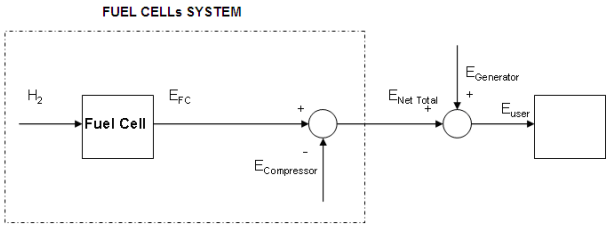


Fig. 7. Definition of the system energies

Fig. 7 sums up the relationships between the whole energy and the net energy supplied by the fuel cells (expressed, respectively as E_{FC} and $E_{Net\ Total}$ in Fig. 7), the energy absorbed by the compressor (expressed as $E_{Compressor}$ in Fig. 7), the energy supplied by the generators, $E_{Generator}$ and the energy absorbed by the users, E_{user}.

The maximization of the Fuel Cells System efficiency allows maximizing the energy that can be provided by a certain amount of hydrogen consumption. In order to maximize the Fuel Cells System efficiency, the key parameters have been the compressor and Fuel Cells working conditions, specifically (see Table 1):

- the Fuel Cell's inlet pressure, PinputFC.
- The ratio between inlet oxygen and consumed oxygen, λ. The value of this parameter ranges between an upper limit, due to NACA air intakes, high power consumption and low humidification problems, and a lower limit, due to chemical and physical reasons.
- The percentage of the aircraft electric power supplied by the Fuel Cells.

Purpose of the tests has been the choice of the best values of the key parameters (input parameters for the simulation, as reported in Table 1) for every mission phase performed by SAvE aircraft.

The methodology that has been adopted to perform the tests can be summed up by the following steps:

- all tests have been performed considering the Hybrid Fuel Cells configuration;
- the whole mission has been split into various phases;
- each mission phase has been performed in a pre-defined mode of operation;
- each mission phase has been simulated quite a few times with different sets of input parameters;
- the best set of input parameters, that guarantee the maximization of the Fuel Cells System efficiency, can then be chosen.

In particular the following mission phases performed in pre-defined modes of operations have been envisaged:

- climb from ground level to 3000 m in normal mode of operation;
- climb from 3000 m to 7000 m in normal mode of operation;
- climb from 7000 m to 14000 m in normal mode of operation;
- descent from 14000 m to 7000 m in emergency 1 mode of operation;
- descent from 7000 m to 3000 m in emergency 1 mode of operation;
- descent from 3000 m to ground level in emergency 1 mode of operation;
- descent in emergency 2 mode of operation;
- cruise in normal mode of operation.

The flight phases of climb (in normal mode of operation) and descent (in emergency 1 mode of operation) have been simulated first, in order to estimate the hydrogen and fuel consumed. Then, on the basis of the remaining amount of hydrogen, the refinement of the cruise flight phase has been accomplished, in order to obtain the maximum energy, which the Fuel Cells System would be able to supply.

TEST 2									
Input									
Parameters	**Value**								**Comments**
Test number:	**2.1**	2.2	2.3	2.4	2.5	2.6	2.7	2.8	
Mode of operation:	**Normal**	Normal	Normal	Normal	Normal	Normal	Normal	Normal	
PinputFC [atm]:	**0,70**	0,80	0,90	1,00	1,00	1,00	1,00	1,00	The compressor is switched on (compressor ratio varies in order to get PinputFC=0,7:1 atm). It has been chosen a lower limit of 0,7 atm to not differ too much from nominal working conditions.
Ratio of inlet oxygen with respect to oxygen spent λ:	**1,7**	1,7	1,7	1,7	2,5	3,0	3,5	4,5	It has been chosen a lower limit of 1,7 because this is the minimum λ working value specified in the FC reference document.
Altitude range [m]:	**3000-7000**	3000-7000	3000-7000	3000-7000	3000-7000	3000-7000	3000-7000	3000-7000	
IAS [km/h]:	**160**	160	160	160	160	160	160	160	
H2 stored at the beginning of the mission phase [kg]:	**9,14**	9,14	9,14	9,14	9,14	9,14	9,14	9,14	Output of Test 1.
Diesel stored at the beginning of the mission phase [kg]:	**1244,85**	1244,85	1244,85	1244,85	1244,85	1244,85	1244,85	1244,85	Output of Test 1.
Percentage of the aircraft electrical power supplied by the FC [%]:	**100**	100	100	100	100	100	100	100	In this phases there is a small generator 2 contribution.
Percentage of turbolence duration with respect to duration of mission phase [%]:	**40**	40	40	40	40	40	40	40	During turbolence there is a higher electrical consumption of actuators.It is a conservative value for typical climatic conditions at that altitude.
Probability that turbolence intensity exceed the typical one for that altitude:	**10^ (-4)**	10^(-4)	10^(-4)	10^(-4)	10^(-4)	10^(-4)	10^(-4)	10^(-4)	
Output									
Parameters	**Value**								**Comments**
Mission phase duration [min]:	**17,93**	17,93	17,93	17,93	17,93	17,93	17,93	17,93	
H2 consumption [kg]:	**0,155**	0,154	0,152	0,151	0,146	0,144	0,144	0,144	
H2 stored at the end of the mission phase [kg]:	**8,98**	8,98	8,98	8,99	8,99	8,99	8,99	8,99	
Diesel stored at the end of the mission phase [kg]:	**1225,79**	1225,79	1225,79	1225,79	1225,79	1225,79	1225,79	1225,79	
Energy supplied by the generators [kWh]:	**0,019**	0,026	0,031	0,034	0,033	0,042	0,068	0,124	It should be the lowest possible value.
Energy absorbed by the compressor [kWh]:	**0,077**	0,116	0,151	0,183	0,259	0,308	0,360	0,462	
Total energy supplied by the FC [kWh]:	**2,290**	2,324	2,357	2,385	2,464	2,504	2,530	2,578	
Net energy supplied by the FC [kWh]:	**2,213**	2,208	2,206	2,202	2,205	2,196	2,170	2,116	
FC system efficiency (ratio of the net energy supplied by FC and the calorific power developed by the H2 consumption) [%]:	**36,07**	36,19	36,43	36,66	38,13	38,36	37,86	36,94	Hydrogen calorific power is: 143000 kJ/kg

Table 1. Test for Normal mode climb 3000-7000 m

Table 1 reports the various sets of input values and the results obtained for the tests performed during climb between 3000 and 7000 m in normal mode of operation. As it can be noted from Table 1, where the column of the selected test has been highlighted in green, the first test has been chosen and its input parameters have been considered the best set for the climb phase between 3000 and 7000 m in normal mode of operation. It is worth noting that, even though in this case the Fuel Cells System efficiency is slightly less than in all other runs, the test has been selected as the best one because the higher value of the Fuel Cells System efficiency of all other runs is due to a higher contribution of energy supplied by the generator and absorbed by the compressor. In fact in the first test the energy supplied by the generators is the lowest. In particular Fig. 8 shows the variation with time of the Fuel Cells System (blue line) and the Fuel Cells (red line) efficiency during climb between 3000 and 7000 m in normal mode of operation. As it can be noted, the Fuel Cells efficiency remains higher

than the Fuel Cells System efficiency throughout the phase because of the energy absorbed by the compressor.

Other tests have also been performed, in order to estimate the main aircraft performance, like maximum endurance and range in case of failures of both engines at different altitudes, when emergency conditions can occur.

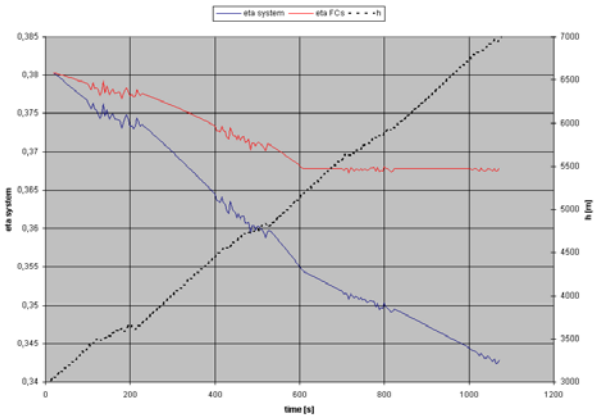


Fig. 8. Fuel Cells System Efficiency and Fuel Cells efficiency (medium Fuel Cells Efficiency)

NORMAL MODE OF OPERATION										
Phase	time [min]	IAS [km/h]	p_min_FC [bar]	lambda	% of total load given by FCs	H2 used [kg]	Diesel used [kg]	eta_syst FC [%]	n° test	Test chosen
0 - 3000 m	15,40	160	Penvironment	4,5 at s.l. and linearly to 3 at 3000m	100	0,064	15,15	46,93	5	TEST 1.3
3000 - 7000 m	17,93	160	0,70	1,7	100	0,155	19,06	36,07	8	TEST 2.1
7000 - 14000 m	53,47	160	0,70	1,7	100	0,328	49,25	34,83	5	TEST 3.1
cruise 14000 m	1927,20	155	0,70	1,7	66	7,520	1093,54	36,34	3	TEST 8.3
14000 - 7000 m	41,12	160	0,70	1,7	100	0,241	5,974	36,35	-	-
7000 - 3000 m	32,35	160	0,70	1,7	100	0,281	4,636	35,95	-	-
3000 - 0 m	29,49	160	Penvironment	4,5 at s.l. and linearly to 3 at 3000m	100	0,116	4,266	46,79	-	-
Total Mission	**2116,95**	**variable**	**variable**	**variable**	**variable**	**8,702**	**1192**	**36,52**		
EMERGENCY 1 DESCENT										
Phase	time [min]	IAS [km/h]	p_min_FC [bar]	lambda	% of total load given by FCs	H2 used [kg]	Diesel used [kg]	eta_syst FC [%]	n° test	Test chosen
14000 - 7000 m	47,67	130	0,70	1,7	100	0,134	0,000	41,75	2	TEST 4.2
7000 - 3000 m	37,00	130	0,70	1,7	100	0,107	0,000	45,21	2	TEST 5.2
3000 - 0 m	33,27	130	Penvironment	4,5 at s.l. and linearly to 3 at 3000m	100	0,041	0,000	54,90	2	TEST 6.2
Total EM 1 descent	**117,93**	**130**	**variable**	**variable**	**100**	**0,283**	**0,000**	**46,55**		
EMERGENCY 2 DESCENT										
Phase	time [min]	IAS [km/h]	p_min_FC [bar]	lambda	% of total load given by FCs	H2 used [kg]	Diesel used [kg]	eta_syst FC [%]	n° test	Test chosen
EM 2 descent	99,40	130	n.a.	n.a.	n.a.	0,000	0,012	n. a.	1	TEST 7.1

Table 2. Tests results

Eventually Table 2 highlights the best set of values of key parameters per each mission phase. It is worth underlying that the fuel cells have been modeled to work in low pressure environmental conditions (0,7 bar). The results obtained are coherent with the state-of-the art literature; however experimental tests would be quite useful to validate the analytical models.

3 Results

After selecting the most appropriate set of key parameters to maximize the Fuel Cells System efficiency, a complete standard SAvE mission (see Fig. 9), which aims at surveying and monitoring the territory of Piemonte, have

been simulated in the Integrated Simulation Environment.

Initial conditions are expressed in terms of:

- IAS, Indicated Air Speed;
- aircraft zero fuel weight, fuel weight and H_2 (maximum 9,2 kg) weight;
- position of flaps and of landing gear;
- aircraft position: altitude, latitude, longitude and heading.

Moreover waypoints have also to be initially set for automatic navigation and the fuel cells working conditions have to be defined per each mission phase according to the results shown in Table 2.

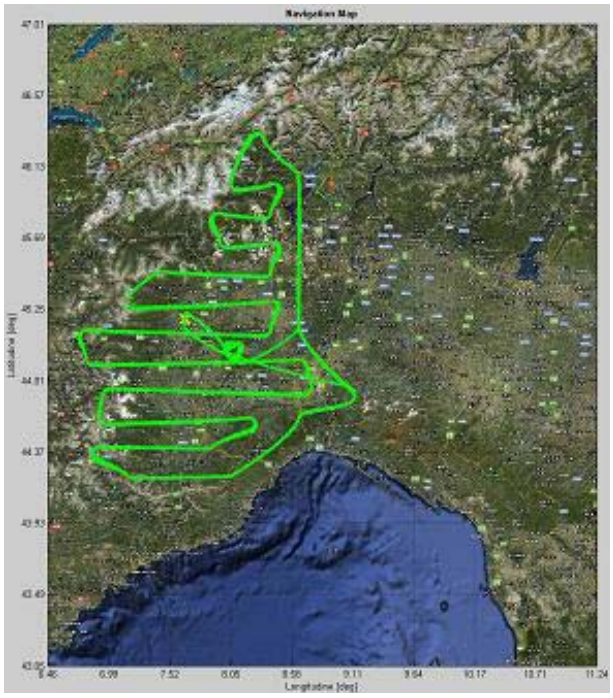


Fig. 9. Example of standard SAvE mission

Fig. 11 shows a typical screen shot of the mission simulation in normal mode of operation, while Fig. 10 and Fig. 12 illustrate examples of results obtained.

Fig. 10 shows the electrical power generation throughout the mission: in particular the red line depicts the total electrical power supplied by all power sources on-board the aircraft, the light blue line depicts the electrical power supplied by the fuel cells, the dark blues line depicts the electrical power supplied by the generators and the green line depicts the electrical power supplied by the batteries.

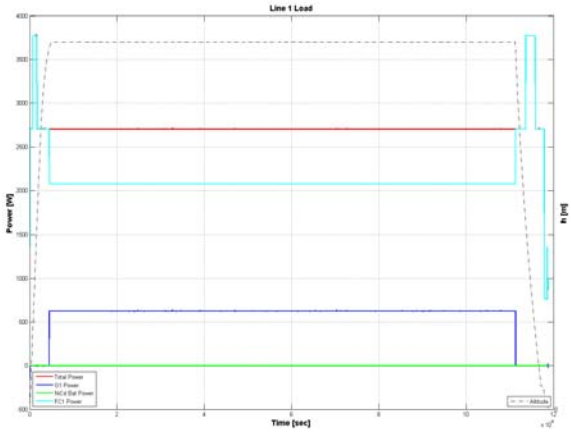


Fig. 10. Line 1: Electrical power generation

It is important to notice that during climb and descent, the whole electrical load is supplied by the fuel cells (see also Table 2), while during cruise, the electrical load is shared between generators (about 30% of the whole load) and fuel cells (about 70% of the whole load).

Fig. 12 shows the avionic elements thermal profile throughout the mission. Attention has to be paid to the phases of climb and descent, when the external air parameters change rapidly. Apart from WBDL_SAT_HVPS (Wide Band Data Link_Satellite), all avionic equipment reach a temperature lower than 80 °C (the maximum allowable temperature for a correct working compliant with components specifications): for this component, a loop heat pipe toward an heat sink may be used.

Eventually Fig. 13 illustrates the flight path of SAvE aircraft in emergency 1 mode of operation. It is worth noting that this is a predefined flight path to let the aircraft approach an emergency landing. Fig. 14 illustrates the electrical power generation as a function of time for the line 2 of the aircraft. The following considerations can be made:

- the emergency 1 mode of operation starts at the beginning of the simulation at about 14000 m of altitude;
- the two shedding buses are cut off;
- the line 1 does not supply the system;
- the line 2 fuel cell supply the whole system;
- the batteries do not contribute to the supply of electrical power.

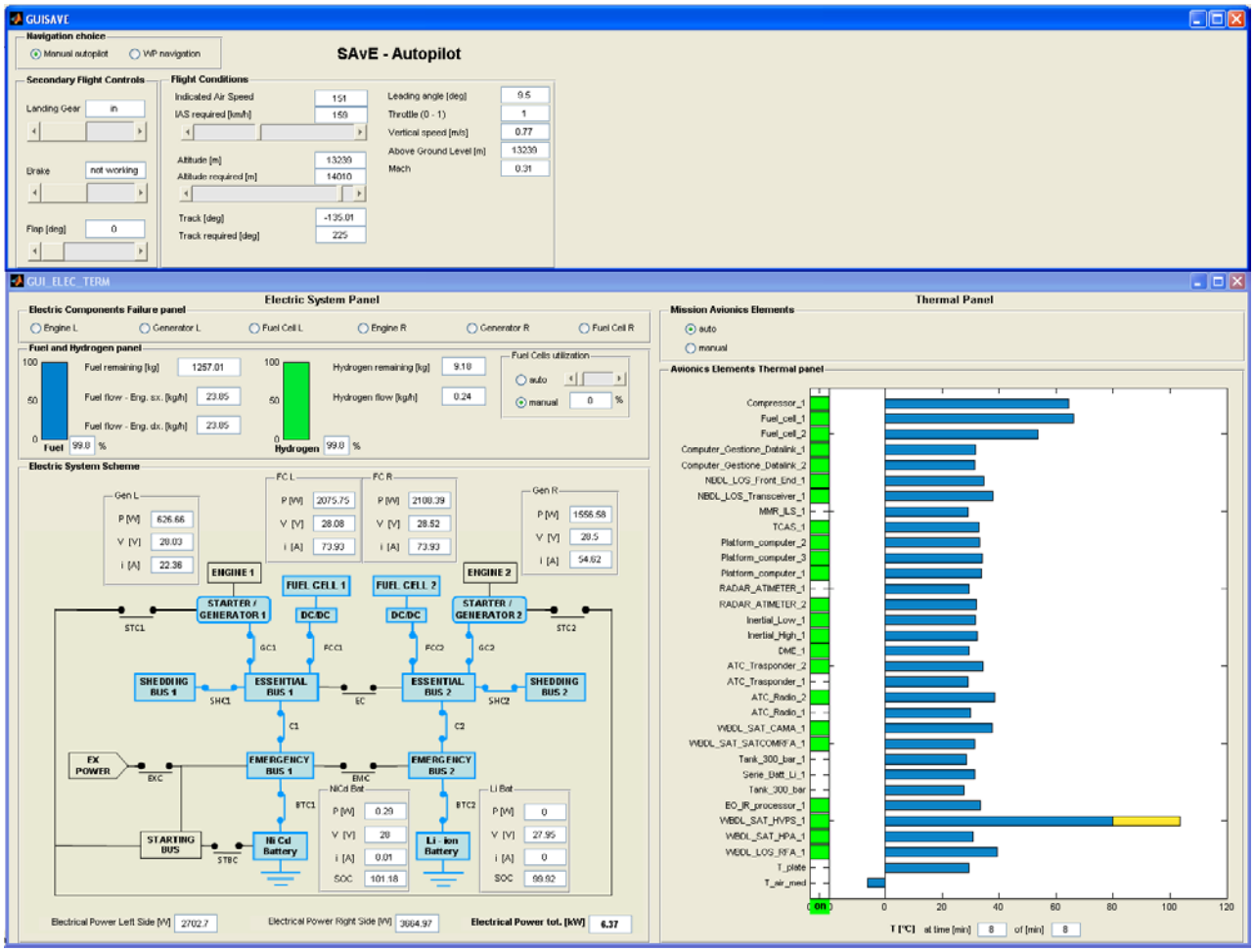


Fig. 11. Screen shot of the mission simulation in normal mode of operation

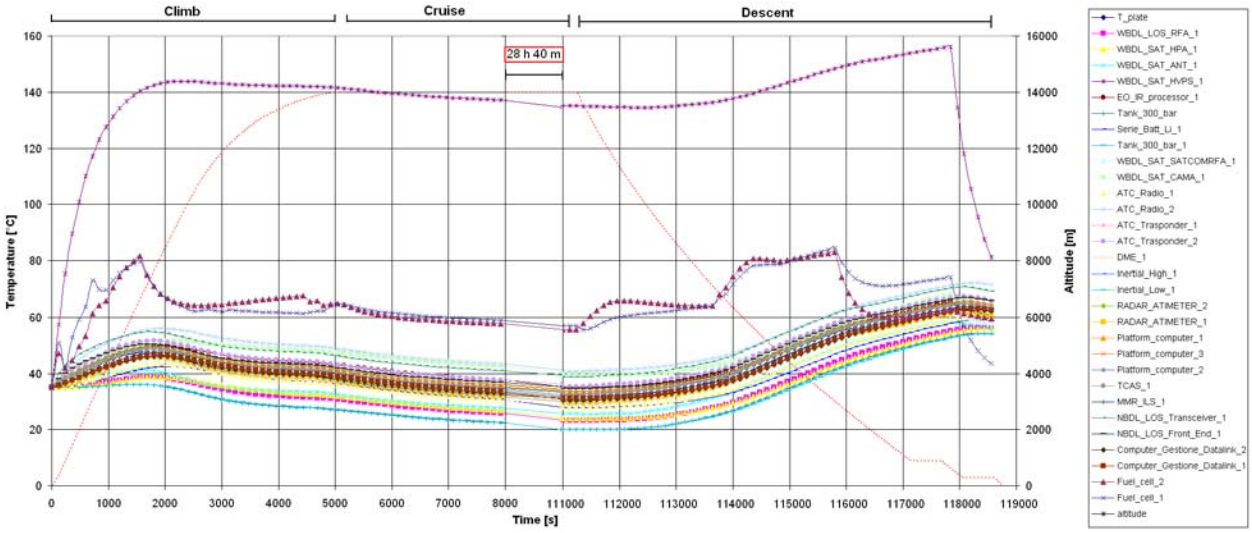


Fig. 12. Avionic elements temperature during whole mission (without WBDL_SAT_HVPS heat dissipation)

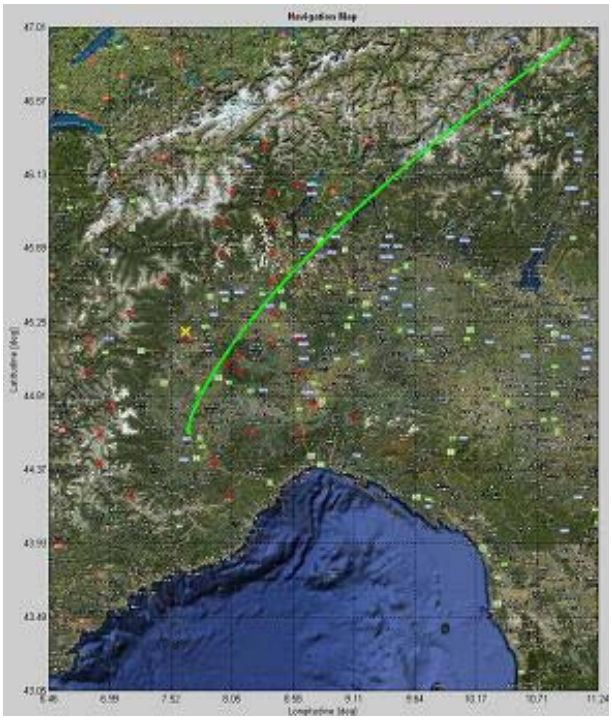


Fig. 13. Screen shot of the flight path in emergency 1 mode of operation

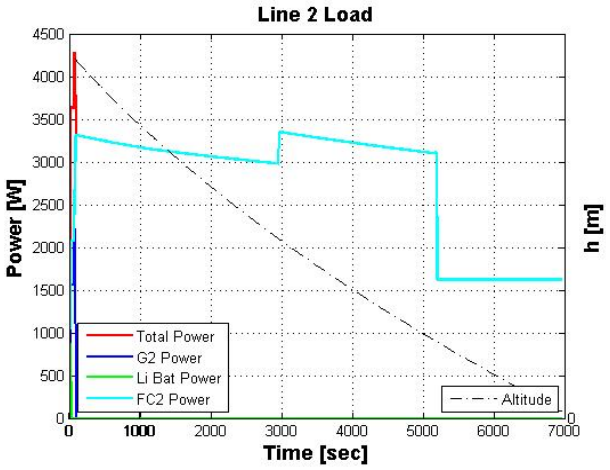


Fig. 14. Electrical power generation vs time for line 2 in emergency 1 mode of operation

4 Conclusion

The paper presents the refined Hybrid Fuel Cells configuration of the all electric Secondary Power System of a future advanced MALE UAV, named SAvE aircraft. The results obtained from the virtual tests within the Integrated Simulation Environment show that the fuel cells, with the appropriate set of key parameters are able to supply about 70% of the entire electric load during cruise and 100% of the electric load during climb and descend in normal mode of operation. Even though these results are encouraging, specific experimental tests of fuel cells working at high altitude are recommended to validate the analytical and numerical models.

References

[1] Borgese A., Faggella D., "CAD model of the reference SAvE aircraft" Alenia Aeronautica/Politecnico di Torino, 2009.

[2] Chiesa S., *Impianti di Bordo per Aeromobili: Impianto Elettrico*, 1989, Turin C.L.U.T. publisher.

[3] Chiesa S., Farfaglia S., Viola N., "Design of all electric secondary power system for future advanced MALE UAV", *International Congress of the Aeronautical Sciences*, Proceedings of the 27th International Congress of the Aeronautical Sciences, Nice, France, 2010.

[4] Chiesa S., *Impianti di Bordo per Aeromobili: Impianti Pneumatico, Condizionamento, Anti-Ghiaccio e APU*, 1995, Turin C.L.U.T. publisher.

[5] De Luca D., Faggella D., Greco F., "Solution Trade Off" Alenia Aeronautica/Politecnico di Torino, 2009.

[6] De Luca D., Faggella D., "Identification Fuel Cells System Design" Alenia Aeronautica/Politecnico di Torino, 2009.

[7] De Luca D., Faggella D., "Preliminary virtual integration" Alenia Aeronautica/Politecnico di Torino, 2010.

[8] De Luca D., Faggella D., Greco F., Ronchetto D., "Simulation Models" Alenia Aeronautica/Politecnico di Torino, 2010.

[9] De Luca D., Faggella D., Greco F., Mannini A., "Model arrangement of the specific application" Alenia Aeronautica/Politecnico di Torino, 2011.

[10] Guynn M.D., Jurns J.M., Millis M.G., Tomsik T.M., Tornabene R.T., Van Overbeke T.J., "Hydrogen Fuel System Design Trades for High-Altitude Long-Endurance Remotely-Operated Aircraft", NASA Technical Reports, Glenn Research Center Cleveland, Ohio, 2009.

[11] Mak A., Meier J. "Fuel Cell Auxiliary Power Study", NASA Technical Report, Volume 1: RASER Task Order 5, Glenn Research Center Cleveland, Ohio, 2007.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Flight tests of solar powered light UAS for very long endurance autonomous flight of solar and fuel cells powered UAV

G. Romeo, M. Pacino and F. Borello

Politecnico di Torino, Dept. of Aerospace Engineering, C. Duca degli Abruzzi 24,Torino, Italy,
giulio.romeo@polito.it - http://www.polito.it/grupporomeo - www.enfica-fc.polito.it

Keywords: *Unmanned Aircraft, solar and fuel cell power, flight tests*

Abstract

*A research is being carried out aiming at the design of **Very-Long Endurance Solar Powered Autonomous Stratospheric UAV (VESPAS-UAV)** and manufacturing of solar powered prototype. The flying model **SESA (Small Electric Solar Unmanned Airplane)** was manufactured to carry out several experimental flight test with a small UAV and to demonstrate some critical technologies and applications (high efficiency solar cells, electric brushless motor, controllers, etc). SESA, powered by solar energy, made its first flight representing the first Solar powered light UAV (weight 35kg, Payload 5-6 kg) flying in Europe. Several flight tests were successfully carried out by POLITO on the two-seat aircraft RAPID200-FC, first all-electric aeroplane worldwide powered by fuel cells. The power (max 40 kW) was supplied either by the Fuel cells (20 kW) and by the Main Battery System (only for climbing phase) installed onboard. Six test flights were successfully carried out showing the positive handling qualities and satisfactory engine performances of this airplane.*

1 Introduction

The total UAV market is growing at a rapid pace and it is imperative that the European community should make a serious effort to attain a significant segment of this market. The wide range of applications for civilian UAVs, will open up a variety of markets for potential sales and economic growth. The Scientific Community could benefit in many ways from employing UAVs in the civilian sphere. The utilisation of UAVs for **border and costal patrol, homeland security, maritime surveillance**, **forest fires mapping, real-time monitoring of seismic-risk areas, volcanoes eruption, etc,** "Eye-in-the-sky" surveillance, will allow better law-enforcement in the protection of citizens and integrity of the borders and will assure the public is aware of imminent disasters and can prepare for their occurrence.

In Sept. 2010, the Boeing Company signed an agreement with the U.S. DARPA to develop and fly the SolarEagle unmanned aircraft for the Vulture II demonstration program. Under the terms of the **$89 million contract**, SolarEagle will make its first demonstration flight in 2014. "SolarEagle (**Fig. 1**) is a uniquely configured, large unmanned aircraft designed to eventually remain on station at stratospheric altitudes for at least five years. "That's a daunting task, but Boeing has a highly reliable solar electric design



Fig. 1 – SolarEagle, Boeing

that will meet the challenge in order to perform persistent communications, intelligence, surveillance and reconnaissance missions from altitudes above 60,000 feet." (Pat O'Neil, Boeing Phantom Works).

Under the coordination of the first author, a research is being carried out since 1996 aiming at the design of **Very-Long Endurance Solar Powered Autonomous Stratospheric UAV (VESPAS-UAV)** and manufacturing of solar powered prototype. This could play the role of pseudo-satellite, with the advantage of allowing more detailed land vision, due to the relative closeness to the land, with continuous earth observation and at a much lower cost than real satellites (**Fig. 2**). Several satellite systems used for earth observation are useless for continuous real-time border surveillance because of their limited spatial resolution.



Fig. 2 – VESPAS-UAV Maritime surveillance

All the Mediterranean Sea border, from Turkey to Spain could be electronically controlled by 8-9 platforms from a high altitude (17-20 km), with very-long endurance (4-6 months) stratospheric UAV (payload up to 150-200 kg, power available for payload up to 1.5-2kW). It is essential to control who and what enters European countries in order to prevent the admission of terrorists and instruments of terror across borders, along coastlines and harbours. Continuous (24h by 24h) border monitoring would be guaranteed and service costs and tedious work would be drastically reduced. A 300km diameter area would be monitored by each of these platforms. UAVs are less expensive (800-1000 €/flight hour) than other manned aircraft used for the borders surveillance (8.000 €/flight hour) drastically reducing service cost and tedious work. NO ONE real very-long endurance stratospheric platform is actually available in Europe.

The main advantage of **VESPAS-UAV** is that this system has less climb and descend events, important when considering interference with the aviation traffic. Any other high-altitude UAV configuration for border surveillance has a very limited endurance (24-48 hours) that would drastically increase any potential collision risk with civil aviation traffic. A double number of UAVs would be requested to guarantee the surveillance service, highly increasing the System Total Life Cycle Cost. Any other MALE UAV has the disadvantage that a much higher number of UAVs are requested to continuously cover all the Mediterranean Sea, since the covered area is decreasing with the square value of the flying altitude increasing tremendously the cost.

POLITO (Co-ordinator Prof. G. Romeo) is carrying on, since several years, one of the few existing world projects on solar powered aerodynamic stratospheric platforms. After a preliminary funding by the **Italian Space Agency**, a very great push to the project has indeed been obtained by the financial support received by the **European Commission** in the field of stratospheric platform (**HeliNet, Capecon, Enfica-FC, Tango**) [1-4].

The possibility of medium-long endurance (4-6 months) for a stratospheric platform can be realised with the application of an integrated Hydrogen-based energy system. It is a closed-loop system: during daytime, the power generated by thin high efficiency solar cells that cover the aircraft's wing and horizontal tail supply power to electric motors for flying and to an electrolyser which splits water into its two components, hydrogen and oxygen. The gases are stored into pressurized tanks and then, during night-time, used as inlet gases for fuel cells stack in order to produce electric DC power and water to be supplied to the electrolyser. Since fuel cells represent the promise of clean and efficient power generation, they are a suitable alternative to conventional energy sources.

Within the EC funded project **TANGO** [4] (Polito Scient. Resp.: Prof. **G. Romeo**) the Heliplat UAV was analysed in cooperation with several satellite systems for few civil applications of the GMES (Global Monitoring Environmental and Security) action. Demonstrations

integrated satellite telecommunication solutions with on-going GMES developments in the framework of **fisheries management**. Inclusion of **UAVs** in the global relay infrastructure enables quasi real time and continuous access to dedicated zones for monitoring or surveillance. A flight test with a scaled solar powered UAV was prepared for final integration with on-going GMES developments in the framework of fisheries management.

2 HELIPLAT® VESPAS

The full scale **HeliPlat®** UAV **(Fig. 3)** was designed by using the most advanced tools for obtaining an endurance of several months (4-6) and being operable in almost all typical environment conditions (wind jet up to 140km/h) at stratospheric altitude (17-20 km). [5-8]. A computer program was developed for designing the platform capable of remaining aloft for very long period of time. The project of the platform was completed up to a quasi-final detail design. A numerical aerodynamic analysis was performed to obtain the highest efficiencies of the whole wing and airplane. Several experimental tests were carried out on Low-Speed-Low-Turbulence Wind-tunnel, obtaining a very good correlation between analytical and experimental results. A first configuration of Heliplat® was worked out, as a result of the preliminary design study. The platform is a monoplane with 8 brushless motors, twin-boom tail type, horizontal stabilizer and two rudders. The design procedure followed in the analysis was based on the energy balance equilibrium between the available solar power and the required power for flying; the endurance parameter has in particular to be fulfilled to minimise the power required for a horizontal flight. Main characteristics for a flight at 38°N latitude and design altitude of 17 km are: Total weight: 8500N; Wing Area: 176m^2; Span:73m; Required Power: 7500W; Aspect ratio=33; Cruise Speed = 71 km/h.

A payload up to 100kg, with available power up to 1500W, could be installed on board for several global monitoring of environmental and security applications (GMES). A numerical aerodynamic analysis was performed to obtain the highest efficiencies of the whole wing and airplane including propellers by using VSAERO software, at the flight Reynolds numbers [6,7].

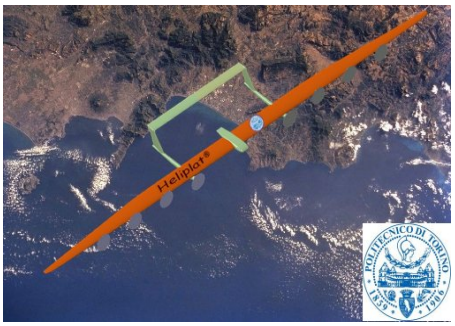


Fig. 3 - HeliPlat ® Configuration

A 1:3 scaled size prototype (wing span 24m, length 7 m) was built in advanced composite material (high modulus CFRP) in order to show the technological feasibility **(Fig. 4).** Several shear/bending/torsion static tests on the whole manufactured scaled-size prototype were performed in our laboratory up to failure load (7.5g) finding a very good correlation with the in-house developed numerical analysis and FEM analysis **(Fig. 5).**



Fig. 4 – 1:3 Scaled-size HeliPlat® UAV



Fig. 5 - 1:3 Scaled-size HeliPlat® UAV – Shear/Bending and Torsion Tests

The results obtained in the CAPECON project [2], confirmed the feasibility of a solar powered stratospheric UAV (SHAMPO) satisfying the requirements of a long endurance stationary flight. A detailed aerodynamic and structural design, the flight mechanics and the

electric systems was completed by the Politecnico di Torino up to a quasi-final design **(Fig. 6)** [7]. A greater aerodynamic and structural efficiency was obtained allowing an higher payload mass (150 kg) and power (1.8 kW).

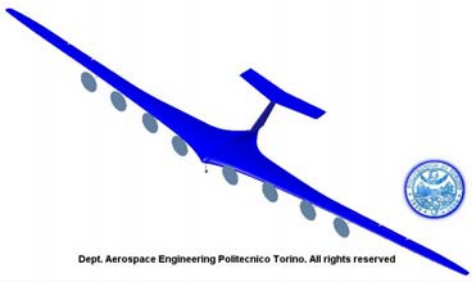


Fig. 6 – CAPECON SHAMPO Configuration

3 Small Electric Solar Unmanned Airplane

The flying model **SESA (**Small Electric Solar Unmanned Airplane) **(Fig. 7)** was manufactured within the EC funded projects CAPECON and TANGO, to carry out several experimental flight test with a small UAV and to demonstrate some critical technologies and applications [9]. Electric brushless motor and LiPo batteries (for climbing phase missions), are used as propulsion system. The level flight power is being achieved from 2 m^2 thin high efficiency (21%) mono-crystalline silicon arrays bonded over 7m wing span**;** a far higher **endurance up to 6-7 hours during June and July** is obtained.



Figure 7 – SESA flight

The wing structure was realized using glass-fibre reinforced plastic for skin and carbon-fibre for main spar **(Fig. 8).** The fuselage structure was designed for a better positioning of the masses inside the fuselage and for a better centre of gravity excursion. The structure was designed to withstand the limit load of 3.8. The structure was built by the "Archemide Advanced Composite" company using fibreglass for the fuse skin and carbon-fibre for a few frames, especially in the connection with the wing spar. A mould was realized for the proper lay-out of the fuse. The boom that connect the fuse and tails has been manufactured in carbon-fibre. The new landing gear was also manufactured by CFRP.

SESA main characteristics are the following: Wing span:7m; Wing area:2 m^2; Total gross weight:35kg; Payload: 5-6kg; Max Solar Power: 370-395W (45°-36°N, June); Max power brushless Motor: 2kW; Level Flight Power: 250 W; Minimum Speed: 36 km/h.



Fig. 8 – SESA manufacture phases.

SESA, powered by solar energy, made its first flight, October 2007, near Torino (45°N latitude) at an altitude less that 500m; see movie on [10]. The plane represents the first European Solar powered light UAV flying in Europe. In the '90, DLR manufactured and flew the UAV scaled solar model "Solitair", but it is no more active. In September 2007, the UAV solar model "Zephir", by Qinetiq, flew in New Mexico. A very great difference which can be seen with respect to other solar powered UAVs is that in SESA all the structures were designed in order to withstand the EASA-VLA regulation; i.e. a maximum limit load of 3.8 has to be withstand by the UAV. The experimental tests carried up to now validated few critical technologies for high altitude very long endurance flight: high efficiency solar cells, electric brushless motor, controllers, video and thermo camera images transmission, telemetry, satellite communication, etc. These results shall be very useful to our group for the future projects.

3.1 Electronic Power system

The flight management electronic power system is reported in **Figure 9**. The power produced by the solar cells is directly supplied to the brushless electric motor for level flight. During take-off and climbing, or for some particular manoeuvres requiring more power, the power is also supplied by the LiPo batteries. The solar panel is composed of a parallel circuit series of 130 solar cells each with efficiency of 21.5% (@1000 W/m^2 and 25°C). The maximum solar panel voltage is 43.5V.

The development of the MPPT electronic device (Maximum Point Power Tracking) was of particular importance for the success of the flight mission in order to optimize the maximum power that could be obtained by the solar cells and to improve endurance **(Fig. 9).**

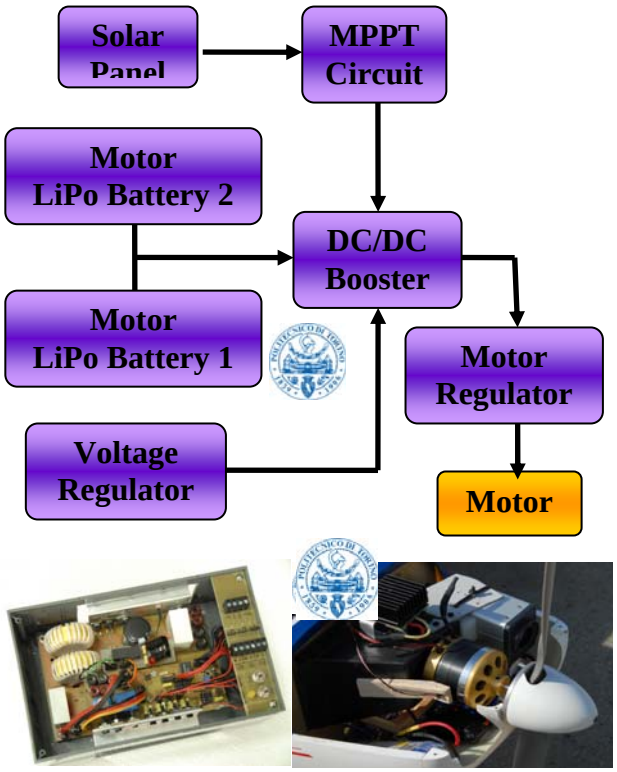


Fig. 9. SESA Electronic Power system and MPPT

The inverter will supply power to the brushless electric motor. The cooling of the inverter, to avoid shut-off of the system as a maximum allowable temperature is reached, is also an important feature. Any solar cell power exceeding the required level flight power will be used to recharge the battery. In order to show the opportunism of introducing such platforms in surveillance or monitoring systems, the model was equipped with a wireless colours camera (40x, resolution 720x576 pixel) and thermo-camera (160x120 pixel); such cameras can transmit (at 1.2GHz), in a range of about 1 – 1.5 km in open air and, through an appropriate capturing peripheral device, directly to a PC**;** the PC could be used to analyze images in real-time, for example to automatically spot forest fires.

3.2 Remote control and telemetry system

Up to 12 servo-actuators could be remote controlled by a radio modem for the UAV flight. The elements controlled were: rudder and tail gear, 2 elevators, 2 ailerons, motor rpm. A telemetry system **(Fig. 10)** was also installed on board to transmit in real time all the most critical data for the safety of the flight to the ground control station. The following data are recorded and wireless sent to the GCS to continuously have the real aircraft flight conditions:

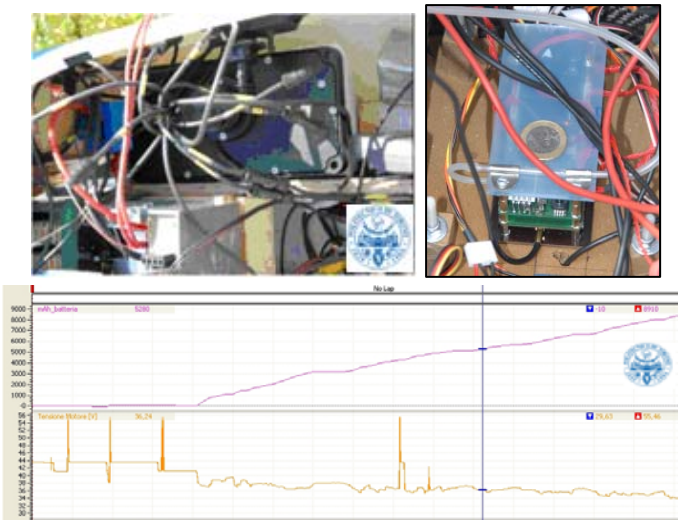


Fig. 10 - Wireless Telemetry system, Data Logger, Autopilot & Data Display (mAh and Motor Voltage).

a) True Air Speed; b) Voltage and Current of the brushless motor; c) Voltage, Current and consumption of the main motor battery; d) Temperature Service of the motor and inverter.

A data acquisition system (Enclosed Dash Logger-EDL2) measured and recorded the most important air vehicle parameters. The data were sent via a radio transmitter in the UAV, to a radio receiver on the ground and, by an RS232, to the GCS. The data were elaborated and displayed to allow the engineers to monitor the

main data while the air vehicle was still in flight. This increased the possibility of detecting early signs of problems, to warn the pilot on the ground and it also helped to prepare for any set up adjustments necessary during the flight.

3.3 Autopilot system and mission architecture

An autopilot system was installed on board for an **autonomously flight up to 40km of distance** by a highly integrated data acquisition, processing and control system which includes all necessary components for aircraft control. Satellite based Communication System was installed onboard for the **Iridium©** satellite communication network aiming to a BLOS flight.
The Autopilot system (built by Mavionics GmbH,D) **(Fig. 10)** consists of three main parts: 1) The TrIMU Sensor Block contains a complete 3-axis Inertia Measurement Unit (IMU) and two pressure sensors for barometric altitude and airspeed determination. It generates up to 12 independent servo control signals. 2) The Navigation Core hosts a sophisticated navigation filter for GPS/IMU data fusion which enables precise and long-term stable determination of the position, velocity and the Eulerian angles and to obtain a reliable attitude determination. 3) a Satellite Navigation Receiver. 16 channel GPS receiver with high sensitivity and integrated ceramic patch antenna. The connection between the Core and Satellite Navigation Receiver is only through power and digital lines thus significantly reducing interference.
The on-board autopilot system communicates with Ground Control (GPS time, position, Eulerian angles, flight speed, etc.) via a dedicated, direct bidirectional data link, using a radio modem which operates in the European 868 MHz band at a rate of 4 Hz. The small size of the demo scaled UAV allowed to install onboard just LEO systems Satellite based Communication System for Iridium© or similar network.
A direct radio connection was used during the take-off & landing and an Iridium satellite-based system for longer flights off-shore. The signal from the 2.4GHz on-board R/C receiver is connected to the autopilot and a special switching mechanism inside the autopilot allow the autopilot to be overridden by the remote control at any time, as long as the remote control transmitter is within range of the aircraft. The range of this system will be 1 to 2 km, which is suitable for the visibility range of the manual pilot. Any possible malfunction of the autopilot or telemetry system is overcome and flight safety is increased to great extent. The same telemetry transceiver (radio modem) system is needed for the ground segment. A directional antenna can be used thus greatly increasing range and reliability of the data connection. The GCS is basically composed of a PC and ground control software which is the user interface to the SESA system for mission planning and runtime control. All the mission planning work is done intuitionally on an underlying map. This allows a very flexible and safe flight path design. A check with respect to the aircraft performance is done automatically to ensure a realistic and safe flight of the UAV.
Several demo flight have been carried out in our airfield near Torino. As can be seen in **Figure 11**, the solar powered airplane is flying autonomously continuous, while the pilot is just controlling that the flight is going according to the planned flight path without any flight control; just in case of emergency the pilot will override the autopilot flight and will take the flight control. The flight path planned for the test flight is plotted in **Figure 12**. The flight path followed by SESA during the autonomous flight is also plotted.

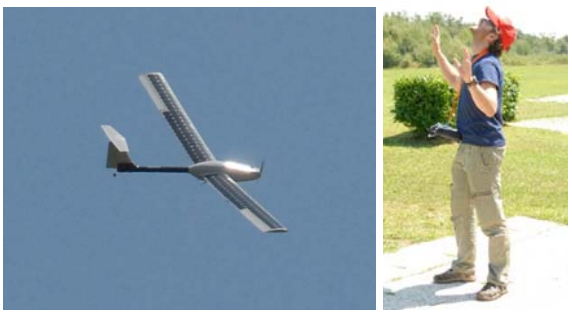


Fig. 11 – Autonomous Flight of SESA UAV

4 UAV Flight Demo for Fishery

In July 2009 a SESA flight demo has been performed to show the system effectiveness. **Fig 13** shows the mission architecture. The mission had the following characteristics:

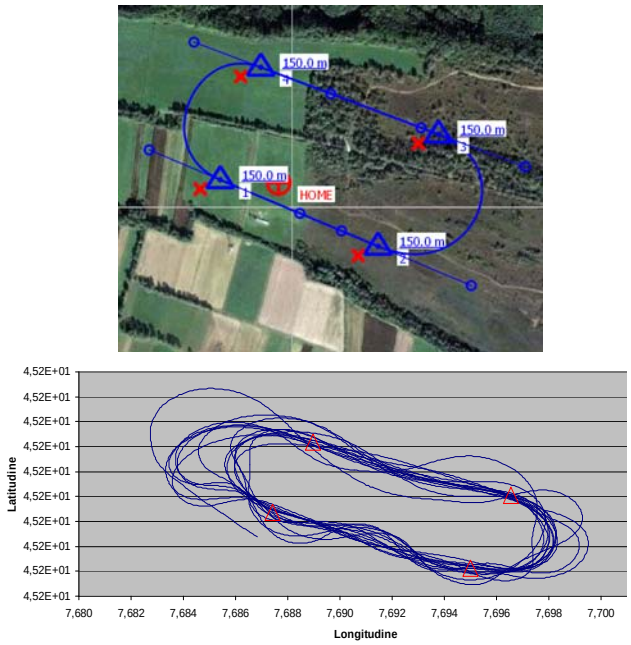


Fig. 12 – Planned Flight path of SESA UAV & Recorded path

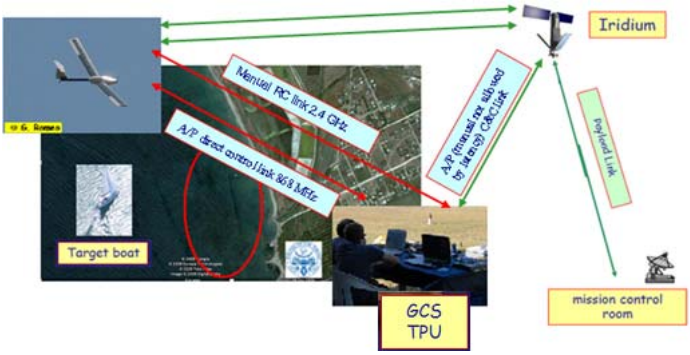


Fig. 13 – Flight area and Mission Architecture.

1. UAV manually controlled for take off.
2. Once in the air, it was switched to auto pilot mode until it reached its destination.
3. The ship positions were sent to the GCS which programmed the UAV to reach the ship coordinates
4. Flying altitude between 150-300 m high.
5. The flight was entirely above sea water.
6. Images of forward and downward cameras transmitted continuously to the GCS.
7. Target coordinates for the auto pilot updated during the flight as the ship moved.
8. When UAV arrives at the ship location, it took some high resolution pictures.
9. Once mission was accomplished, UAV reprogrammed to fly back to the take-off site.
10. When plane reached remote control and was within the line of sight of the operator, the UAV switched back to manual mode to safely land. At a UAV distance of less than 5km from the GCS, a double redundancy of transmission is expected for safety reasons.
11. Total duration of the flight was around 3h. Two hours to take off, to get into position and to take picture and one hour to get back and land.

Talks with the Italian CAA authorities concerning flying of UAV showed some safety issue to take into account. Since SESA weight is 35kg and since it shall be used for research and scientific purposes, an EASA certificate is not necessary. The following items were indeed pursed:
a) See & Avoid system on board; b) Flight over non populated area (although Maximum kinetic energy less than 95 KJ); c) Direct visual manual control necessary to obtain a safe flight and to avoid double-triple redundancy. SESA shall be followed during the demo flight by a boat with the pilot on board for any UAV emergency remote control. d) No catastrophic failure condition shall result from the failure of a single component; the allowable Quantitative Probabilities (per 1 flight hour) is less than 10^{-6}.

4.1 Flight Tests for Fishery Demo

Flight tests were carried out from S. Giorgio airfield (5km south Lido di Tarquinia). The airplane was equipped with: a) video camera looking forward for see & avoidance purpose but also for looking boats and ships. Zoom control is also possible by operator. Images are continuously wireless sent to the GCS and can be view by the operator. b) photo/video camera looking downward; video images are continuously wireless sent to the GCS and can be view by operator. Every 5 or 10 or 15 seconds, the camera take a picture and they are stored on board. Video is continuously wireless sent to GCS. The quality of the picture is good.
Fig. 14 shows some of the pictures taken from downward camera for identifying the boats.
SESA was used, in cooperation with Iridium satellite system, as a reconnaissance aircraft of the sea patrol sending real time images and videos to the GCS.



Fig. 14 - Looking Downward Camera Photos & zoom

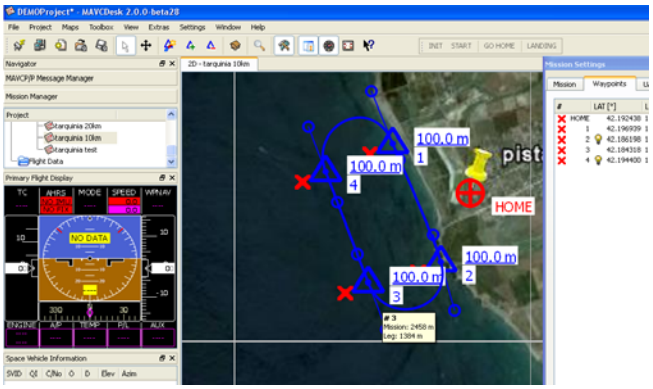


Fig. 15 – Planned Flight path of SESA UAV

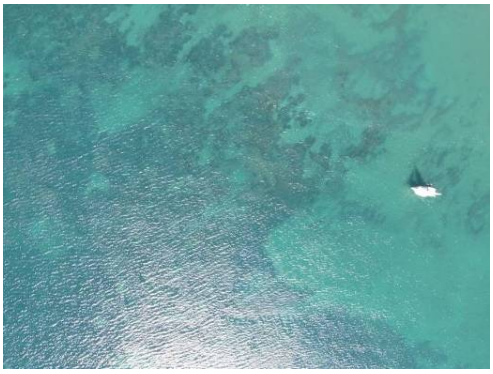


Fig. 16 – Photo of boat from downward camera

The GCS is composed by two laptops; one is controlling the flight path planned in the autopilot system but also shows the movie taken by the cameras looking downward; the second shows the movie taken by the cameras two looking forward. The flight path planned for the test flight is plotted in **Fig. 15**. One of the many pictures taken from the onboard looking downward camera is reported in **Fig. 16.** A maximum altitude of about 225m was reached during flight. The sailboat detected by the camera was a 43ft long boat. The flight path maintained by the airplane during the demo for continuous following the boat is plotted in **Fig. 17**; also reported are the GPS coordinate received by Iridium including the altitude too.

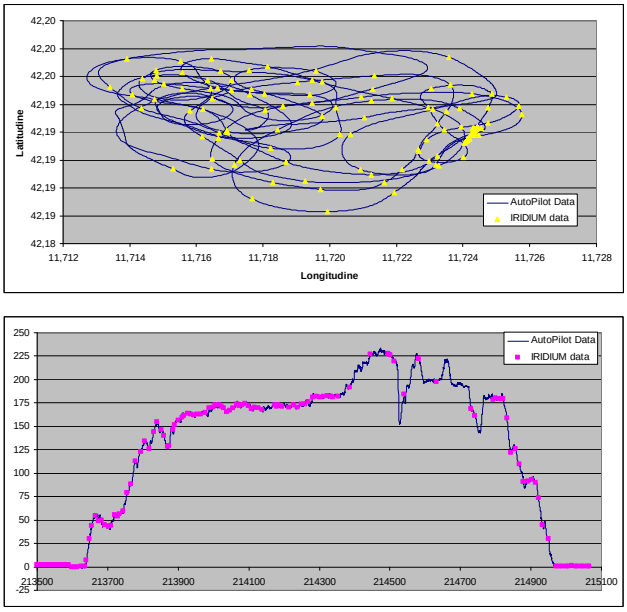


Fig. 17 – Recorded Flight path & Altitude (Autopilot and Iridium) of SESA UAV

5 Flight test of fuel cell powered aeroplane

The main critical topic to be solved for a very-long endurance UAV flight should be the electric power source generated by fuel cell system. RAPID 200-Fuel Cell, **first aeroplane in Europe and in the World** powered by a fuel cell fuelled by hydrogen, funded by European Commission and coordinated by Prof. Romeo of Politecnico di Torino, successfully concluded flight tests. **(Fig.18).** The speed and endurance world records were established for Electrically-

powered Aeroplane (FAI Sporting Code - Class C) Fuel Cells operated (ZERO EMISSION).



Fig 18 - Rapid 200-FC. Flight test

The completely electrical power system was successfully tested during the experimental flights. The rotation speed of 84 km/h was obtained within 184 m of taxi at power of 35 kW. After take-off, speed was than increased up to indicated values of 110-120 km/h. On 26th May 2010, a new speed world record of 135 km/h (flown of 4 consecutive runs over a 3 km course - according to FAI Sporting Code, Class C-Aeroplane) was established. Also several higher speed at 145-150 Km/h for tens of seconds were measured during free flights. The flights, fuelled by gaseous hydrogen at 350 bar pressure, reached an endurance of 40 minutes. The previous record was established by Boeing Research & Technology Centre (Madrid) in their world first hydrogen flight (110 km/h for 20 minutes with a motor-glider, Class D, FAI Sporting Code); also DLR flew in 2009 with Antares-H2 motor-glider (Class D, FAI Sporting Code) powered by fuel cells. The results obtained during flights can be considered as a further step in the European and World Aeronautics Science in introducing a completely clean energy [11-14].

The aircraft has an entirely electric 40 kW propeller; power is supplied to the propeller through 20 kW hydrogen gas fuel cells. To guarantee absolute safety of the operations, the aeroplane also has a second source of energy that consists of a set of 20 kW Li-Po batteries which are able to guarantee alternative or supplementary power during take off and initial climbing. The electric engine is fed through the generation of high energy currents in an ionization and hydrogen re-combination system (PEM) which has a final product of 100-110 Amps of electrical current at 200-240 V.

The aircraft and the electric and energy system were developed by the team which consists of the 9 companies and enterprises involved in the European Community “ENFICA-FC” project (**Environmentally Friendly Inter City Aircraft powered by Fuel Cells).** The team was made up of the Politecnico di Torino (IT) (Co-ordinator & Design of the modified aircraft and experimental test flights), Skyleader (CZ), Intelligent Energy (UK), APL (UK), Mavel Elettronica (IT). IAI, ULB, UNIPI and Evektor

The objective was that of building an aeroplane that works on hydrogen, taking advantage of the “fuel cell” technology at present available to create a demonstrator aircraft that is able to connect cities through flights while totally eliminating the environmental impact.

At the same time, more theoretical type studies have been carried out. These will not have an immediate practical application in the initial stages because of the present technological limits, but have the aim of using zero emission propellers in the future to equip aircraft for 20-30 passengers in the regional and intercity sector.

The starting up, functioning under power, taxiing and flight tests of the aircraft with its definitive propeller were carried out along the 1400 metre runway at the Reggio Emilia airport. The hydrogen at high pressure (350 bar), the 20 kW LiPo battery set and the electric engine have proved to be reliable and offer continuity of supply of the more than 40 kW power necessary for take off. The sophisticated electronic control system ensures redundancy and makes both sources of energy available to the pilot in order to guarantee the safe functioning of the single propeller, even in the case of breakdown of the hydrogen current generators. The airplane (whose final lay-out was achieved with the technical assistance of the Italian Skyleader importer – T&T Ultralight) has a wing opening of about 10 metres; with the systems at present available, the airplane has autonomy of 1 hour and can reach a cruising speed of 150-180 km/h, thanks to the hydrogen alone.

6 Conclusion

The following conclusion should be issued as result of the many years of research in the field of solar powered UAV:

- Possible realisation of VESPAS-UAV at least for low latitude sites in Europe and for 4-6 months continuous flight.
- Forest Fire Early detection, Border Patrol and Fishery monitoring would be possible at much cheaper cost and higher resolution than actual systems, and it would be obtained continuously.
- Showed feasibility of very light CFRP structural elements. Good correspondence between experimental, analytical and FEM analysis.
- The experimental tests successively validated several critical technologies for high altitude very long endurance flight: high efficiency solar cells, electric brushless motor, controllers, video and thermo camera image transmission, telemetry system, Solar Power production and control, Wireless transmission
- Autonomous flight control of UAV as by autopilot as well as by satellite communication.
- The feasibility of an UAV system, integrated with a satcom system was shown, to achieve Fisheries patrolling in a faster, accurate and more reliable manner.
- For small UAV limited in payload weight, it should be necessary to improve the image compression system to transmit the required images by satcom system.
- The extensive experimental flight campaign carried out during the ENFICA-FC project, proved that fuel cell technologies represent a promising future innovation in aeronautics as a key-enabling technology for all-electric, zero emission, low noise aircraft.

7 Acknowledgements and References

The authors acknowledge the important contribution of European Commission by the funding programmes:

1. **HELINET: "**Network of the Stratospheric Platforms for Traffic Monitoring, Environmental Surveillance and Broadband Services". **EC 5° FP**–2000-03; Leader of the HeliPlat Design and Manufacturing: Prof. G. Romeo
2. **CAPECON: "**Civil UAV Applications & Economic Effectivity of Potential Configuration Solutions". **EC 5° FP**. - WP Leader of 3 HALEs Design: Prof. G. Romeo.
3. **ENFICA-FC: "**ENvironmentally Friendly Inter City Aircraft powered by Fuel Cells. **EC-6FP**-2005-30779-AERO-1, Ott. 2006 - 09. Coordinator: Prof. G. Romeo
4. **TANGO***:* Telecommunications Advanced Networks for Gmes Operation (SIP5-CT-2006-030970-2006-09). Coord. EADS Astrium.
5. Romeo G., "Design of High Altitude Very-Long Endurance Solar-Powered Platform for Earth Observation and Telecommunication Applications". *J. Aerotecnica Missili e Spazio***,** Vol.77, N.3-4, July-Dec.98. pp. 88-99.
6. Romeo G., Frulla G., Cestino E., Corsino G., "HELIPLAT: Design, Aerodynamic and Structural Analysis of Very-Long Endurance Solar Powered Stratospheric UAV". *AIAA Journal of Aircraft*, Vol.41, No.6, Nov-Dec 2004, pp. 1505-1520.
7. Romeo G., Frulla G., Cestino E., "Design of high altitude long endurance solar powered unmanned air vehicle for multi-payload and operations", *Journal of Aerospace Engineering Part G., Proc. IMechE*, Vol.221, 2007, pp.199-216.
8. Tuzcu I., Marzocca P., Cestino E., Romeo G., Frulla G., "Stability and Control of a High-Altitude-Long-Endurance UAV", *Journal of Guidance, Control, and Dynamics*, Vol.30, No.3, May-June 2007, pp.713-721.
9. Romeo G., Pacino M., Borello F., "Autonomous Flight of Solar Powered Light UAS". *UAS 2010.* Paris, June 14-18, Blyenburgh & Co., Paris.
10. *http://www.polito.it/grupporomeo*
11. *http://www.enfica-fc.polito.it/.*
12. Romeo G., Borello F., Correa G., "Set-Up and Test Flights of an All-Electric Aeroplane Powered by Fuel Cells", *AIAA Journal of Aircraft*, 2011, Vol. 49, N.8 (July-Aug. 2011). DOI: 10.2514/1.C031271
13. Romeo G., Borello F.: "Design and Realization of a 2-seat Aircraft Powered by Fuel Cell Electric Propulsion". *The Aeronautical Journal*, May 2010, Vol. 114, No 1155, pp. 281-297., ISSN-0001-9240.
14. Romeo G., Cestino E., Borello F., Correa G.: "Engineering Method for Air-Cooling Design of Two-Seat Propeller Driven Aircraft Powered By Fuel Cells", *J. of Aerospace Eng.*, 2011, Vol. 24, pag. 79-88, ISSN: 0893-1321.

© Copyright 2011 by G. Romeo. Published by AIDAA/CEAS with permission. All informations are Propriety of Prof. G. Romeo, Politecnico di Torino and cannot be reproduced without his authorization. All right reserved.

3rd CEAS Air&Space Conference
21st AIDAA Congress

The return of Diesel engine in the aeronautic propulsion

Mauro Varetti,
Avio S.p.A. Torino Italy
Sara Bagassi
University of Bologna, Italy

Keywords: *Propulsion, Diesel engine*

Abstract

The recent developments in piston engine technology have increased performances as specific power and specific fuel consumption; environmental concern and increasing fuel cost are the main driving factors of this evolution.

New generations of turbo-diesel engines have been developed for automotive applications, being now competitive with their gasoline counterpart.

The use of Diesel engine for aeronautical propulsion is a discussed topic. In the past few years there have been a number of attempts to produce diesel engines for aircraft/rotorcraft and to this aim the three most important objectives to be pursued are reliability, compactness and low weight to power ratio.

In this paper an outline of the most challenging aeronautical applications for diesel engines and an overview of the most promising technical solutions are provided.

1 Historical background

The applications of diesel engines to the aeronautical propulsion belongs to the aeronautical era since the beginning of the flight.

In 1914 the German company Junkers built the first prototype of a Diesel engine for aeronautical propulsion applications, the 4 cylinder MO3. This prototype never flown but was the forerunner of the well known Jumo Diesel engine family used on several German aircraft like Do 18 flying boat equipped with the Jumo 205 that obtained in 1938 the record of the longest non-stop leg flying from Devon (England) to Caravelas (Brazil) for a distance of 8391 km.

Worth to be mentioned is the Junkers 86 P equipped with the turbocharged Jumo207 A1 capable to fly normally at 12000 m, higher than the ceiling capability of the enemy chase; only in August 1942 there is a record of the first Junkers 86 P hit by a modified Supermarine Spitfire MkV while flying at 14500 m.

A version with ceiling at 16000 m was later developed but the end of the war came first of the operational service.

Also in the USA Diesel engine application were developed like the Packard DR-980 [pic2]. The first flight is reported in the 1928 powering the Stinson SM-1DX "Detroiter" aircraft, while in 1931 a Bellanca CH-300 equipped with this engine flown non-stop for 84 hours and 32 minutes. This record was surpassed only 55 years later during the Rutan Voyager non-stop round the world trip.

These two engines had very different architectures.

The Jumo was the most unconventional and innovative anticipating technical solutions that still now inspire new aero Diesel engines designs.

The solution of the opposed cylinder in a common can arrangement is peculiar. Since the Diesel engine requires much more compression ratio than the spark ignition one, the stroke must be significantly higher which in turn implies an average piston speed higher.

To minimize the piston speed and the relevant mechanical tribological drawbacks the idea was to split the stroke in two parts assigning to each piston the task to accomplish part of the work and at the same time to open and close respectively the exhaust port and the flushing port in a uni-flow arrangement.

Hence in the '30 diesel engines applied to aeronautic propulsion gave impressive performance, still now unsurpassed, in terms of specific fuel consumption respect to other types of engines.

The Diesel engine powered aircraft were able to fly longer "legs" and guarantee longer flight endurance; at the same time, due its inherent architecture and thermodynamics, the Diesel engines were capable to anticipate performance in terms of climb rate and ceiling altitude reached first by turbo-compound spark ignition engines and turbine engines later on.

The main drawback of the Diesel engine, compensated by better specific fuel consumption, was the lower power to weight ratio.

After the 2nd WW end, the aircraft commercial speed and the single engine power requirements, were possible only using turbine jet powered aircraft.

A remarkable application of Diesel cycle to an aeronautic engine in a complex turbo-compound arrangement was the magnificent Napier Nomad which expressed, at that time, the best specific fuel consumption, with a power to weight ratio of 1,4 kW/kg delivering an equivalent shaft horse power of 2338 kW (3135 hp) SL ISA.

The emphasis placed at that time on the aircraft commercial speed killed first the propeller based propulsion and in turn all the relevant engine applications.

Hence the propeller driven aircraft were almost discontinued, except for the General Aviation, where the requirement of lighter powerplants like the spark ignition high octane gasoline fuelled engines, were preferred against the heavy Diesel engines.

The typical usage of such aircrafts with limited requirements on range did not put emphasis on the competitive advantage of Diesel engine in terms of specific fuel consumption.

Low fuel cost played a big role in the evolution of the technology in this field as well.

2 The Diesel engine evolution in non aeronautical application

The Diesel engine evolution continued only in the ground and marine application where the weight and dimensions were not an issue.

In the automotive application due to the lower fuel consumption the Diesel engines has been widely used for trucks, agricultural equipment, trains and fishing boats powerplants especially in Europe were gasoline was often much more expensive than gasoil due to different end user taxation rules.

The oil crisis, with severe increase of $/barrel price of crude oil, drove the evolution in the propulsion both in air and surface transport applications.

On the aircraft side, aeronautical engineers "re-discovered" the propulsion efficiency; this has been enhanced through the following years both with turbofan with higher and higher by pass ratio and the return to the application of propeller driven low speed regional aircraft.

In the automotive field, thanks to the wider market in continuous increase, technical evolution and car application simply exploded to such an extent that today every car manufacturer offers the same model of car

equipped both with petrol and gasoil fuel an some time with different version of the engine.

The original Diesel engine penalties were fought to make this engine as performing as the equivalent spark ignition engine at least for private car use.

Note that recently also some car competitions like the long endurance high speed ones (Le Mans, Daytona, Sebring) are now regularly won by diesel powered sport cars.

The power to displacement ratio, originally lower in the Diesel engine when compared to spark ignition one, has been successfully fought adopting the turbo-charging technology after the positive experience gained on other applications like the already mentioned aeronautical one.

Note that the same technology applied to a spark ignition engine does not give the same advantages because of the total allowable pressure ratio (turbo-charger+ volumetric).

As a result either engines with the same power can be mounted into the same car hood avoiding expensive customization.

Cold starts have been made possible with the heated glow used during engine turn on.

As a matter of fact it is much more frequent that a spark ignition engine has troubles in the ignition chain than a compression ignition engine. This inherent reliability of Diesel engine is a remarkable plus for the aircraft applications.

Even though not necessary to win the competition against the spark ignition engine, the direct fuel injection into the combustion chamber has been introduced in the diesel engine to assure better fuel consumption with respect to the indirect injection into a prechamber.

This solution adopted first in the commercial application like engines for trucks migrated later on into cars.

The drawback of this solution is the excessive noise (hammering) due to the the uncontrolled fuel burn chemistry.

Finally, the innovation that boost to the maximum extent the Diesel engine diffusion is the common rail technology.

Before the fuel injection into the cylinder was possible thanks to small pumps camshaft driven in synchronization with crankshaft with fuel pressure depending on the speed.

Since fuel flow must be modulated not only as a function of rotational speed but also according to the power requirement that can vary at constant speed, the old fuel flow supply and metering system were unable to cope with the optimum fuel injection law.

So fuel burn inefficiency and pollutants were present outside the optimum compromise design point of such old systems.

In 1988 a patent was issued by Magneti Marelli with the idea of decoupling the fuel flow requirement from the fuel pressure requirement being the latter obtained by the accumulation of fuel pumped at very high pressure into a common small tank and leave the metering of the right amount of fuel into each individual chamber piloted by an electronic control unit.

Nowadays all the Diesel engines have the common rail distribution system as standard equipment.

This invention made hence possible to modulate the fuel flow independently from the crankshaft angle and speed, by the timing of opening and closing of the valves and even to modulate the fuel flow during the combustion for optimum burning taking into account several parameters (temperatures, fuel composition, power settings etc..).

It is worth to be noted that all these evolutions took place around the four stroke diesel engine arrangement but most of them can be transferred to two stroke arrangement for lower rotational speed not geared application like the aircraft propeller.

Aside of these innovation peculiar of Diesel engine application, other innovation took place in the reciprocating engine technology that can be incorporated into an aeronautical engine.

The original struggling of the cold/hot arrangement of the camshaft, rod, roster and valve, was solved with the introduction of the play recovery with hydraulic compensation in every condition, irrespective of differential temperature transients or even wear of the mating parts.

The evolution of this idea leads to the possibility of changing the opening/closing timing of the valves changing hydraulically the length of the valve stem, in a way to virtually change the cam profile (FPT Multiair).

Last but not least, a sensible improvement has been achieved in the materials and manufacturing technology of the reciprocating engine parts.

As a conclusion of this step of Diesel development, the aircraft propulsion can benefit of aeronautical diesel engines with:

- Even better fuel consumption
- Environmental emission (no leaded fuel)
- Comparable power/weight ratio (w.r.t. S.I.)
- Higher reliability (eg no ignition system)
- Higher safety (used fuel less hazardous)
- State of the art technology well advanced.
- Growth potential still to be exploited.

3 The specific advantages of Diesel and the promising applications

The recent developments in piston engine technology have increased performances as specific power and specific fuel consumption; environmental concern and increasing fuel cost are the main driving factors of this evolution. New generations of turbo-diesel engines have been developed for automotive applications, being now competitive with their gasoline counterpart.

The above described common rail technology, originally developed by Centro Ricerche Fiat and Magneti Marelli, put the diesel performances at a competitive level with spark ignition engines. Diesel turbocharged/turbo compound engines, burning diesel or jet fuels, have great performances; last generation of automotive common rail diesel has very flat torque and efficiency curves.

In the past few years there have been a number of attempts to produce diesel engines for aircraft/rotorcraft and to this aim the three most important objectives to be pursued are reliability, compactness and low weight to power ratio.

In the past, emissions were not very critical in aviation applications, since requirements were not as stringent as in automotive applications and power output is kept very close to the maximum all flight long. Nowadays environmental concern must be taken into account in developing new aeronautical engines.

Fuel management and combustion control system are critical for optimum efficiency. Since common rail technology relies on a electronic control unit, governing a high pressure injection systems and supercharging devices, the development of a FADEC (Full Authority Digital Engine Control) complying with aeronautical standards is the key factor to set diesel technology as a good candidate for aeronautical applications for power ratings in the range of 100-450HP. The development of such FADEC units allows also a full integration with modern avionics and advanced strategies for managing failures.

Diesel engines have now reached a technology level that is appealing also for helicopter applications. Currently almost all helicopters with installed power above 250HP have turbo-shaft engines. These engines are very reliable and have a specific power out of reach for piston engines, on the other hand they are quite thirsty for fuel.

Today's diesels are quiet, and the specific fuel consumption is cut by 40 percent compared to existing turbine engines. So fixed and rotary wing applications requiring low-medium power, and high range/duration performances (as in UAS) are the best candidates for Diesel technology, reducing operating cost and environmental impact. Another advantage is the better performance of diesels at high altitude, compared with the decrease in power with altitude of turbo-shafts, requiring a higher power rating to compensate losses due to reduced air density at altitude.

As a consequence Increased consideration is given for using diesels in unmanned aircrafts (both fixed and rotary wing UAS), were low SFC is a key factor for achieving long

endurance and FADEC technology fits well with advanced Flight Management Systems.

Also for General Aviation this technology is appealing: as stated in (1) "retrofitting a diesel engine to run on Jetfuel or Kerosene, reduce Gallons/Hour by some 30%, eliminate ignition systems (magnetos, spark plugs) and their problems, eliminate mixture control, increase TBO to 2,400-3,000 hours, increase performance between 6,000 and 12,500 ft., and drastically reduce Operating Costs".

In Europe increased gasoline highly taxed fuel costs, compared to the JP4, JP8, and Jet (A1) engines, help to revive interest in the diesel concept. Moreover the good advantage for the aeronautical development is also in safety, in fact JP4, JP8, Jet (A1) and diesel are less inflammable than gasoline.

Following the dramatic improvement of diesel propulsion over the past two decades, and the success gained in the automotive industry, some diesel engines have been developed, certificated for aeronautical applications and entered into production, with some success in equipping general aviation fixed wing aircrafts.

Nevertheless, state-of-the-art of aeronautical diesel engines comprises a number of studies, many prototypes, some demonstrators and a limited number of "products".

While the more advanced diesels in production are for automotive applications and are four stroke, with in-line o V shaped architectures, when new aeronautical diesels are designed "from scratch", the range of proposed solution is far more variable, ranging from two-stroke to four stroke, Boxer, opposed-piston, or even opposed-piston-opposed cylinder, Wankel, liquid or air cooled, etc. This is due to the influence of theoretical performances levels and solutions already implemented in the past (see the famous Junkers diesels), but some of them are affected by serious technical drawbacks, as demonstrated by the abortion of some programs, as for example the GAP diesel engine of Teledyne Continental Motors and it's industry team partnered with NASA Glenn to develop a highly advanced piston engine.

Looking at the few studies that have been able to become products, sometimes certified, it is evident that up to now the successful approach is based on:

- automotive derived engine taking advantage of reliability of the components, reducing development and certification costs, reducing also the unit cost of many parts which are produced I volumes achievable only in the automotive industry
- four-stroke engine (as opposed to two-stroke
- liquid cooled engine (as opposed to air cooled): improving thermal control reducing thermal stresses and easing a lightweight construction

A new interest has emerged for two-stroke diesels: some new programs are under development with a conventional opposed cylinder architecture or the revolutionary OPOC (Opposed Piston Opposed cylinder) engine. It has two opposing pistons in two contiguous cylinders, connected to a common crankshaft in the middle of the engine. An Electrically Controlled Turbocharger (see next paragraph) provides boost on demand. It's a two-stroke engine with no valves, yet still achieves 90% scavenging efficiency with less oil consumption than a four-stroke engine. In other words, it's a two-stroke engine that can meet the strictest emissions standards. It can be made as a spark-ignited or compression-ignition engine, and the diesel version can meet emission standards without using urea.

The manufacturer (Eco Motors LtD) claims that it's a design which is half the size and uses half the parts of a conventional piston engine. It can be built for 20% lower cost and 30% lower investment than traditional internal combustion engines (ICE's). And it should provide a 15% improvement in fuel economy over current advanced designs.

4 Technical evolutions

Electrically Controlled Turbocharger

This development in turbocharger technology incorporates an electric motor into the turbo assembly. In essence, it provides a supercharger, driven by the electric motor, as an adjunct to the exhaust-driven turbocharger. Boost pressure can be created by the electric motor, the turbocharger, or both. The ECT effectively eliminates turbo lag because the electric motor provides much faster turbine response, and also provides boost when there is low energy from the exhaust flow. The motor is actuated by an electronic controller, which can be integrated with the engine control unit. When it is being spun by the turbocharger, the electric motor acts as generator, producing electricity.

Hybrid engines

EADS Innovation Works presents model of a concept helicopter with a diesel-electric hybrid propulsion system as part of the Group's research to make helicopters more environmentally friendly.

The diesel-electric hybrid concept is one of the projects that are grouped under the name of eCO2avia by EADS Innovation Works. Highly efficient electrical motors driving the rotors, combined with OPOC (Opposed Piston, Opposed Cylinder) diesel engines, reduce fuel consumption and emissions by up to 50%. Take-offs and landings are possible on electrical power alone, resulting in lower noise levels and improved flight safety.

The main components of this hybrid system are multiple diesel-electric motor-generator units, a pair of high-performance batteries and a power electronics unit controlling the energy flows for best efficiency. The OPOC diesel engines, designed and built by EcoMotors International in the US, offer a fuel economy improvement of up to 30% compared to today's helicopter turbine engines.

The OPOC engine's power output shafts are fitted with advanced, weight-optimised generators delivering electrical current to a power electronics unit, which manages the distribution of the electricity to the electrical motors driving the main rotor and the tail rotor as well as the other user systems on the helicopter. High-performance batteries can store sufficient energy to enable the helicopter to take off and climb or approach and land on electrical power alone.

Several kinds of combustion engines could be integrated into such a hybrid system.

Compound Cycle Engine

As an example of this evergreen concept, we cite the NASA Compound Cycle Engine (CCE) Program (1986) [2].

This CCE is a highly turbocharged, power compounded power plant which combines the lightweight pressure rise capability of a gas turbine with the high efficiency of a diesel.

When optimized for a rotorcraft, the CCE will reduce fuel burned for a typical 2 hr (plus 30 min reserve) mission by 30 to 40 percent when compared to a conventional advanced technology gas turbine. The CCE can provide a 50 percent Increase in range-payload product on this mission.

5 The present and the future global scenario

The main drivers of the transportation (ground, sea, sky) scenario are fuels availability and price; at the same time environmental sustainability must be taken into account as well.

The EU has set up a system to control the total amount of CO2 emission named Emission Trading System (ETS) with total cap, reduced year by year, and a system of prizes and penalties to promote in all the companies and nations the global reduction.

So the improvement in CO2 emission will be driven not only by the pure economic trade off between fuel price and engine amortization cost but also by the need to comply with ETS system

either purchasing or selling emission rights or to avoid fines for threshold limit violations.

Similar systems are going to be extended in other countries like USA .

As matter of fact in 2009 the General Aviation Manufacturers Association (GAMA) joined the operator community policy represented by the International Business Aviation Council (IBAC).

They committed to the following steps to reduce emissions:

- Carbon-neutral growth by 2020;
- An improvement in fuel efficiency of
- an average of 2 percent per year from
- today until 2020;
- A reduction in total carbon emissions of
- 50 percent by 2050 relative to 2005.

Now it has been demonstrated that the Diesel specific fuel consumption is much better both at optimum design point and even better at partial loads. An overall fuel efficiency with respect to the Avgas fuelled engines of 30% is hence immediately achievable simply replacing the engine.

If we can figure that all the existing engines are suddenly converted to Diesel cycle, the General Aviation target contribution to the above objective is almost achieved since now.

Furthermore taking into account that for power level up to 500 HP Diesel engine can compete also with small turbo-shaft on the basis of specific fuel consumption and cost of ownership, one can imagine a scenario in which Diesel engines, conceived for General aviation (fixed and rotary wing and airships as well), can give a short medium term solution readily available now.

Short References

[1] http://www.dieselair.com/

[2] Compound Cycle Engine Program, NASA Technical Memorandum 88879, 1986

CEAS 2011

Flight Dynamics and Flight Testing

Gyroplane Longitudinal Flight Dynamics

H. Duda, I. Pruter, C. Deiler, H. Oertel, A. Zach
German Aerospace Centre DLR, Institute of Flight Systems, Department Flight Dynamics and Simulation, 38108 Braunschweig, Germany

Keywords: *flight dynamics, gyroplane, gyrocopter, handling qualities, modeling and simulation*

Abstract

This paper describes recent research activities at DLR's Institute of Flight Systems on the flight dynamics and performance of modern lightweight gyroplanes. The main causes of hazardous situations or accidents with gyroplanes are discussed. DLR's gyroplane simulation model is presented briefly; it contains subsystems for the gyroplane body and rotor aerodynamics, the landing gear, the engine/propeller, the control system, and the environment. A special flight test program has been conducted with a gyroplane of the type MTOsport operated at DLR to gather data for model validation. The model validation process aiming at the application in a flight simulator for pilot training is described briefly. The flight stability and controllability is analyzed by means of eigenvalues and transfer functions. Some often discussed topics are investigated like the effect of the tail plane, the vertical location of the propeller thrust line in relation to the centre of gravity and the issue of pilot involved oscillations.

Nomenclature

α_R	Rotor plane angle of attack
a_z	Vertical acceleration
β_R	Rotor flapping angle
β_{1c}	Rotor longitudinal flapping angle
δ_{stick}	Pitch stick deflection
F_{stick}	Pitch stick force
F_R	Rotor force
γ	Flight path angle
h_{CG_PPB}	Vertical distance between CG and PPB
h_{TB_PPB}	Vertical distance between TB and PPB
h_{CG_Prop}	Vertical distance between CG and Propeller
I_y	Body pitch moment of inertia
I_R	Rotor moment of inertia around TB
m	Gyroplane mass
m_R	Rotor mass (rotor head and two blades)
m_{blade}	Rotor blade mass
l_{CG_PPB}	Horizontal distance between CG and PPB
l_{CG_stab}	Horizontal distance between CG and horizontal stabilizer
l_{TB_PPB}	Horizontal distance from TB to PPB
$M_{R,flap}$	Rotor aerodynamic flapping moment
$N_{R,aero}$	Rotor aerodynamic rotational moment
$N_{R,ext}$	External rotor rotational moment
r_{CG_TB}	Distance between CG of each blade and TB
r_R	Rotor radius
Θ	Body pitch angle
$\Theta_{R,bl0}$	Rotor blade pitch angle
$\Theta_{R,tilt}$	Rotor head tilting pitch angle in body fixed system
$\Theta_{R,rh}$	Rotor head tilting pitch angle
Θ_R	Rotor plane pitch angle
t_R	Rotor blade chord
V_{TAS}	True airspeed
w_{DW}	Rotor downwash velocity
$w_{R,eff}$	Effective vertical airflow through the rotor plane
w_R	Vertical component of the true airspeed with

	respect to the rotor plane
ω_R	Rotor rotational speed

Abbreviations

BCAR	British Civil Airworthiness Requirements
CAA	Civil Aviation Authority
CG	Centre of Gravity
DLR	German Aerospace Centre
GPS	Global Positioning System
PIO	Pilot-Involved Oscillations
PPB	Pitch Pivot Bolt
PPO	Power Pushover
PTL	Propeller Thrust Line
RPB	Roll Pivot Bolt
TB	Teeter Bolt
UK	United Kingdom
rpm	Revolutions per Minute

1 Introduction

The boom of lightweight gyroplanes in Germany triggered by the introduction of the MT03 in the year 2004 initiated a research program at DLR's institute of flight systems. Research goal is the further improvement of gyroplane flight safety by means of developing a specific flight training device and further understanding of its very unique flight dynamics.

A gyroplane is an aircraft that gets lift from a freely turning rotary wing and which derives its thrust from an engine-driven propeller. Historically, this type of aircraft has been known as autogyro and gyrocopter. It was developed by Juan de la Cierva and in 1923 it was the first rotary wing aircraft flying. Early gyroplanes were powered by engines in a tractor (pulling) configuration and were relatively heavy, like the C-30 from 1932 [1], Fig. 1.



Fig. 1 De la Cierva's C-30 (1934).

The first gyroplane flying – the C-4 - was fitted with a nontilting rotor and a set of ailerons mounted on a stub spar projecting from the sides of the fuselage. Pitch and directional control was then achieved by conventional airplane surfaces, with an elevator and a rudder used at the tail.

After significant improvements in helicopters, gyroplanes became largely neglected. In the 1950s there was some revival of interest in the gyroplane by Igor Bensen's home-built gyroplane kits with an open airframe and the Fairey Company in Britain. Two *Fairey Gyrodyne* prototypes led to the Rotodyne, which was the world's biggest gyroplane with a cabin big enough for 40 passengers. During early 1960s, single- and two-seater gyroplanes were developed for the private aviation market; one of it gaining a starring role in a 1967 James Bond film – the *little Nellie*, a Ken Wallis WA-116.

1.1 Recent Activities

Recently, there have been two companies - *Carter Aviation Technologies* and *Groen Bothers Aviation, Inc* - in the United States that have begun to exploit gyroplane capabilities using modern technologies. The Carter test vehicle is a hybrid aircraft using a rotor and a fixed wing. The rotor provides nearly all of the lift during takeoff and landing, the wing produces most of the lift at higher airspeeds[1]. Conventional flight control surfaces (ailerons, elevator, and rudder) are used. *Groen* has developed the world's first turbine-powered gyroplane - the Hawk 4 - providing short takeoff and nearly vertical landing capabilities. It uses a swash plate with collective and cyclic pitch, which gives the aircraft excellent control and manoeuvrability [1]. Unlike the Carter machine roll and pitch motions are achieved through rotor control.

In Europe several manufacturers sell single- and two-seater gyroplanes for the private aviation market, such as Spanish *ELA Aviation*, Italian *Magni Gyro*, German *Rotortec GmbH* or *AutoGyro GmbH*; the latter produced and sold almost thousand machines of the types MT03, MTOsport and Calidus since 2004. This gyroplane boom in Germany can be explained by the fascinating flying characteristics of a gyroplane in combination with the performance and robustness of the MT03/MTOsport vehicles which is even toped by the Calidus, Fig. 2.



Fig. 2 AutoGyro's Calidus (2009).

1.2 Airplane or Helicopter?

The body structure of a gyroplane is closer to a helicopter than to an airplane due to the rotor system, which is however, of significantly lower complexity as in a helicopter: no transmissions, gearboxes, tail rotors or drive shafts are needed. It is this simplicity that makes a gyroplane a relatively inexpensive flying vehicle. A

gyroplane has horizontal and vertical tail planes with a rudder as well as a main engine providing forward thrust by a propeller like an airplane. Additionally a gyroplane performs more like an airplane in terms of speed, range, payload, stability, and safety.

From a cruising flight performance point of view the gyroplane is something between a helicopter and an airplane: over a speed range from approximately 30 to 70 mph, the power required by the gyroplane is predicted to be 30% less than that for the helicopter [2]; compared to an airplane its cruising flight performance is poor due to the significantly lower glide ratio: an aerodynamically effective gyroplane (like the Calidus) may have a glide ratio of about five while a comparable fixed wing airplane may have twice as much or more.

1.3 Gyroplane Flight Controls

In the very beginning gyroplanes had been controlled in the same way as fixed wing aircraft that is by deflecting the air flowing over moving surfaces such as ailerons, elevator and rudder [1]. This technique provided limited authority especially at low airspeeds. The Cierva C-30 was the first gyroplane to have a tilting rotor hub and a control rod from the hub to the pilot's cockpit for controlling the rotor plane. This was known as *direct control*. Modern gyroplanes primary flight controls are

Roll and pitch control: Most of today's gyroplanes are controlled by tilting rotor systems. Fig 3 shows the rotor head of the MTOsport, which can be tilted around the *Pitch Pivot Bolt* (PPB) and the *Roll Pivot Bolt* (RPB) by control stick movements; by this the direction of the rotor lift force can be controlled. A pneumatic trim system provides additional control forces in order to improve flight comfort. No elevator is applied for pitch control; the horizontal tail plane is for increased pitch damping.

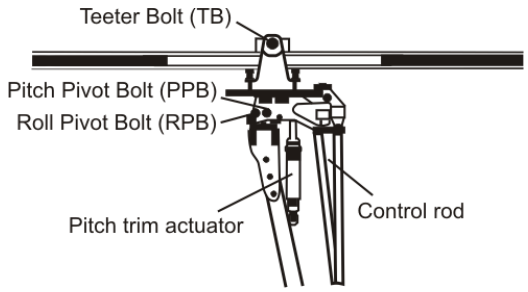


Fig. 3 AutoGyro's MTOsport tilting rotor system.

Yaw control: The rudder is operated by foot pedals in the cockpit which are additionally connected to the nose wheel. It provides a means to control yaw movement of the aircraft similar to the rudder of an airplane and is used to maintain coordinated flight, to compensate for yaw moments due to propeller slipstream effects and to perform crosswind landings. Consequently, many gyroplane rudders are located in the propeller slipstream and provide excellent control while the engine is developing thrust. This type of rudder configuration, however, is less effective and requires greater deflection when the engine is idled or stopped.

Engine power control: The throttle is conventional to most powerplants, and provides the means to increase or decrease engine power and thus, propeller thrust.

Collective blade control: The collective blade control provides a means to vary the rotor blade pitch angle, and is available only on more advanced (expensive) gyroplanes, such as *Groen*'s Hawk4. It allows jump takeoffs when the blade inertia is sufficient. Also, control of in-flight rotor rotational speed is possible to enhance cruise and landing performance, but not needed. The MTOsport investigated herein does not have any collective blade control.

2 Gyroplane Accidents and Hazardous Conditions

Though the accident rate of gyroplanes is quite high it is commonly accepted within the community that compared to other aircraft, the gyroplane is just as safe and very reliable. The most important factor, as in all aircraft, is pilot proficiency. Based on flight school experience the accident rate is relatively high for low pilot proficiency (which means less than 50 flight hours after examination) while above 50 hours the accident rate is low compared to lightweight airplanes.

The following hazardous flight situations or main accident causes are known from accident analyses and flight experience:

2.1 Pilot Involved Oscillations (PIO)

According to the US American military specification MIL-STD-1797 [3] "*… PIO are sustained or uncontrollable oscillations resulting from the efforts of the pilot to control the aircraft*". PIO can be considered as a closed loop destabilization of the aircraft-pilot loop. It is known since the early days of aviation for all types of air vehicles while the introduction of digital fly-by-wire flight control systems has increased its potential [4]. Time delays due to the flight control computers and rate limiting of the electro hydraulic actuators were identified to be the main cause of PIO problems with modern fly-by-wire aircraft [5].

In the case of gyroplanes the PIO problem is basically known to occur in the pitch axis and at high airspeeds. Its main cause is considered to be a low pitch damping due to a missing or too ineffective horizontal tail plane. The increased likelihood of PIO at higher airspeeds is because pitch control becomes more sensitive [6]. However, it is possible to design a gyroplane without any PIO tendencies also at higher airspeeds. No PIO accident or incident is known to the authors with MT03/MTOsport gyroplane types.

2.2 Power Pushover (PPO)

If rotor force is rapidly removed, some gyroplanes have a tendency to pitch forward abruptly. This is often referred to as a forward *tumble*, *buntover*, or *power pushover*. Removing the rotor force is often referred to as unloading the rotor and can occur if the nose of the gyroplane is pushed forward rapidly after a steep climb (or at high airspeeds). A PPO can occur on gyroplanes that have the propeller thrust line (PTL) above the centre of gravity (CG) and do not have an adequate horizontal tail plane [6]. In this case, when the rotor is unloaded, the propeller thrust magnifies the pitching moment around the CG. Furthermore a negative pitch moment (nose down) due to the propeller is generated caused by the decreased angle of attack. Unless a correction is made, this nose down pitching action could become self-sustaining and irreversible. An adequate horizontal tail plane slows the pitching rate and allows time for recovery.

2.3 Control Stick in Forward Position during Takeoff Roll

The takeoff procedure to be applied for a modern lightweight gyroplane with a tilting rotor system is as follows:

1. Move the control stick in full forward position.
2. Prerotate the rotor up to 200 rpm minimum.
3. Move the control stick to full aft position.
4. Release the brake and apply full throttle.
5. Accelerate and take off at about 300 rpm rotor speed.

The takeoff acceleration with the rotor disc tilted aft allows airflow through the blades to accelerate the rotor to flight rpm. If the pilot does not perform step 3 the rotor is not loaded and by this looses rotational speed; if in this case the pilot further continues the acceleration, severe blade flapping may occur which may lead to catastrophic situations, such as destruction of the gyroplane or uncontrolled takeoff with severe pitch reactions. Some fatal accidents occurred due to these obvious pilot errors [7].

2.4 Flying behind the Power Curve

If a gyroplane is slowed to the minimum level flight speed, it requires full throttle just to maintain altitude. Operating in this performance realm, sometimes referred to as the “backside of the power curve” can be hazardous at low altitudes, since no power reserve is available when the gyroplane starts to decent due to downwind. For this reason, it is essential to practise slow flight at altitudes that allow sufficient height for a safe recovery. Unintentionally flying a gyroplane on the backside of the power curve during approach and landing can be extremely hazardous. Severe accidents occurred recently by flying too slow at too low altitudes.

3 Data Gathering

For data gathering and simulation model validation an MTOsport gyroplane was used – the D-MTOS.

3.1 Instrumentation

The D-MTOS gyroplane has been instrumented by DLR according to the requirements for simulator validation, Fig. 4 [8], [9].



Fig. 4 MTOsport (D-MTOS) instrumented for simulator validation flight tests at DLR Braunschweig: Measurement equipment installed at rear seat location.

All relevant flight parameters were measured at 100 Hz: an inertial platform provides the roll, pitch and yaw rates, the corresponding attitude angles, the longitudinal accelerations, GPS positions, and the ground speed. A laser altimeter measures the height above ground. The roll and pitch stick forces are measured by strain gauges; the corresponding deflections by potentiometers. The pedal forces and deflections are measured as well. Taps of the basic avionic provide the rotor rotational speed, propeller rotational speed and the pneumatic trim pressure. The thrust lever position is measured by a spring potentiometer. Static and dynamic pressures are measured as well. In order to limit the instrumentation effort no nose boom and no sensors for rotor flapping angle are installed.

Together with the inertial platform a data acquisition system, a battery and an interface control box are mounted in place of the rear seat, Fig. 4. In good accessibility for the pilot a panel with switches and indicator lights is integrated, enabling him to power on the measurement devices, to start and stop the data acquisition and to supervise the state of the system.

3.2 Flight Tests

Special flights were conducted comprising several ground and flight trials to gather data about rotor motion, taxiing, steady state flight conditions, turns, decelerations, accelerations, ground effects, dynamic manoeuvres,

takeoff, touch down, gliding performance, and tests with variable pitch propeller [10]. All eleven flights and ground trials were conducted within the period of March to April 2010.

4 Gyroplane Modelling at DLR

4.1 Rotor Modelling

The rotor aerodynamics are calculated by the strip method, or blade element theory, such that the individual airflow at 10 blade elements is calculated considering the rotational speed, the vertical component of the true airspeed and the flapping motion. The rotor blades are considered to be rigid. The lift/drag characteristics of the NACA8H-12 airfoil known from literature are used as starting value [11]. Based on the MTOsport flight tests conducted its lift/drag characteristics are modified slightly such that the rotor model has a similar performance as the real rotor during vertical autorotation flight, which means

- rotor thrust of about 3850 N,
- rotational speed of about 300 rpm,
- vertical descending velocity of about 7.4 m/s.

The MTOsport is descending with 7.4 m/s at a weight of 3850 N (390 kg) during vertical autorotation.

An inflow model with constant induced velocity according to [12] is applied. The downwash velocity w_{DW} is a function of rotor radius r_R, the rotor thrust force F_R and the true airspeed V_{TAS}

$$w_{DW} = f(r_R, F_R, V_{TAS}) . \qquad (1)$$

The effective vertical airflow through the rotor plane $w_{R,eff}$ is determined by

$$w_{R,eff} = w_R - w_{DW} . \qquad (2)$$

The vertical component of the true airspeed with respect to the rotor plane is determined by

$$w_R = \sin\alpha_R \cdot V_{TAS} . \qquad (3)$$

α_R is the rotor plane angle of attack. In case of vertical autorotation the following numbers are obtained from flight test:

- $w_R = 7.4$ m/s by measurement,
- $w_{DW} = 6.6$ m/s by inflow model,
- $w_{R,eff} = 0.8$ m/s by Eq. (2).

The downwash velocity is decreasing significantly at higher airspeeds [12]. In case of the MTOsport gyroplane model the downwash velocity during steady state flight at 40 m/s true airspeed is almost zero.

The rotor rotational speed ω_R is calculated by the following equation

$$\dot{\omega}_R = \frac{1}{I_R}(N_{R,aero} + N_{R,ext}) . \qquad (4)$$

The rotor aerodynamic rotational moment $N_{R,aero}$ is determined by the sum of all tangential forces of all blade elements multiplied by their corresponding distances to the rotor shaft. The external rotor rotational moment $N_{R,ext}$ is considering the prerotator and the rotor brake systems.

The flapping angle β_R is calculated by the following equation

$$\ddot{\beta}_R = \frac{1}{I_R}(M_{R,flap} - 2 \cdot r_{CG\_TB}^2 \cdot m_{blade} \cdot \omega_R^2 \cdot \sin\beta_R) \qquad (5)$$

The rotor aerodynamic flapping moment $M_{R,flap}$ is determined by the sum of all vertical forces of all blade elements multiplied by their corresponding distances to the teeter bolt. The second term in this equation represents the flapping moment due to centrifugal effects.

4.2 Overall Simulation Model

An overall simulation model of the MTOsport gyroplane is developed at DLR [13, 14] aiming at its implementation in a training simulator. This simulation model is implemented under Matlab/Simulink based on the modelling standards at DLR's institute of flight systems. It contains subsystems for the gyroplane body and rotor aerodynamics, the landing gear, the control system, and the environment. A comprehensive model of the Rotax912 engine and IVOprop propeller is developed as well [15].

Besides the rotor model described above the main difference to fixed wing aircraft modelling is that the body (fuselage, mast, horizontal tail, and engine/propeller) and the rotor are modelled as two individual masses, which are connected at the pitch and roll pivot bolts. Furthermore it is mandatory to build the model of a gyroplane with tilting rotor system such that the pitch and roll stick forces are the input signals to the simulation and not the control positions.

In order to understand this issue consider the fuselage changing its pitch attitude due to an external disturbance while the rotor plane does not. In this case the control stick must move without a control force applied by the pilot. This situation can not be handled utilizing a model with stick deflection as input signal. More details about the modelling approach can be found in [13]. The rudder pedal and throttle position are used as input signal as well.

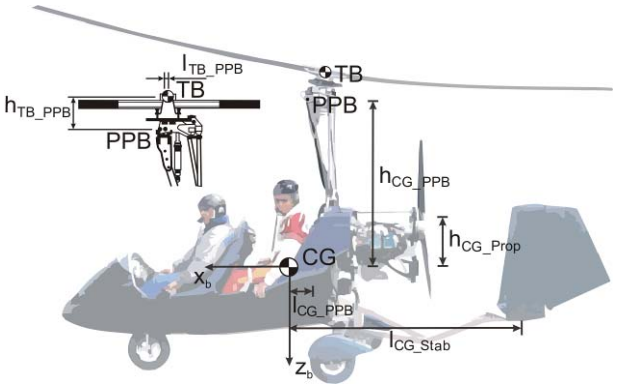


Fig. 5 MTOsport gyroplane, geometric definitions.

The relevant geometric parameters for the MTOsport gyroplane model for the longitudinal motion are presented in Fig. 5 and Tab. 1.

Tab. 1 MTOsport longitudinal simulation model parameters

Body data		
m	Gyroplane mass	390 kg
h_{CG_PPB}	Vertical distance between CG and PPB	1.59 m
h_{CG_Prop}	Vertical distance between CG and Propeller	0.27 m
l_{CG_PPB}	Horizontal distance between CG and PPB	0.06m
l_{CG_stab}	Horizontal distance between CG and tail plane	1.66 m
h_{TB_PPB}	Vertical distance between TB and PPB	0.21 m
l_{TB_PPB}	Horizontal distance from TB to PPB	0.024 m
I_y	Body pitch moment of inertia	350 Kg m²
Rotor data		
m_R	Rotor mass (rotor head and two blades)	32 kg
m_{blade}	Rotor blade mass	10 kg
I_R	Rotor moment of inertia around TB	150 Kg m²
r_{CG_TB}	Distance between CG of each blade and TB	2.38 m
r_R	Rotor radius	4.20 m
t_R	Rotor blade chord	0.20 m
$\Theta_{R,bl0}$	Rotor blade pitch angle	2.50 deg

4.3 Model Validation

The most relevant longitudinal flight parameters during straight and level steady state flight are presented in Fig. 6 in comparison to the simulation model.

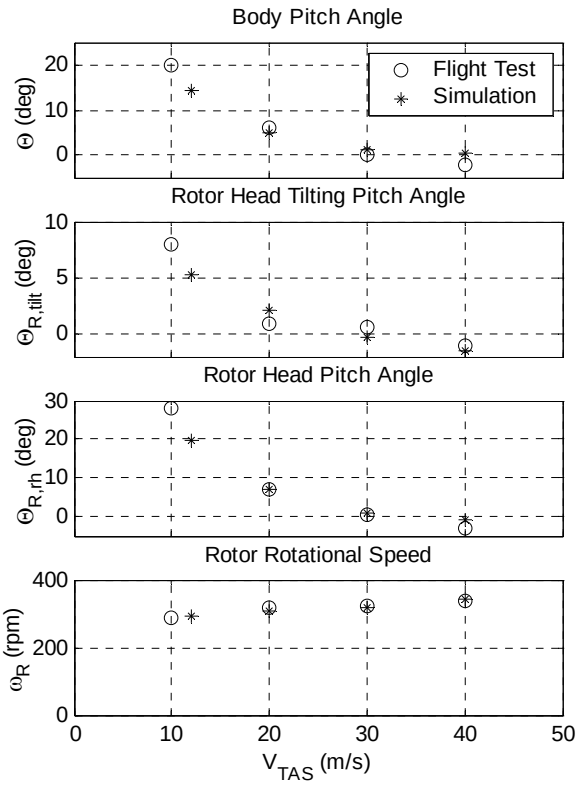


Fig. 6 MTOsport longitudinal steady state level flight parameters; flight test vs. simulation.

The rotor head tilting pitch angle $\Theta_{R,tilt}$ is not measured directly but calculated by the measured pitch stick deflection δ_{stick} based on the nonlinear transmission ratio obtained by ground calibrations. The rotor head pitch angle $\Theta_{R,rh}$ is determined by

$$\Theta_{R,rh} = \Theta + \Theta_{R,tilt} \qquad (6)$$

Fig. 6 indicates that the rotor head pitch angle is relatively small at higher airspeed regime; at 40 m/s true airspeed it is even negative, which is to be explained due to the decreasing downwash velocity at higher airspeeds [12]. However the rotor plane pitch angle Θ_R is relevant for its angle of attack

$$\alpha_R = \Theta_R - \gamma \qquad (7)$$

The rotor plane pitch angle is

$$\Theta_R = \Theta_{R,rh} + \beta_{1c} \qquad (8)$$

The additional rotor pitch angle due to flapping β_{1c} (which is not measured but simulated) is about 4 deg at 40 m/s true airspeed, hence the rotor plane angle of attack is slightly positive in this flight regime.

The rotational speed of the rotor and its change due to true airspeed is represented well by the (quite basic) simulation model; the rotational speed increases by about 50 rpm between 10 to 40 m/s true airspeed.

Several dynamic manoeuvres are used to validate the MTOsport model, such as doublets in the roll, pitch and yaw axes as well as dynamic throttle movements. As typical dynamic manoeuvre a pitch doublet is chosen, Fig. 7. The input signal for the simulation model is the measured pitch stick force. The output signals considered in this case are pitch stick deflection, pitch rate, vertical acceleration and rotor rotational speed. The simulation model represents these important flight dynamic parameters adequately well.

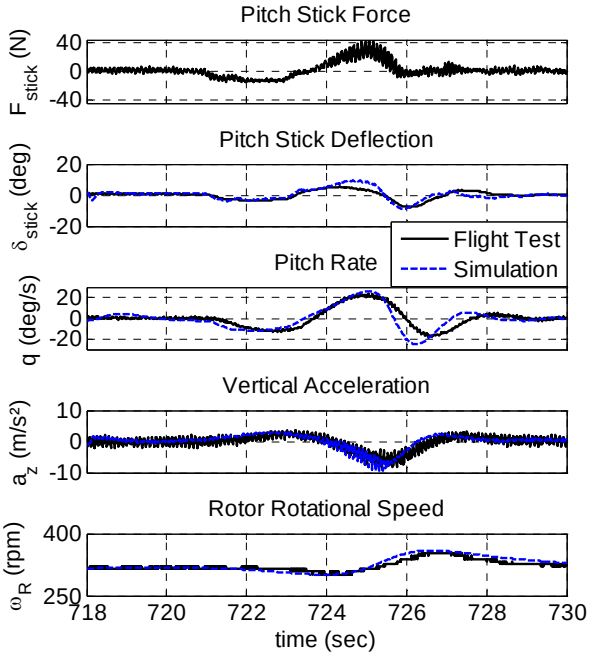


Fig. 7 MTOsport model validation: pitch stick doublet at 20 m/s true airspeed (m=390 kg).

5 Flight Dynamics Analysis

From a scientific perspective there have been few recent studies of gyroplanes flight dynamics. Research has been conducted recently at Glasgow University developing advanced mathematical models of the gyroplane validated by flight measurements conducted on specially instrumented gyroplanes like the VPM M16 [16]-[18] and the Montgomerie Merlin [19]. Further studies about the initial development and validation of a flight simulation model of a Rotorsport UK Ltd.. In [19] the following is stated:

- The most significant configurationally effect is the vertical location of the centre of gravity (CG) with respect to the propeller thrust line (PTL).
- The stability characteristics of a gyroplane can be considered as a mix of helicopter and fixed wing aircraft modes; gyroplanes can have a fixed-wing-like lightly-damped phugoid oscillation and a cross-coupling exists between longitudinal and lateral/directional degrees of freedom.
- Wind tunnel data showed that the tail plane aerodynamics have a very limited effect on longitudinal dynamics (at low air speeds).
- PIO tendency is most probably caused by the relatively high frequency, lightly damped or even unstable phugoid.

In the analysis presented below these statements are investigated amongst others based on DLR's flight validated MTOsport simulation model presented above.

5.1 Vertical CG location above the PTL

The vertical location of the centre of gravity (CG) with respect to the propeller thrust line (PTL) has been subject to long term debates. The UK Civil Aviation Authority (CAA) undertook a series of flight tests to assess the handling qualities of certain single seat gyroplane types. As part of this investigation, the vertical displacement between PTL and CG was determined for each machine. Based on these investigations the British Civil Airworthiness Requirements (BCAR Section T) was updated and the following statement can be found under Sub-Section B (AMC T 23 Interpretative Material) [21]:

"…For gyroplanes that have a thrust line/centre of gravity limit of offset greater than ±2 inches, normal distance from the propeller thrust line, the CAA shall be consulted to establish design/equivalent safety standards that would ensure safety objectives continue to be achieved.

Tab. 1 indicates that the MTOsport's CG is located about 27 cm (10.6 inches) below its PTL. The flight experience with this popular gyroplane type (more than 500 MTOsport gyroplanes flying all over the world) shows definitely no problems due to this offset.

The analysis of the MTOsport eigenvalues based on the simulation model presented above prove this experience from practice, Fig. 8.

Within the airspeed range of 12 to 40 m/s the short period mode is damped well and the phugoid mode is stable as well except for very low airspeeds of 12 m/s. The eigenfrequency of the phugoid mode is about 0.2 rad/s which is comparable to fixed wing airplanes. For rotorcraft and fixed wing aircraft military specifications exist [3, 24]. In this paper the fixed wing specifications are considered. From [3] the following requirements for phugoid are known:

- Level 1: damping ratio above 0.04.
- Level 2: damping ratio above 0.
- Level 3: time to double above 55 s.

At higher airspeeds Level 1 handling qualities are obtained with respect to phugoid damping [22]. At 12 m/s true airspeed Level 3 handling qualities are not fulfilled when compared to the fixed wing specification; however considering the rotorcraft specification Level 1 is fulfilled also in this case.

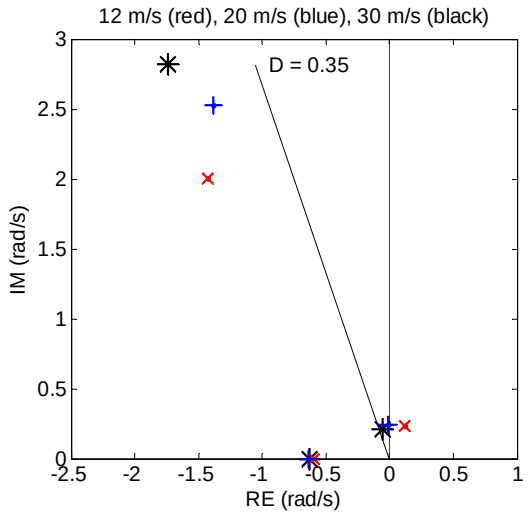


Fig. 8 MTOsport eigenvalues longitudinal motion at 12 (red x), 20 (blue cross), and 30 (black star) m/s true airspeed (m=390 kg).

The short period damping requirements for category A&C flight phases according to [3] are

- Level 1: above 0.35.
- Level 2: above 0.25.
- Level 3: above 0.15.

The MTOsport has Level 1 handling qualities with respect to short period damping for all airspeeds investigated, which means it is satisfactory without any improvements.

The physical background for proposing the PTL below the CG is as follows: consider only three forces acting on the gyroplane 1) the rotor thrust force, 2) the propeller thrust force and 3) the weight force; a PTL below the CG leads to a rotor thrust line located behind the CG. In this case an external disturbance, which increases the rotor thrust force, therefore generating a negative pitching moment and reducing the angle of attack and hence the rotor thrust force; this represents a stable aircraft behavior as desired and can be view as a kind of "power augmented stability".

A gyroplane with its CG behind the rotor thrust line may behave like a tail heavy fixed wing aircraft, which is unstable. In the case of the MTOsport no unstable mode is found (except at very low airspeed). There are two main reasons: first the tail plane produces a downward oriented force resulting in a pitch up moment such that the rotor thrust line almost runs through the CG, Fig. 9, second the design of the rotor head has significant influence on flight stability.

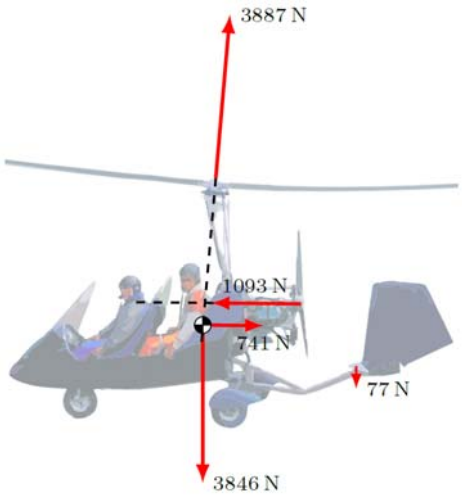


Fig. 9 MTOsport simulation of steady state flight at 30 m/s true airspeed: the rotor thrust line is running through the CG (m=390 kg).

The MTOsport rotor head is designed such that the rotor shaft is moved some centimeters behind the pitch pivot bolt (PPB) around which the rotor is tilting, Fig. 5. This design causes a pitch down moment if the rotor force is increased due to an external disturbance or a pilot control activity, hence stabilizes the rotor itself. This effect is considered in the DLR model presented above and has been proved as a very sensitive configurationally parameter (l_{TB_PPB}) with respect to flight stability.

5.2 Mix of Helicopter and Fixed Wing Aircraft Modes?

The longitudinal modes (short period and phugoid) of the MTOsport are very similar to those of an aircraft, Fig. 8. This is also true for the lateral/directional modes with roll, spiral and dutch roll eigenvalues. The piloting technique required to fly this type of gyroplane is in most situations very similar to those of an airplane except ground handling, takeoff, touch down, and vertical autorotation.

The significant difference to airplane dynamics is rotor rotational speed mode, an eigenvalue at approximately -0.7 to -0.6 rad/s. The relatively high frequency flapping mode is not relevant for the flight dynamics. In order to gain a better understanding of this mode the transfer function of pitch rate by pitch stick force is explored at cruising speed of 30 m/s, Fig. 10.

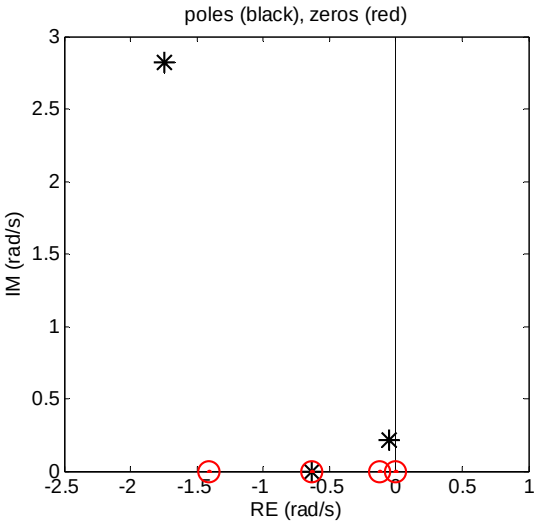


Fig. 10 MTOsport transfer function pitch rate by pitch stick force at 30 m/s true airspeed (m=390 kg).

Four zeros are located on the real axis, two near the origin and one at around 1.4 rad/s and one near the rotor rotational speed pole. The first three zeros also exist for the fixed wing airplane pitch rate transfer function. The additional zero near the rotor rotational speed pole is unique for the gyroplane. Due to the nearby location the pole and zero cancel each other in first approximation and the same pole/zero combination as for airplanes emerge. If this zero would not exist the pitch response of the gyroplane would be very sluggish, since the rather low frequency pole adds a lot of phase and amplitude decrease to the system response.

5.3 Tail Plane Effect on Longitudinal Flight Dynamics

The tail plane adds pitch damping to the gyroplane and therefore it contributes significantly to improve the damping ratio of the short period mode. In case of the MTOsport the short period damping ratio is about 0.4 for an air speed of 30 m/s. This represents Level 1 handling qualities as stated above. Removing the tail plane from the simulation model allows an estimation of the pitch dynamics of an MTOsport without a tail plane. Fig. 11 shows nonlinear simulations with and without the tail plane for an air speed of 30 m/s.

The stronger pitch rate oscillations with tail plane removed indicate that the short period damping ratio is reduced significantly. The short period damping ratio of an MTOsport without a tail plane would only be about 0.2 which leads to poor handling qualities (Level 3) and the likelihood of PIO may increase.

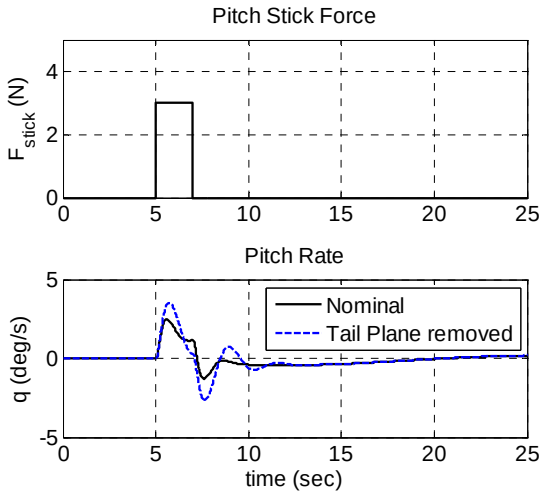


Fig. 11 Effect of tail plane – MTOsport nonlinear simulation at 30 m/s true airspeed (m=390 kg).

5.4 PIO due to phugoid motion?

The analyses presented herein show that the MTOsport is definitely not PIO prone, which fits the experience from real flight. No PIO accident or incident is known to the authors with MT03/MTOsport gyroplane types.

In [6] it is stated that gyroplane PIO is more likely to occur and can be more pronounced at higher airspeeds. This may be due to insufficient short period damping resulting from too small tail planes and increased pitch control sensitivity at higher airspeed. In case of the MTOsport model an increasing phugoid damping is found with increasing airspeed (see Fig. 8).

Further investigations need to be done in this area.

5.5 Backside of the Power Curve

Operating a gyroplane on the backside of the power curve can be hazardous at low altitudes, since no power reserve is available when the gyroplane starts to decent due to an external disturbance. In [3] the flight path stability is defined by means of the gradient $d\gamma / dV$ which is the steady state flight path angle change due to airspeed at a constant throttle position. The following levels of handling qualities are specified:

- Level 1: $d\gamma / dV < 0.06$ deg/kt
- Level 2: $d\gamma / dV < 0.15$ deg/kt
- Level 3: $d\gamma / dV < 0.24$ deg/kt

Fig. 12 shows the total x-force and the maximum thrust available over the entire airspeed range.

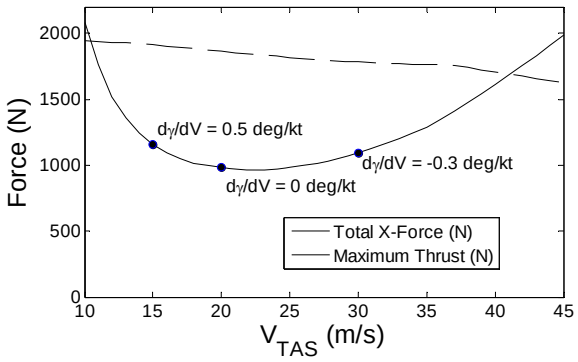


Fig. 12 MTOsport flight path angle stability; total force in x-direction and maximum propeller thrust available over the airspeed range (m=390 kg).

The total x-force is the sum of external forces in body fixed x-direction containing of the body drag, the rotor drag and the effective gravity force. The total x-force is minimal between 20 and 25 m/s true airspeed. Below this value the gyroplane is flying on the backside of the power curve. In this region the flight path angle gradient is increasing rapidly with reduced airspeed and at 15 m/s Level 3 is not fulfilled anymore with respect to the flight path stability criterion. This finding based on the simulation model is well in agreement with the experience from real flight such that a gyroplane is not to be operated at very low airspeeds near the ground.

6 Conclusion

Due to the increasing interest in gyroplanes, a research program at DLR's Institute of Flight Systems was initiated, aiming at the further improvement of gyroplane flight safety. A new flight simulation model for AutoGyro's MTOsport has been developed and validated with flight test data. This simulation model is used for a new gyroplane flight training device which is currently under development.

The following conclusions can be drawn based on the analyses for the MTOsport presented herein:

- Based on the flight test data analysed, the flight dynamics of the MTOsport resemble the one of a small airplane; the same characteristic modes in longitudinal motion – the short period and phugoid, and even the same zero location of the pitch rate to stick force transfer function. In the lateral motion the same is valid, but not presented in this paper.
- The often discussed vertical offset between propeller thrust line (PTL) and centre of gravity (CG), which is part of the British Civil Airworthiness Requirements (BCAR Section T) is not the most relevant configurationally parameter. The requirement to limit the offset between CG and PTL to ±2 inches can not be confirmed in case of the MTOsport. The effects of the tail plane and the specific rotor head design must be taken into account, too.
- The damping effects of the tail plane are significant in case of the MTOsport. The short period handling qualities would be degraded from Level 1 to Level 3 if the tail plane would be removed. Too low short period damping has been most likely the cause for pilot involved oscillations (PIO) with gyroplanes in the past.

References

[1] Leishman J. G., "Development of the Autogiro: A Technical Perspective," *Journal of Aircraft*, Vol. 41, No. 4, 2004, pp. 765-781.

[2] Mc Cormick B. W., "A Numerical Analysis of Autogyro Performance," Biennial International Powered Lift Conference and Exhibit, Virginia, 2002.

[3] Anon., "Flying Qualities of Piloted Aircraft," US Department of Defence, MIL-HDBK-1797, 1997.

[4] Anon., "Aviation Safety and Pilot Control, Understanding and Preventing Unfavourable Pilot-Vehicle Interactions," 1997.

[5] Duda H., "Prediction of Pilot-in-the-Loop Oscillations due to Rate Saturation," *Journal of Guidance, Navigation, and Control*, Vol. 20, No. 3, May-June 1997.

[6] Anon., "Rotorcraft Flying Handbook," U.S. Department of Transportation, Federal Aviation Administration, Flight Standards Service.

[7] Anon., "Investigation Report BFU 3X121-0/09," Bundesstelle für Flugunfalluntersuchung (BFU), Braunschweig, March 2010

[8] Oertel H., Slanina P. and Zach A., "Data Acquisition System" (in German), DLR Institute of Flight Systems, IB 111-2011/11, Braunschweig, 2011.

[9] Zach A., "Evaluation Results of Calibrations to the Gyroplane Measuring System" (in German), DLR Institute of Flight System, IB 111-2010/10, Braunschweig, 2010.

[10] Pruter I. and Duda H., "Flight Tests for System Identification of a MTOsport in the Project GyroTrain" (in German), DLR Institute of Flight Systems, IB 111-2010/57, Braunschweig, 2010.

[11] Schaefer R. and Smith H. A., "Aerodynamic Characteristics of the NACA 8-H-12 Airfoil Section at Six Reynold Numbers from 1.8 x 10^6 to 11 x $10^{6,}$" Langley Aeronautical Laboratory, National Advisory Committee for Aeronautics, Technical Note NACA TN 1998, Langley Air Force Base, Va., 1998.

[12] Arnold U., "Investigation of Gyroplane Flight Mechanics" (in German), Master Theses, Institute of Flight Systems, Braunschweig, 1988.

[13] Pruter I. and Duda H., "Modeling und Simulation of Gyroplane Flight Dynamics" (in German), Deutsche Gesellschaft für Luft- und Raumfahrt (DGLR), Aachen, 2009.

[14] Pruter I. and Duda H., ”A new Flight Training Device for Modern Lightweight Gyroplanes,” American Institute of Aeronautics and Astronautics, AIAA – Modelling and Simulation Technologies Conference, Portland 2011.
[15] Mößner M., “Development and Validation of an Engine and Thrust Model for the MTOsport Gyroplane” (in German), DLR Institute of Flight Systems, IB 111-2010/46, Braunschweig, 2010.
[16] Houston S. S., “Longitudinal Stability of Gyroplanes,” *Aeronautical Journal*, January 1996.
[17] Houston, S. S., “Identification of Autogiro Longitudinal Stability and Control Characteristics,” *Journal of Guidance, Control and Dynamics*, Vol. 21, No. 3, May-June 1998, pp. 391-399.
[18] Houston, S. S., “Identification of gyroplane lateral/directional stability and control characteristics from flight test,” *Proc Instn mech Engrs*, Vol. 212, Part G, 1998, pp. 271-285.
[19] Thomson D. and Houston S. S., “Advances in the Understanding of Autogyro Flight Dynamics,” American Helicopter Forum, Montreal, Canada, 2008.
[20] Jump M., Kendrick S., Perfect P., Padfield G. D. and Ridgway G. R., “Rapid Prototyping and Validation of an Autogyro Simulation Model,” CEAS European Air and Space Conference, Manchester, 2009.
[21] Anon., “British Civil Airworthiness Requirements, Section T Light Gyroplanes,” Issue3, 2005.
[22] Deiler C., “Analysis of Gyroplane Flight Dynamics at the Limits of the Flight Envelope” (in German), DLR Institute of Flight Systems, IB 111-2011/12, Braunschweig, 2011.
[23] Lopez C. A. and Wells V. L., “Dynamics and Stability of an Autorotating Rotor/Wing Unmanned Aircraft,” *Journal of Guidance, Control and Dynamics*, Vol. 27, No. 2, March-April 2004, pp. 258-270.
[24] Anon., “Aeronautical Design Standard Performance Specification Handling Qualities Requirements for Military Rotorcraft”. ADS-33E-PRF, 21 March 2000.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

On Some Trim Strategies for Nonlinear Aircraft Flight Dynamics Models with the Open Source Software JSBSim

James M. Goppert and Inseok Hwang
School of Aeronautics and Astronautics, Purdue University
West Lafayette, Indiana, USA

Agostino De Marco
Department of Aerospace Engineering, Università degli Studi di Napoli "Federico II"
Via Claudio, 21 – 80125 Napoli, Italy

Abstract

JSBSim is an open source flight dynamics library written in the C++ programming language. In this paper, we implement a simplex-based trimming algorithm and compare it to the existing interval method used by JSBSim. The test cases include a Boeing 737, an F-16, and a Multiplex Easy Star UAV (unmanned aerial vehicle). The existing trim algorithm requires less computation, but is rudimentary and fails to solve difficult trim problems. We briefly discuss the usage of our added GUI (graphical user interface) for the trimming and linearization algorithms. In addition, we give an overview of the JSBSim block for the open source Scicos block diagram environment.

1 Introduction

JSBSim is a high-fidelity, 6-DoF (Degree-of-Freedom), general purpose flight dynamics model (FDM) software library written in the C++ programming language [1]. The library routines propagate the simulated state of an aircraft given inputs provided via a script or issued from a larger simulation application. The inputs can be processed through arbitrary flight control laws, with the outputs generated being used to control the aircraft. Aircraft control and other systems, engines, etc. are all defined in various files in a codified XML format. The library consists of approximately 70,000 text lines in 185 files. The total lines of the program code are estimated at about 50,000 lines. JSBSim development began in 1997 and has an international team of active developers and user contributors. JSBSim is released presently as an open source software (OSS) under the LGPL license, and can be incorporated into a larger flight simulation architecture (such as FlightGear, or OpenEaagles). It can also be run as a standalone batch application when linked with a stub routine.

Thanks to a large worldwide community of users of this free simulation and modelling tool, many advances have taken place, and a variety of uses have been demonstrated [2, 4]. In this paper, we test and compare two trimming algorithms built on top of the FDM software: one existing and one we built. We also show an example of interfacing JSBSim with ScicosLab, a free open source software package providing a multi-platform environment for scientific computation [3].

The goals of this paper are: (*i*) to emphasize the potential of JSBSim as a tool for the analysis of aircraft flight dynamics models; (*ii*) to provide an example of integration of non-commercial tools at the disposal of aerospace engineers, in competition with similar commercial off-the-shelf software; and (*iii*) to emphasize the poten-

tial of JSBSim as a valuable tool in Flight Dynamics teaching, due to the fact that FDMs are easily built in XML format, and models may range from low to high fidelity. In this respect, we will present a GUI in order to give a clear perception of the problem of trimming and introduce the subject of model linearization.

This paper is organized as follows: Section 2 will discuss the trimming of the aircraft dynamics. Section 3 will discuss the linearization of the aircraft model. Section 4 will present the JSBSim ScicosLab block.

2 Aircraft Flight Dynamics Model Trimming

The rigid aircraft equations of motion can be expressed in the following form:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{1}$$

where $\mathbf{x}$ is the aircraft state vector, $\mathbf{u}$ is the input vector, $\dot{(\,)}$ is the time derivative operator, and $\mathbf{f}$ is a vector-valued function. Equation (1) is a typical state equation of a dynamical system, used by control engineers to model a plant. Finding an equilibrium flight condition, or *trim* condition, means finding that particular state $\mathbf{x}_{\text{eq}}$ in combination with a particular set of inputs $\mathbf{u}_{\text{eq}}$ such that, when substituted in the right-hand side of (1), those components of $\dot{\mathbf{x}}$ representing accelerations (linear and angular) are zero.

Trim defines conditions for both design and analysis based on aircraft models; in this sense, trim conditions are considered analysis points. In simulations, these analysis points establish initial conditions comparable to flight conditions. Based on aerodynamic and propulsion system models of an aircraft, trim analysis can be used to provide the data needed to define the operating envelope or the performance characteristics. Linear models are typically derived at trim points. Control systems are designed and evaluated at points defined by trim conditions.

In order to develop a linear model, it is necessary to find an equilibrium point. Most aerial and ground vehicles have many degrees of freedom with a wide range of time constants. Consequently, trimming is often a numerically hard problem. Time constants associated with the propulsion system are typically slow, while time constants associated with the control surfaces are typically fast. We employed the JSBSim library to create a trimming program based on the Nelder-Mead simplex algorithm [6]. When creating this trimming routine for JSBSim, we followed the work of Stevens and Lewis [8], evolving from the work of De Marco *et al.* [4]. The present trimming code can be applied to trimming of any JSBSim modelled aircraft. Since the JSBSim model format is written in XML, it enables easy modelling of any experimental data or new aircraft.

2.1 Design Vector

It is numerically unfeasible to solve for trim conditions without imposing constraints to reduce the dimension of the design vector for the aircraft. In the unconstrained problem, the design vector $\mathbf{d}$ consists of the state and input vectors, $\mathbf{x}$ and $\mathbf{u}$, without the components which are the trim design objectives. Constraint equations are introduced in the next subsection.

To have an idea of how the problem of trimming an aircraft model is posed, let us consider the case of a constant altitude coordinated turn. If we wished to trim an aircraft for a coordinated turn, we would assign V_t to the desired airspeed, ψ to the desired heading, the triplet (latitude, longitude, altitude) to the desired position, and we would set the body roll rate p, body pitch rate q, and body yaw rate r to match the desired yaw rate in the navigation frame (i.e. the turning rate $\Omega = \dot{\psi}$). Moreover, we would make sure the side slip angle β is zero (turn coordination) and the climb angle γ is zero (constant altitude).

The search for a trim condition involves the determination of a sutable functioning state of the propulsive system. We call 'rpm' the revolution per minute rate of the motor and 'prop pitch' the angle of the propeller blade, in case of propeller driven airframe. The latter quantity is an aircraft state when it is variable (typically on larger aircraft).

Equilibrated flight conditions are ensured by a set of appropriate aerosurface regulations and a

suitable throttle setting. These are all aircraft primary controls. We call 'aileron', 'elevator', and 'rudder' the normalized aerosurface control input, that vary from -1 to 1; we call 'throttle' the normalized throttle setting, varying from 0 to 1.

For a propeller driven aircraft the full state vector is:

$$\mathbf{x} = \big[V_t, \alpha, \beta, p, q, r, \ldots \\ \psi, \theta, \phi, \text{altitude, longitude, latitude, } \ldots \\ \text{rpm, prop pitch}\big]^T \quad (2a)$$

For a conventional aircraft architecture, the control vector is the following:

$$\mathbf{u} = \big[\text{throttle, aileron, elevator, rudder}\big]^T \quad (2b)$$

It can be augmented with other aerodynamic controls when a specific configuration is under study, or when particular flight conditions are investigated, e.g. a trim in landing configuration with flaps and spoilers extended.

In general, the design vector is:

$$\mathbf{d} = \left[\widetilde{\mathbf{x}}^T, \widetilde{\mathbf{u}}^T\right]^T \quad (2c)$$

where $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{u}}$ are vectors obtained by taking out from $\mathbf{x}$ and $\mathbf{u}$ some trim design objectives. For instance, in most cases the airspeed V_t is a design quantity, i.e. trim conditions are serched for at given flight speeds. In some cases, when a wings-level flight condition is desired, the roll angle ϕ is known and set to zero. The coordinated turn itself as discussed above, takes out from $\mathbf{x}$ at least six components. In some other cases a trim condition is desired for an assigned value of a control input, e.g. when a failure of one of the aerosurfaces is investigated.

2.2 Constraints

Stevens and Lewis [8] provided the crucial constraint equations and the design vector that are used in the final trim algorithm. The first constraint is the rate of climb constraint.

2.2.1 Rate of Climb Constraint

The rate of climb constraint solves for the pitch angle θ in terms of the flight path angle γ, the angle of attack α, the roll angle ϕ, and the side slip angle β. It is given by the following equation:

$$\tan\theta = \frac{ab + \sin\gamma\sqrt{a^2 - \sin^2\gamma + b^2}}{a^2 - \sin^2\gamma} \quad (3a)$$

where

$$\theta \neq \pm\pi/2 \\ a = \cos\alpha\cos\beta \quad (3b) \\ b = \sin\phi\sin\beta + \cos\phi\sin\alpha\cos\beta$$

2.2.2 Turn Coordination Constraint

The turn coordination constraint solves for the roll angle ϕ such that there is no side slip in a turn.

It can be shown that the following equation holds in the general case of nonzero side slip [8]:

$$\tan\phi = \Gamma\frac{\cos\beta}{\cos\alpha}\times \\ \frac{(a - b^2) + \tan\alpha\sqrt{c(1-b^2) + \Gamma^2\sin^2\beta}}{a^2 - b^2(1 + c\tan^2\alpha)} \quad (4a)$$

where

$$\Gamma = \frac{\dot{\psi}\,V_t}{g} \\ a = 1 - \Gamma\tan\alpha\sin\beta \quad (4b) \\ b = \frac{\sin\gamma}{\cos\beta} \\ c = 1 + \Gamma^2\cos^2\beta$$

Here V_t represents the tangential velocity in the turn, g is the acceleration of gravity, and the constant $\dot{\psi}$ is the yaw rate during the turn.

The turn coordination constraint results from equation (4) by enforcing the condition: $\beta = 0$.

2.2.3 Reduced-Order Design Vector

In our trim algorithm we always assign the airspeed and impose the constraints recalled in Section 2.2. This reduces the dimension of the design vector, which is defined as:

$$\mathbf{d} = \big[\alpha, \beta, \ldots \quad \text{throttle, elevator, aileron, rudder}\big]^T \quad (5)$$

This means that we search for an equilibrium solution over a design space of dimension 6.

2.3 Numerical Trimming Methods

In this section, we will discuss several trimming methods to compare against the rudimentary interval method currently implemented in JSBSim. The first method, the Nelder-Mead simplex algorithm, was implemented for the JSBSim library. The controlled random search (CRS) and the simulated simplex annealing (SSA) methods are discussed for background. Trimming routines provided by ScicosLab are also briefly described.

The Nelder-Mead simplex algorithm was first proposed by J. A. Nelder and R. Mead in 1965 [6]. If x is a generic variable point in the design space, a *simplex* is a finite set of values of x. With an n-dimensional design vector, a simplex of $(n+1)$ points is employed to find the minimum of a scalar function.

In case of the aircraft trimming algorithm, the function to be minimized is a non-negative cost function $J(x)$ that vanishes in equilibrium conditions. It can be defined in different ways, the simplest expression is:

$$J = \dot{V}_t^2 + \dot{\alpha}^2 + \dot{\beta}^2 + \dot{p}^2 + \dot{q}^2 + \dot{r}^2 \quad (6)$$

where the single terms in the sum are given by the first six components of the state function $\mathbf{f}$ in Equation (1).

In Figure 1, the three procedures of the Nelder-Mead technique are graphically shown. The first procedure is *reflection*. This involves reflecting the maximum cost vertex about the face containing the minimum cost vertex. The next procedure is *contraction*. This contracts the maximum cost vertex toward the center of the simplex. The last step is *multi-dimension contraction*, and this contracts all vertices about the minimum cost vertex.

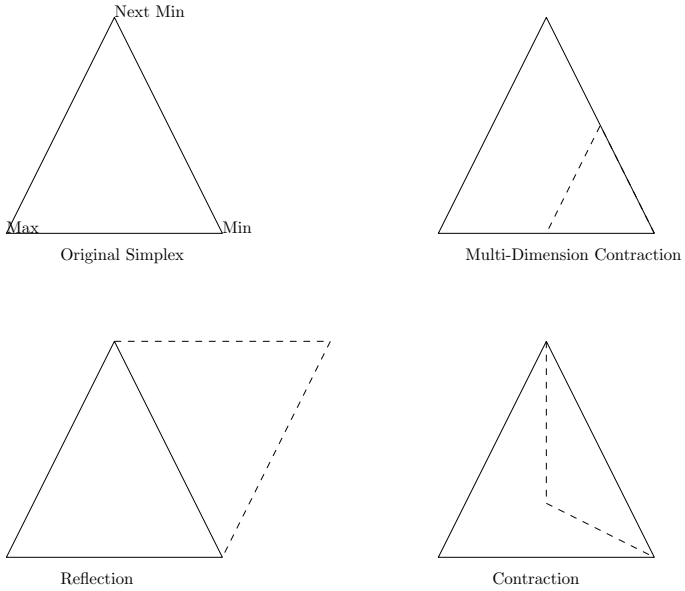


Fig. 1 The three different steps of the Nelder-Mead simplex algorithm.

Another numerical method that is more likely to discover global minima is the controlled random search (CRS) method. This algorithm constructs a simplex from random points in the possible solution domain. It can take longer to converge due to the larger initial population size. When reflection occurs in the CRS method, the maximum cost vertex of the set is replaced by the reflected vertex if the reflected vertex has lower cost [7].

The simplex simulated annealing (SSA) method was proposed by Kvasnicka and Pospichal in 1997 [5]. This method is similar to the CRS method, however there are two distinct differences. First, it randomizes the reflection process by making the reflection distance normally distributed. Second, it uses a criterion to replace the maximum cost point with reflected points with even higher cost. The probability of this event decreases as the scheduled "temperature" variable decreases.

The methods supported by Scilab/ScicosLab, the conjugate gradient method and the quasi-Newton approach, failed to converge for the aircraft test cases. Once the throttle was set, the engine, yaw rate, and pitch rate did not reach a steady state.

Alg. 1 The Trim Method Employed for JSBSim, a probabilistic Nelder-Mead Approach

$S \leftarrow$ set of $(n+1)$ user defined points of D
while $|x_H - x_L| > \varepsilon$ **do**
  $x^\star \leftarrow$ randomized-reflection(S)
  if $f(x^\star) < f(x_L)$ **then**
    $x^{\star\star} \leftarrow$ randomized-expansion(S)
    if $f(x^{\star\star}) < f(x^\star)$ **then**
      $x_H \leftarrow x^{\star\star}$
    else
      $x_H \leftarrow x^\star$
    end if
  else
    $x_{0,H} \leftarrow x_H$
    randomized-contraction(S)
    if $x_H \geq x_{0,H}$ **then**
      randomized-multi-contraction(S)
    end if
  end if
  if iterations $>$ max iterations **then**
    $S \leftarrow$ re-sample population, centered at current minimum vertex
  end if
end while

When relying on ScicosLab's routines to trim an aircraft, it is important to remember that the routines are unaware of aircraft dynamics and therefore do not know how to impose the rate of climb and turn coordination constraints. Therefore, it is strongly suggested to separate the slow time constants from the fast time constants in the diagram. For instance, the motor lag block can be separated from the dynamics so that the lag system can be linearized separately (see Figure 10 and 11). After linearization, the motor lag and aircraft dynamics blocks are combined through series multiplication of state-space systems.

When implementing the standard Nelder-Mead simplex algorithm for the JSBSim library, it was noted that convergence of the simplex can follow a deterministic path that can reduce the likelihood of discovering global minima upon reinitialization of the simplex. To overcome this problem, a probabilistic stepping was introduced to the simplex algorithm. By randomizing the process, similar to the CRS/SSA methods, the aforementioned effects are negated.

The details of the resulting algorithm that has been implemented can be seen in Algorithm 1, where D is the design space and S is the simplex. In this algorithm x_H is the high point or maximum cost vertex. x_L is the low point or low cost vertex. The next lowest point is denoted x_{NL}. The reflection point is $x^\star$. The expansion point is $x^{\star\star}$. The lowest and next lowest points are used to compute the direction for the simplex to reflect.

Alg. 2 The constrained cost function algorithm implemented for JSBSim

$x \leftarrow$ constrain(D)
$\text{cost}_0 \leftarrow f(x)$
propagate(x)
$x \leftarrow$ constrain(D)
$\text{cost}_1 \leftarrow f(x)$
while $|\text{cost}_0 - \text{cost}_1| < \varepsilon$ **do**
  $\text{cost}_0 \leftarrow f(x)$
  propagate(x)
  $x \leftarrow$ constrain(D)
  $\text{cost}_1 \leftarrow f(x)$
end while
return cost_1

2.4 Cost Function

The Nelder-Mead Simplex algorithm can be extended to handle constrained function minimization by setting the cost to a high value if the simplex steps beyond the constrained region [6]. This means that the design space D is subject to a constraint operation at each iteration of the minimization process. For the algorithm employed for the JSBSim library, when a constraint is violated, the constraint is re-imposed and the cost is computed using the constrained design vector. This has the same effect as above, and the simplex will be unlikely to continue stepping over the boundary since the cost will not decrease in that direction.

Since JSBSim models aircraft at a very low level including propulsion system states, simply enforcing the rate of climb constraint and the coordinated turn constraints is not sufficient to uniquely define the state vector. For instance, the internal rpm of the engine/motor is not addressed in constraint Equations (4) and (3) . What we do know about these other states is that they

must reach a steady state value. Therefore, if we repeatedly enforce the constraints and propagate the simulation, an equilibrium solution with the propulsive system in steady state can be found.

The constrained cost function minimization process can be seen in Algorithm 2. Here the cost is recomputed until the difference between the current cost and previous cost is less than a threshold ε.

The developed JSBSim trimming utility can be readily used for trimming various types of aircraft at various states. For example, we successfully used it to trim a nonlinear model of an F-16, both for steady level flight and for coordinated turning flight.

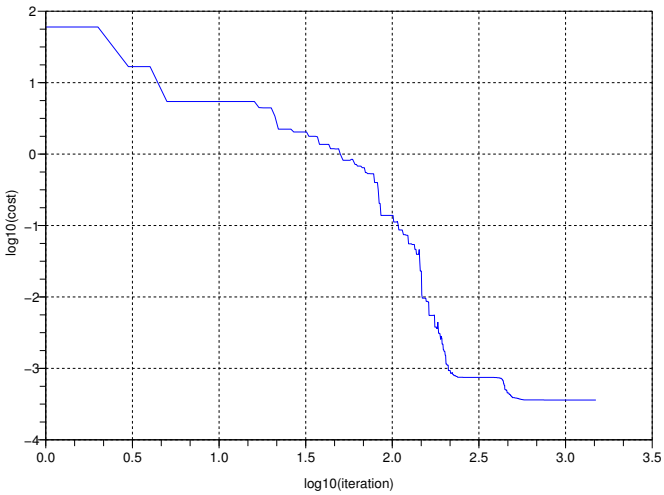


Fig. 2 F16 Simplex Cost Convergence History for 30 deg Bank Turn @ 500 ft/s

2.5 Comparison of Simplex Trimming to Existing Interval Based Method in JSBSim

The interval based trimming method currently implemented in JSBSim assumes that each variable of the trim vector (angle of attack, throttle, elevator, etc.) can independently control a state derivative. For instance, angle of attack is responsible for reducing the altitude derivative. Although the trim vector components are not truly independent, the method is able to find many trim solutions through iterative adjustment of the trim vector until the derivatives approach zero.

The Nelder-Mead simplex algorithm does not make assumptions about the relationship between the variables of the trim vector and the state derivatives of the aircraft. As the simplex converges, the trim vector is adjusted in multiple dimensions simultaneously. A cost function of multiple state variables is used to evaluate how stationary the current trim solution is [see (6)].

The interval based method will attempt to set any derivative to zero, if it is able to, with the associated component of the trim vector. When a trim condition is not possible for an aircraft (i.e. not enough rudder authority to coordinate a turn), it is often desirable to allow some state derivatives to be non-zero in order to reduce other non-zero derivatives. An example of this can be seen in Table 1. In the table, $\dot{u}$ (the derivative of forward velocity), and $\dot{w}$ (the altitude derivative) are both much larger in the interval method than the simplex method. The simplex method, however, takes more time due to its multi-dimensional technique. The simulation of the trim conditions verifies that both methods can easily find the stationary point of the 737 (see Figure 5 and Figure 6).

Table 1 Boeing 737 Trim Comparison, 10 deg Bank Turn, 750 ft/s

	simplex	interval
angle of attack, deg	2.106	2.11
throttle, %	0.7866	0.79
elevator trim, %	−0.1887	−0.19
side slip, deg	0.01505	0
ailerons, %	−0.003156	0
rudder, %	−0.005718	−0.03
$\dot{w}$, ft/s^2	1.09E−06	**0.00041**
$\dot{u}$, ft/s^2	−3.412E−08	**0.00021**
$\dot{q}$, rad/s^2	1.098E−08	−3E−08
$\dot{v}$, ft/s^2	3.227E−07	−8E−08
$\dot{p}$, rad/s^2	−2.373E−09	−1E−07
$\dot{r}$, rad/s^2	−7.037E−09	1.4E−18
computation time, s	**2.4**	0.01

One drawback of the simplex method is that in a state where a stationary trim solution cannot be found, many iterations can occur with little improvement to the overall cost. This situation is illustrated in Figure 2. From the 1000th till the 2000th iteration, the cost decreases minimally. A large number of iterations causes an increase in the computation time as shown in Table 2. Al-

Table 2 F16 Trim Comparison, 30 deg Bank Turn, 500 ft/s

	simplex	interval
angle of attack, deg	7.897	7.19
throttle, %	0.3347	0.34
elevator trim, %	−0.2018	−0.19
side slip, deg	1.096	0.58
ailerons, %	0.02449	−0.01
rudder, %	−1	0
$\dot{w}$, ft/s^2	0.002899	0.00018
$\dot{u}$, ft/s^2	0.0178	8.8E−06
$\dot{q}$, rad/s^2	2.991E−06	7.8E−08
$\dot{v}$, ft/s^2	−0.9425	−4E−06
$\dot{p}$, rad/s^2	2.55E−05	4.1E−08
$\dot{r}$, rad/s^2	−0.0006498	**−0.111**
computation time, s	**13.96**	0.02

Table 3 Easy Star UAV Trim Comparison, No Ailerons, 10 deg Bank Turn, 40 ft/s

	simplex	interval
angle of attack, deg	1.19	2.78
throttle, %	0.514	0.51
elevator trim, %	−0.7226	−0.7
side slip, deg	−0.8595	−0.05
rudder, %	−0.03273	−0.01
$\dot{w}$, ft/s^2	**0.9013**	**1.19**
$\dot{u}$, ft/s^2	−0.07468	**−0.132**
$\dot{q}$, rad/s^2	−0.02834	−2E−05
$\dot{v}$, ft/s^2	0.2737	−3E−05
$\dot{p}$, rad/s^2	−0.002053	**−0.181**
$\dot{r}$, rad/s^2	−4.072E−05	3.6E−08
computation time, s	**5.92**	0.7

though a stationary trim solution was not found by either method, the simplex method was able to trim the aircraft considerably better. The results of the trim simulation are shown in Figure 3 and Figure 4.

The Easy Star UAV provides a unique challenge to the trimming problem. The aircraft has no ailerons. Therefore, side slip is necessary to counteract the torque from the propeller. Although a stationary trim point could not be found, the simplex method outperforms the interval method (as seen in Table 3). We penalized the attitude derivatives heavily in the simplex cost function, so the attitude is more stationary, and the resulting turn behavior is more desirable. The results of the trim simulation are shown in Figure 7 and Figure 8.

3 Dynamics Linearization

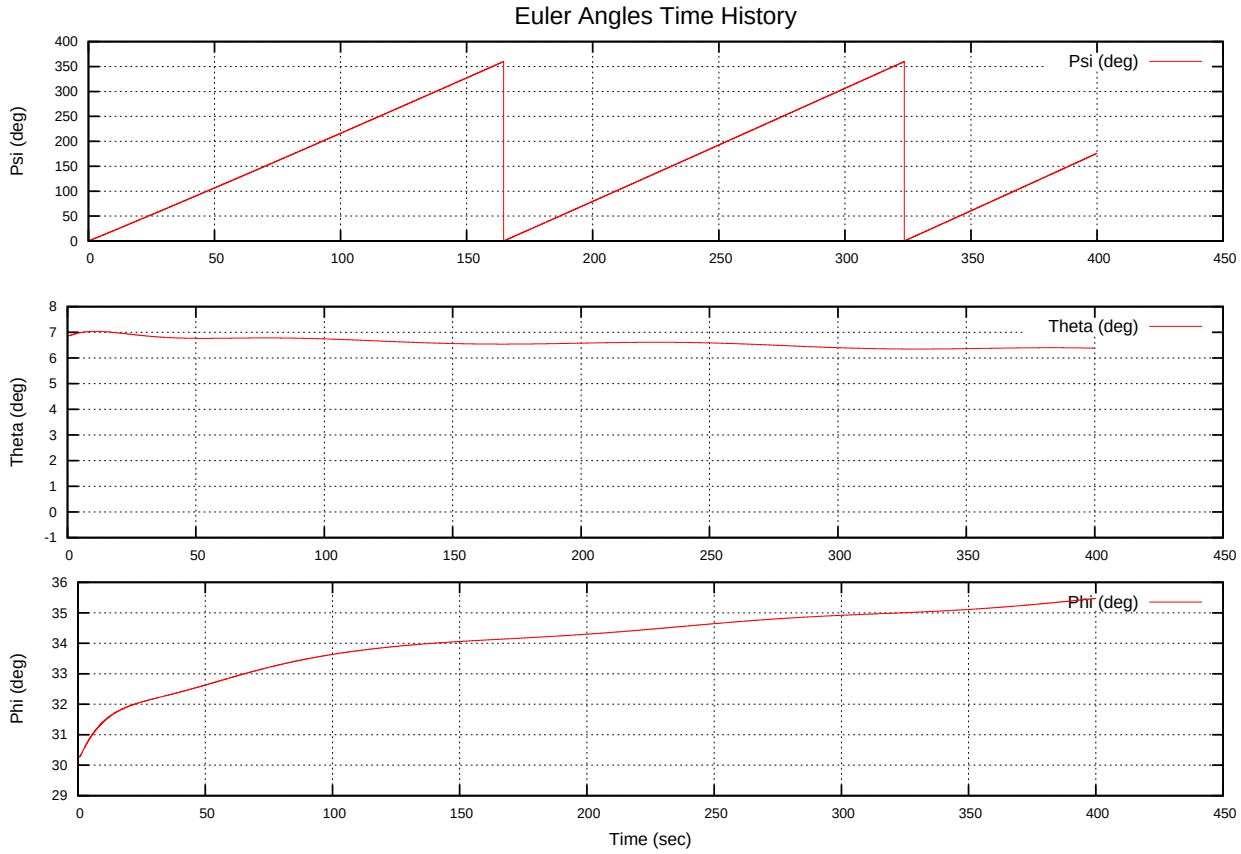
We also have added the ability to compute the state space linearization of the JSBSim dynamics. This functionality has been implemented in a new C++ class named `FGStateSpace`. This class provides a simple interface for selecting components of the input and state vectors and outputting the state and its derivative. If a derivative is provided by JSBSim, that value is used. If a derivative is not calculated in JSBSim, then the default state space component class will implement a finite difference approximation for the derivative.

A multi-platform GUI interface to JSBSim's new linearization and trimming capabilities has been created to help users easily obtain trim conditions and linear models for aircraft in an intuitive manner. This program has been called *Mavsim Trim* and some screenshots are shown in Figure 9.

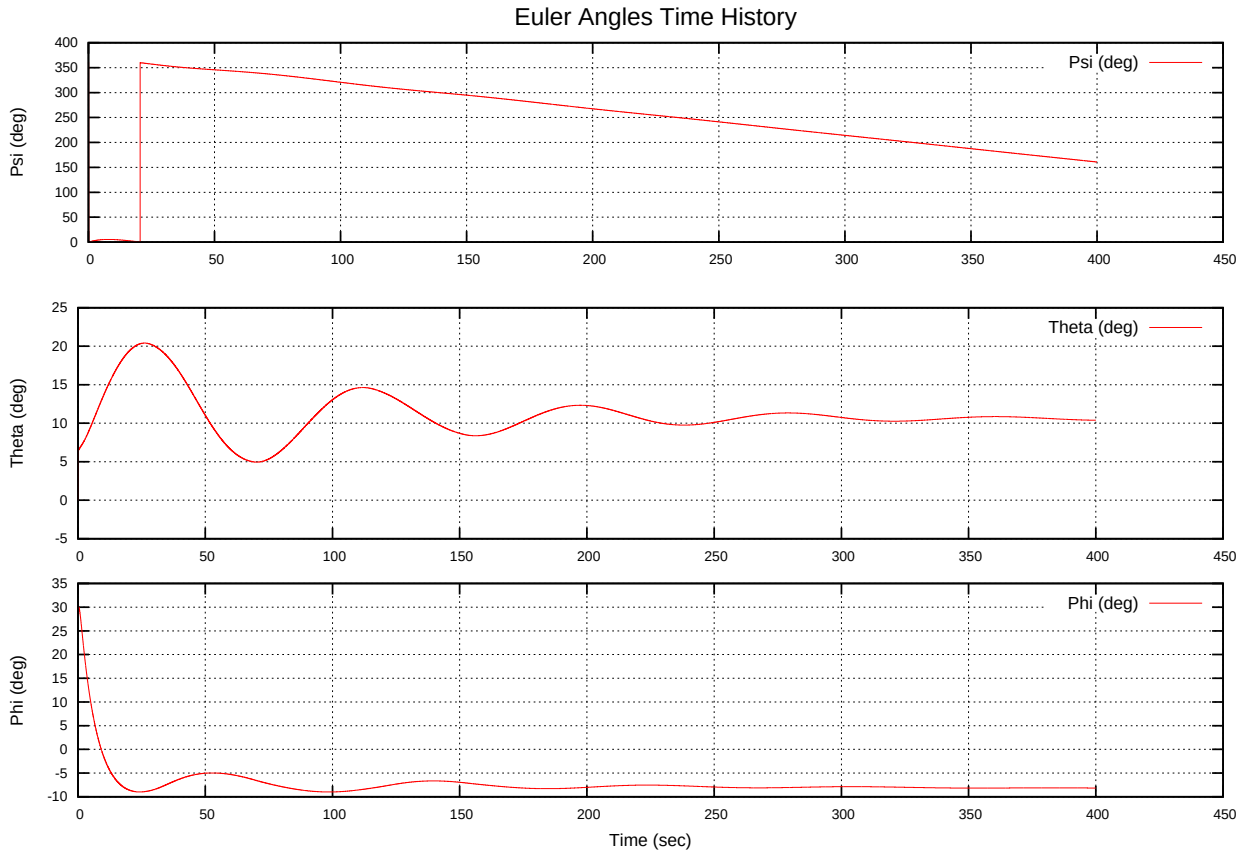
4 JSBSim ScicosLab Block

ScicosLab is an open source scientific computation program. It provides a block diagram environment, Scicos, similar to Simulink from MathWorks. The JSBSim Scicos block enables rapid control design and testing. Hardware-in-the-loop blocks have been developed for use with several current autopilots that employ the MAVSim protocol language. This provides a completely free and open source tool-chain for rapid autopilot development. The JSBSim Scicos block can simulate any JSBSim modeled aircraft and forward the output data to other ScicosLab blocks. The block also receives input. This enables hardware-in-the-loop or software-in-the-loop feedback loops. An autopilot feedback loop for the Easy Star UAV is shown in Figure 10. JSBSim simulates the dynamics of the EasyStar.

The measurable states are forwarded to the control system block within Scicos. The output

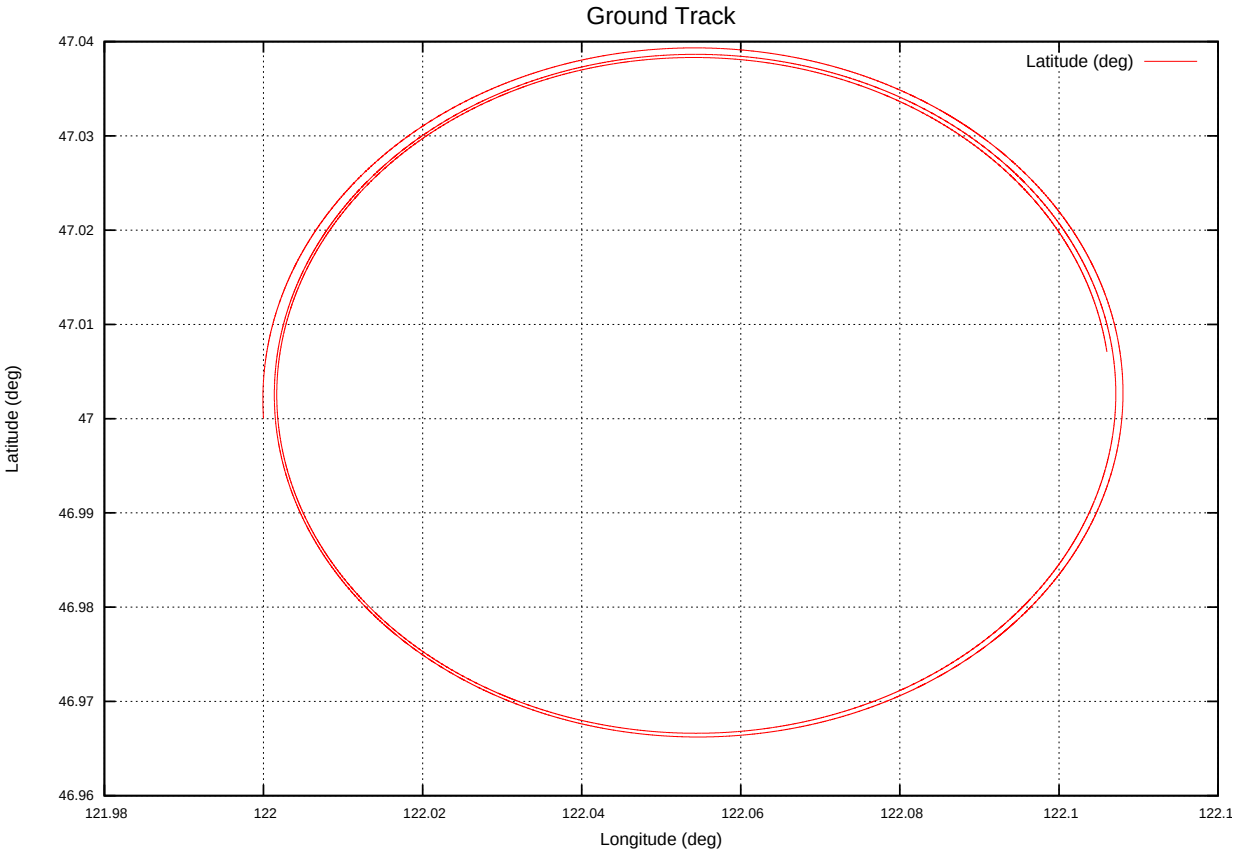


(a) Simplex Method

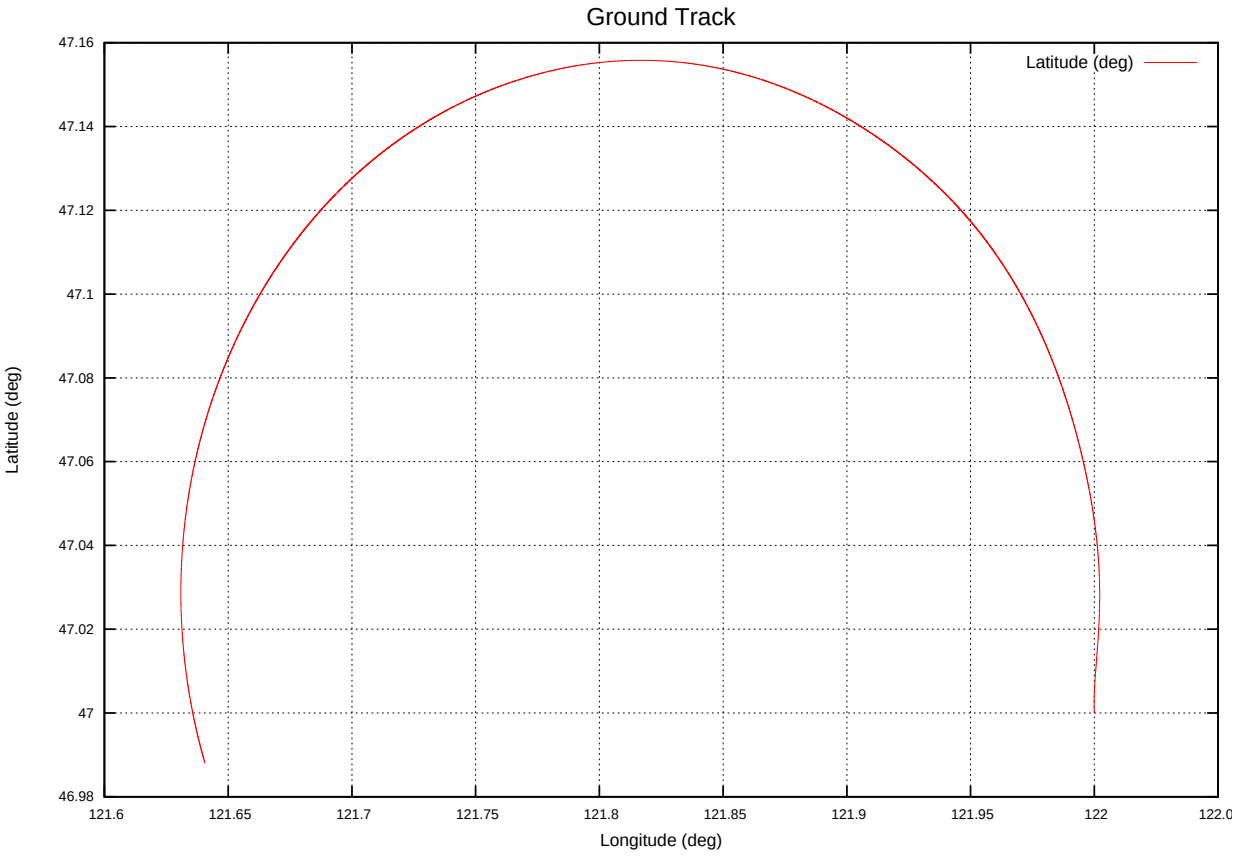


(b) Interval Method

Fig. 3 F16 Attitude History for Trimmed 30 deg Bank Turn @ 500 ft/s

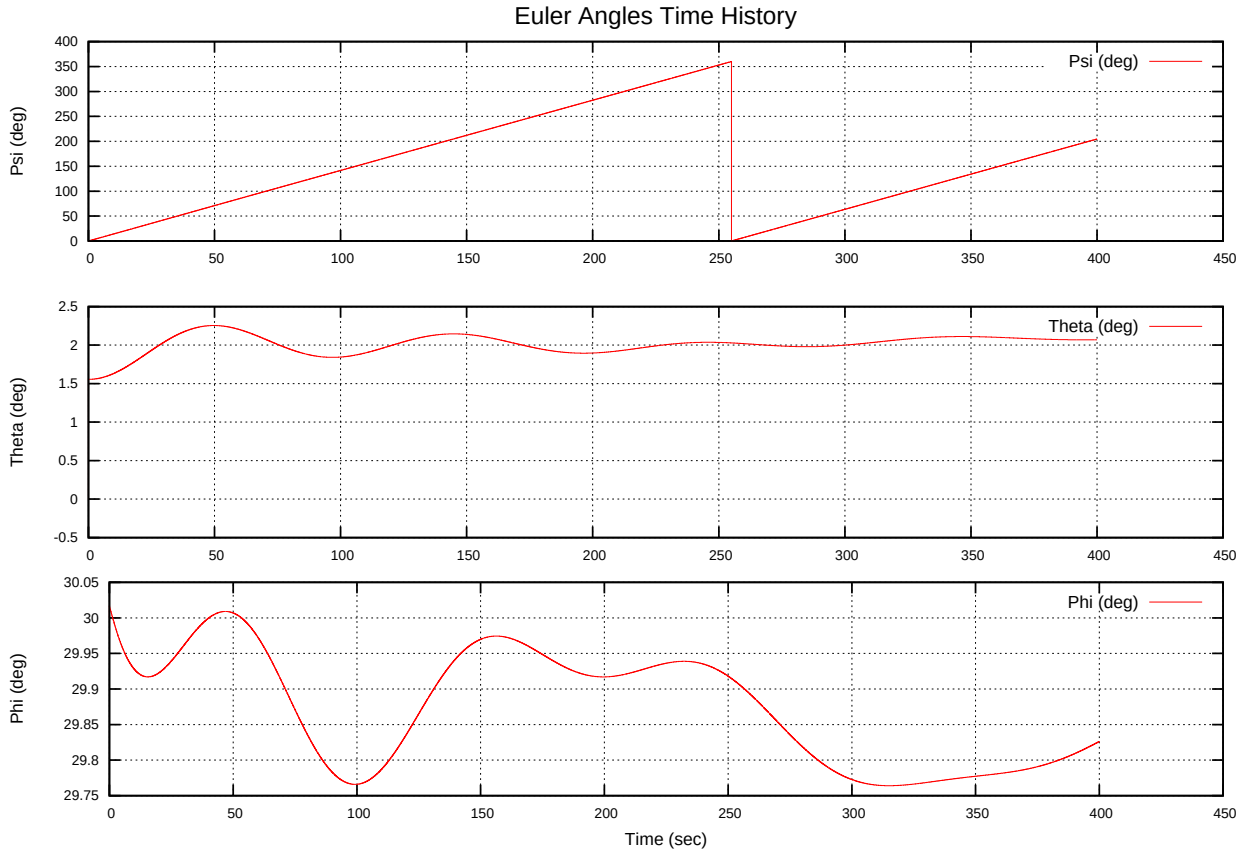


(a) Simplex Method

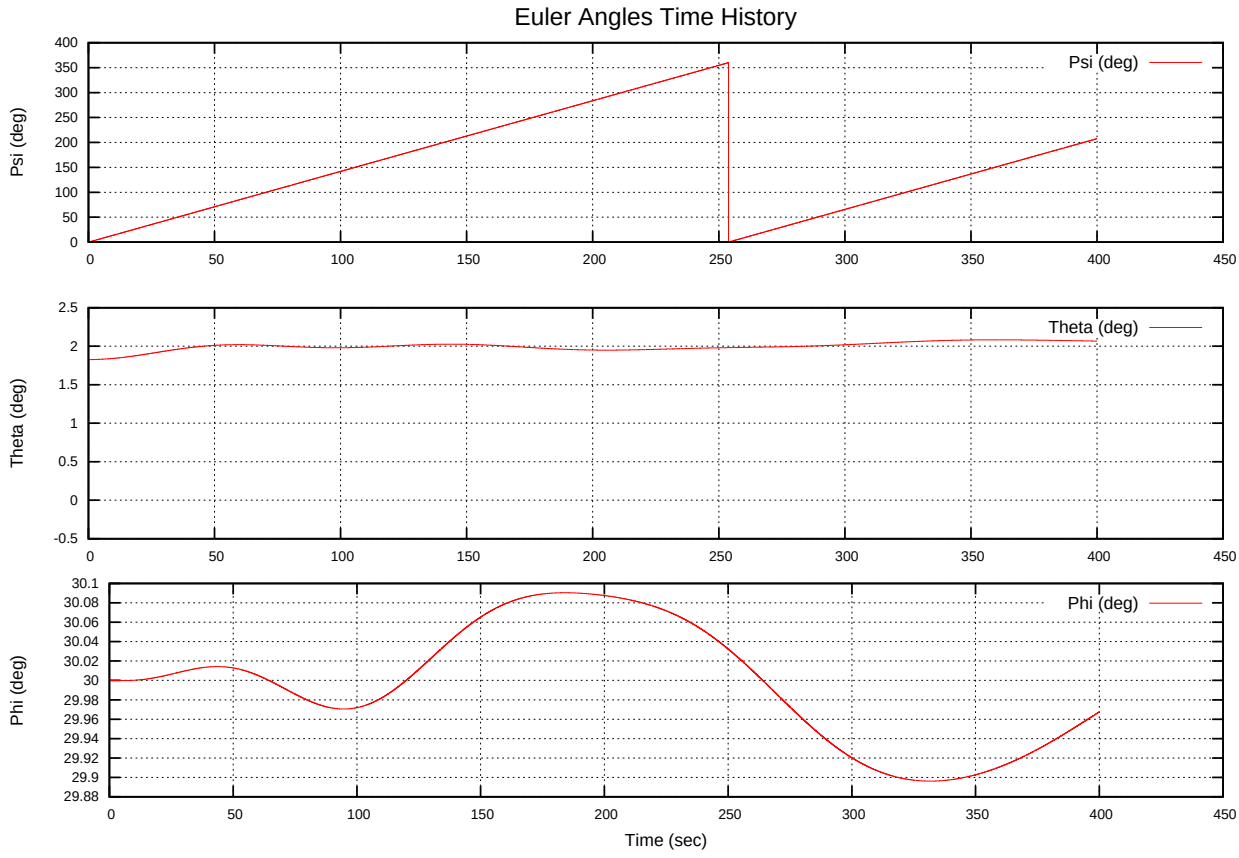


(b) Interval Method

Fig. 4 F16 Position History for Trimmed 30 deg Bank Turn @ 500 ft/s

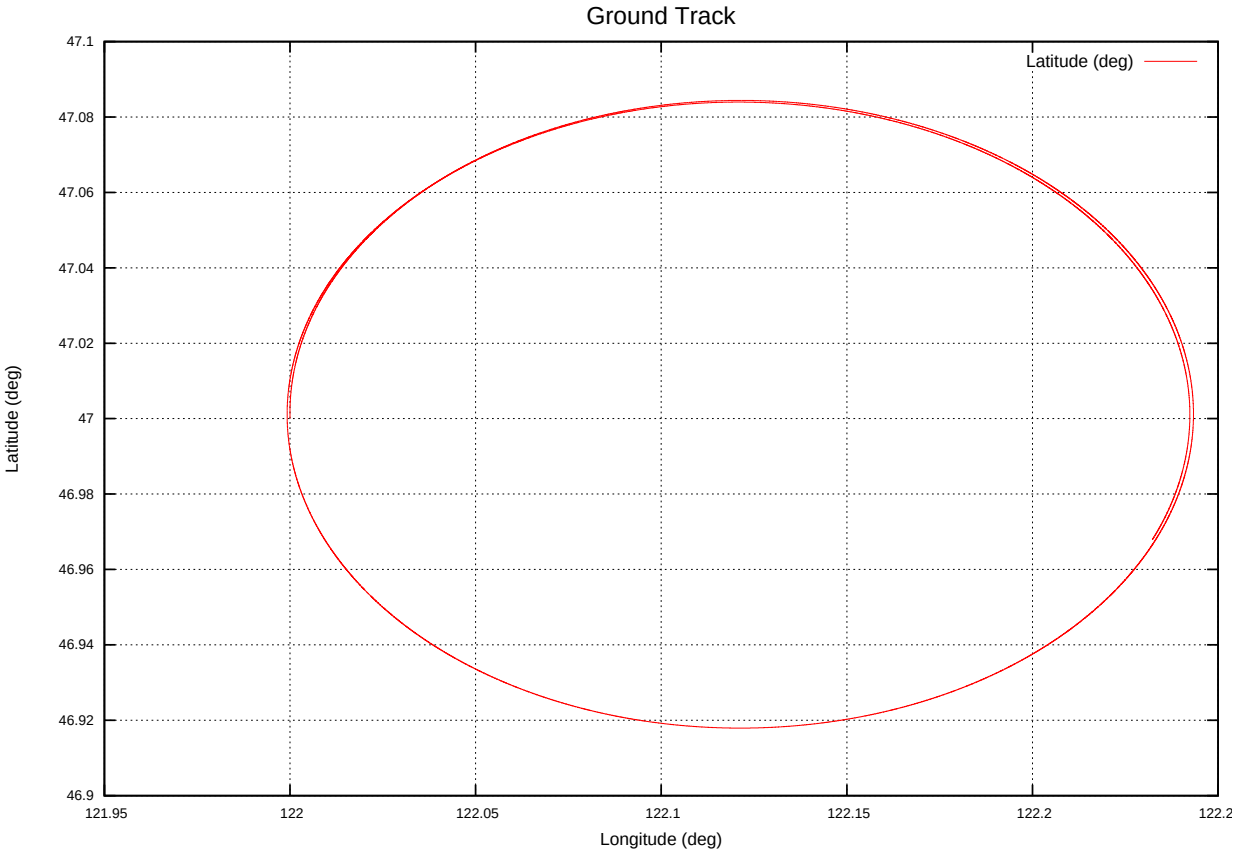


(a) Simplex Method

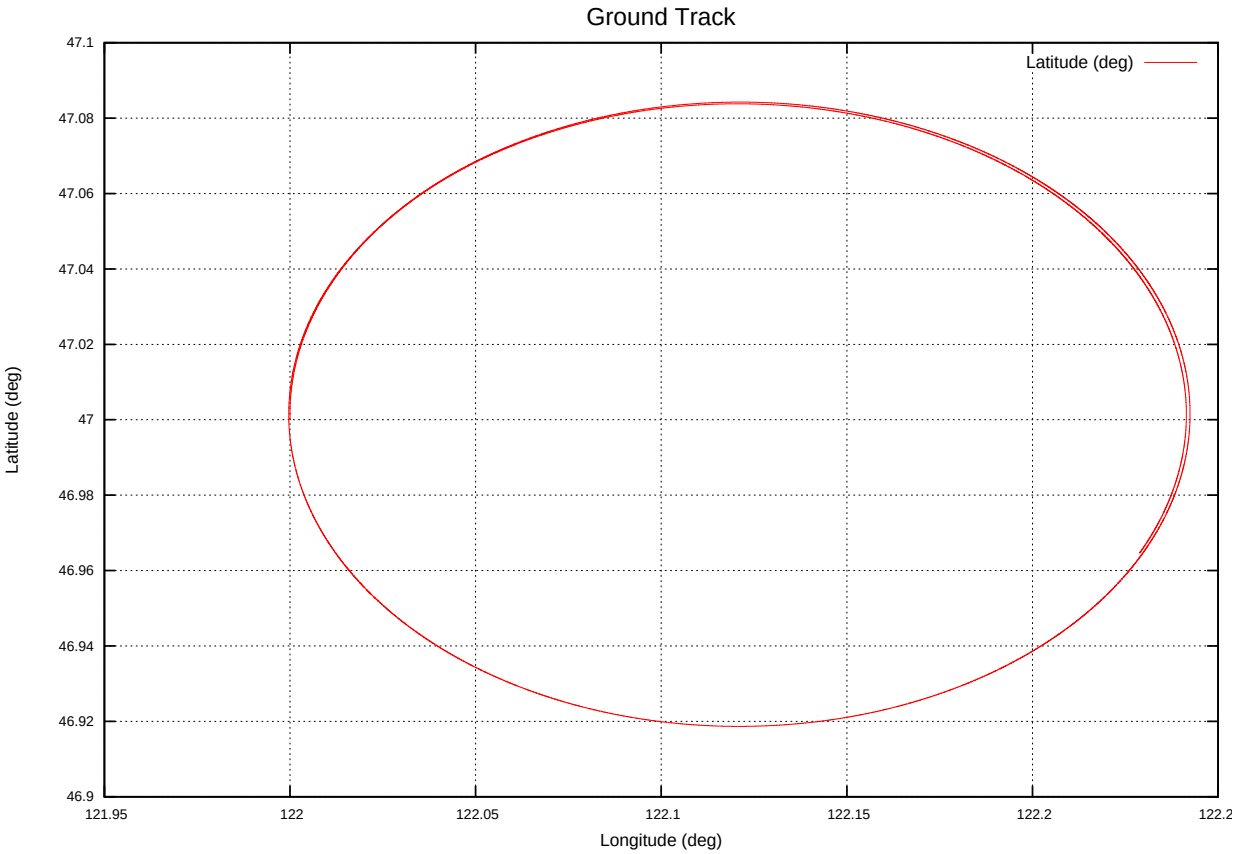


(b) Interval Method

Fig. 5 Boeing 737 Attitude History for Trimmed 10 deg Bank Turn @ 500 ft/s

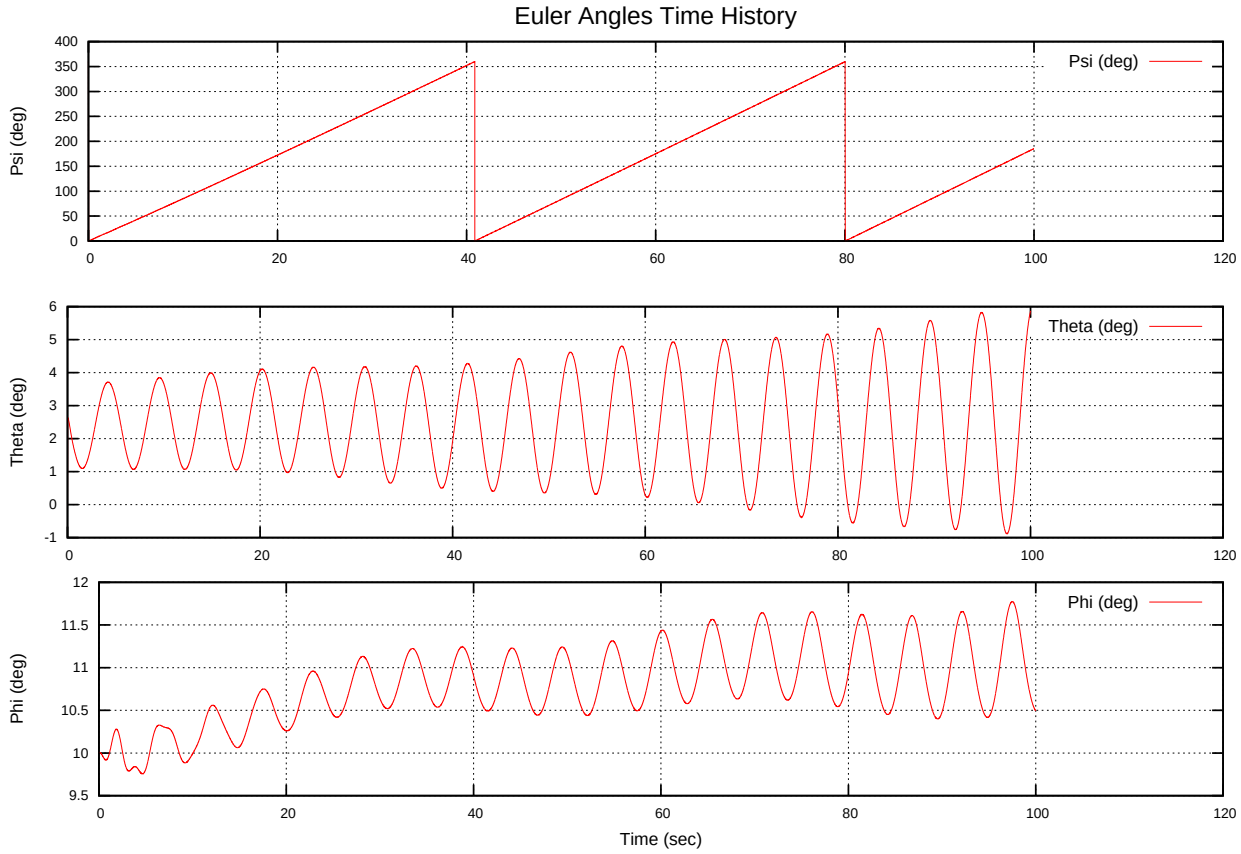


(a) Simplex Method

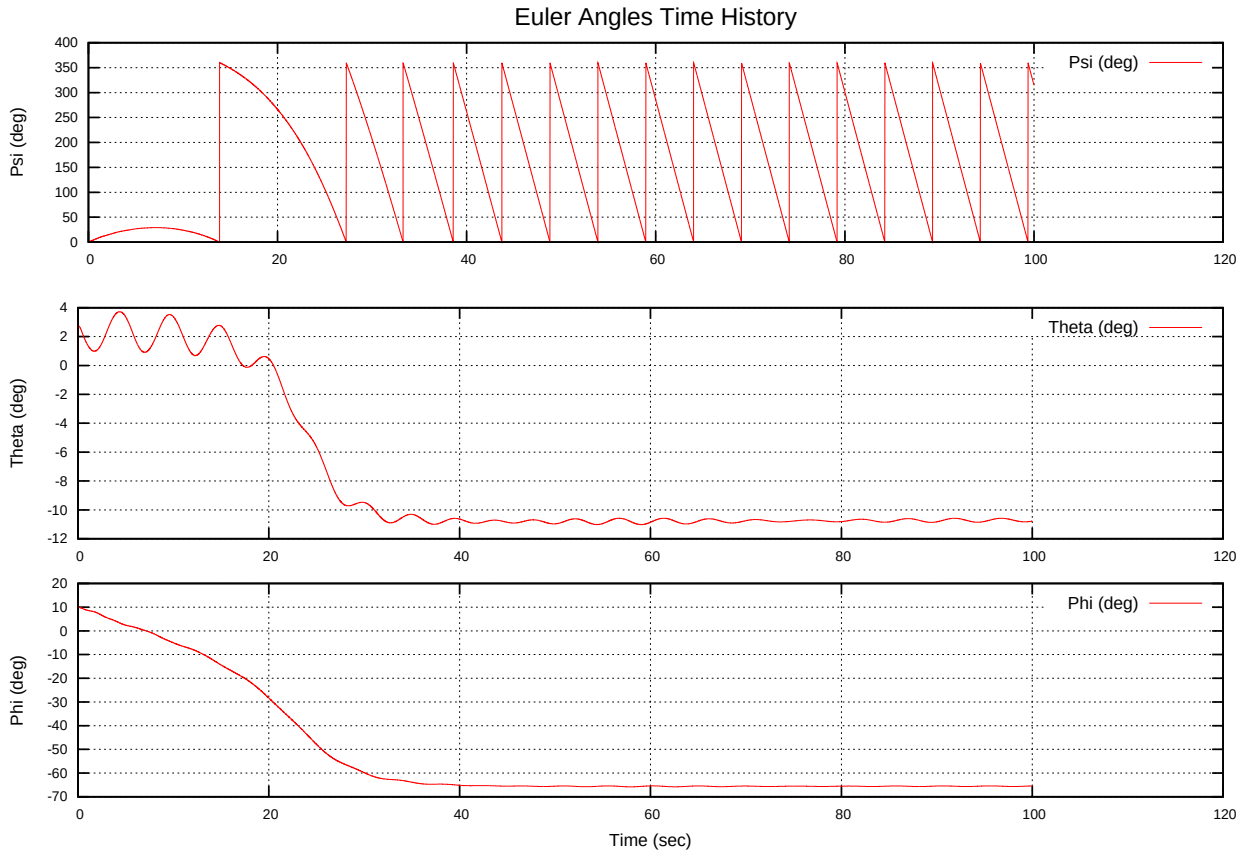


(b) Interval Method

Fig. 6 Boeing 737 Position History for Trimmed 10 deg Bank Turn @ 750 ft/s

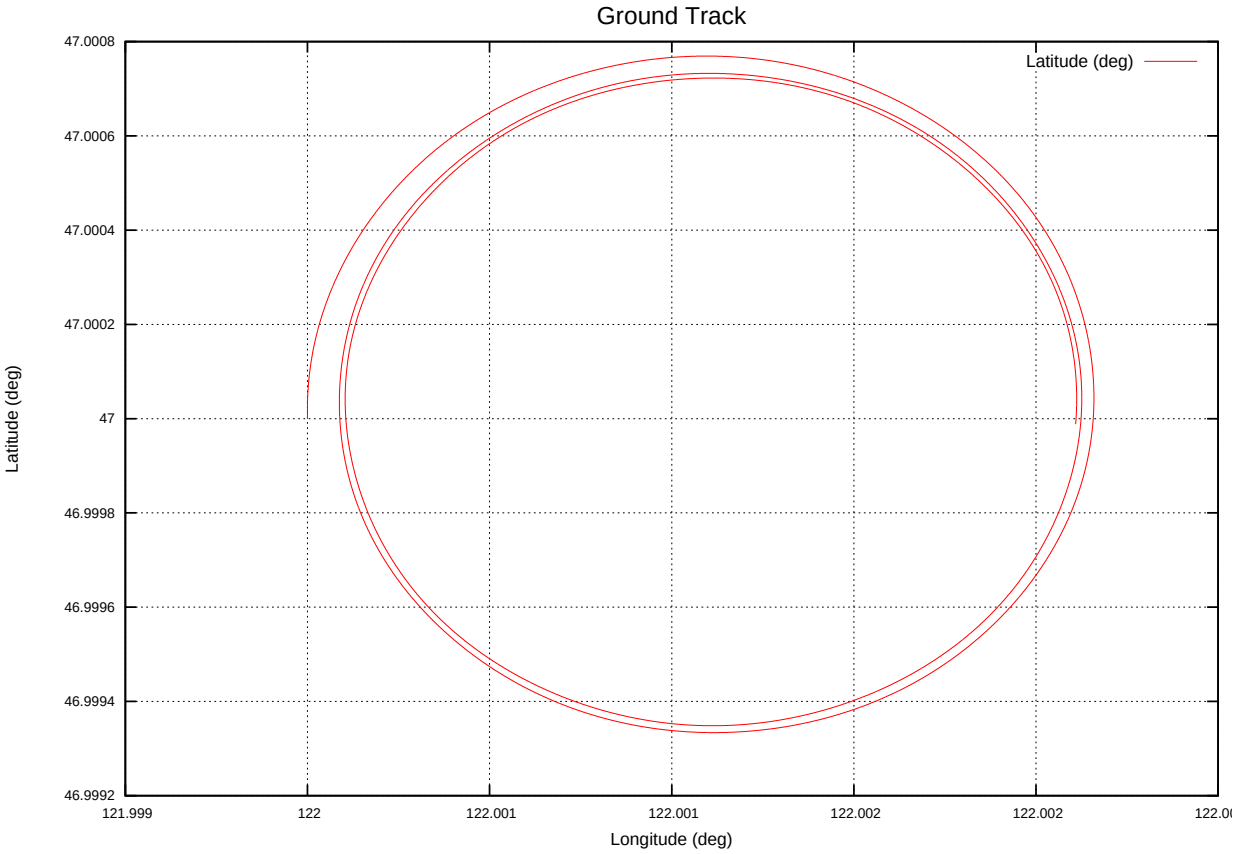


(a) Simplex Method

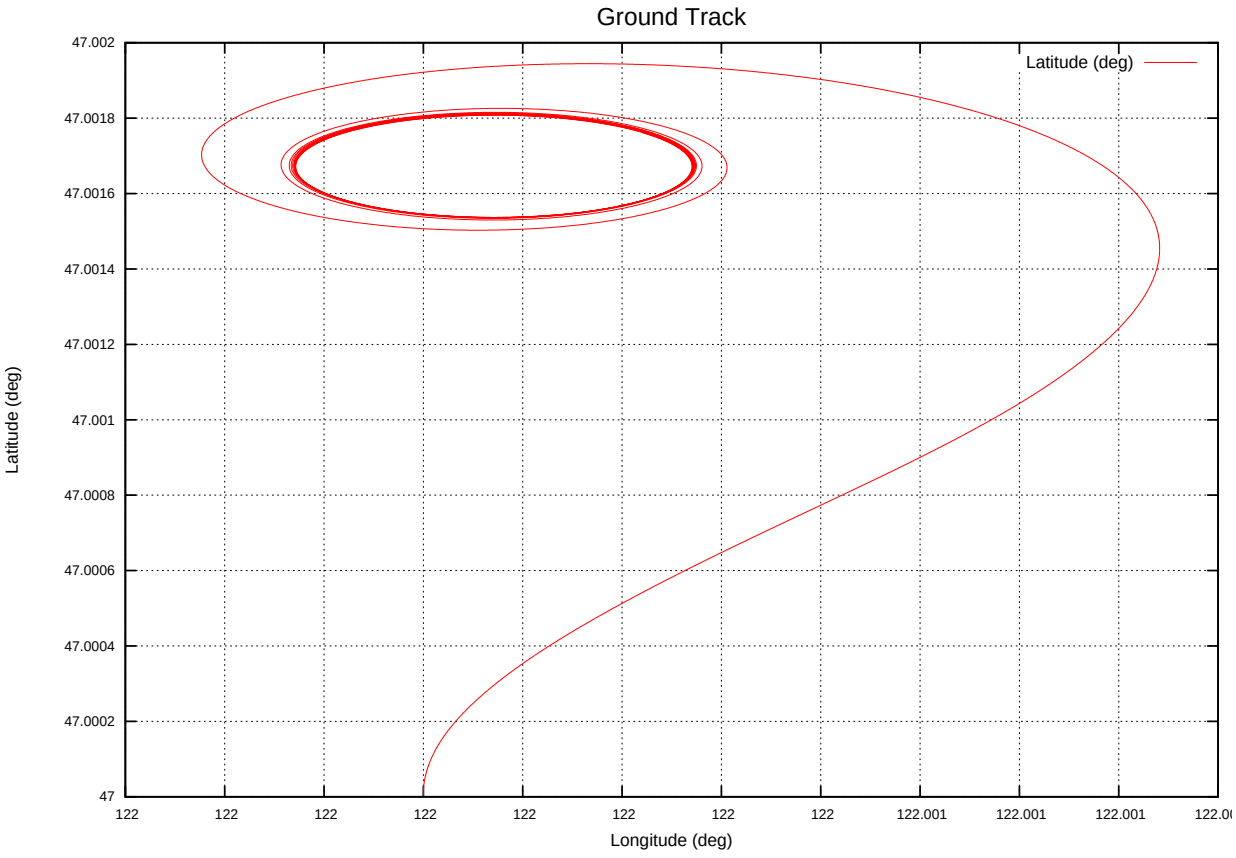


(b) Interval Method

Fig. 7 Easy Star UAV Attitude History for Trimmed 10 deg Bank Turn @ 40 ft/s

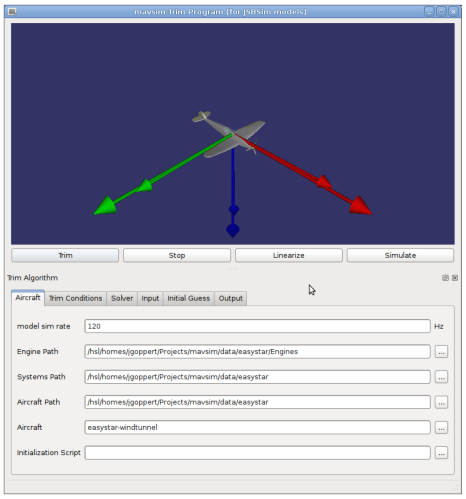


(a) Simplex Method

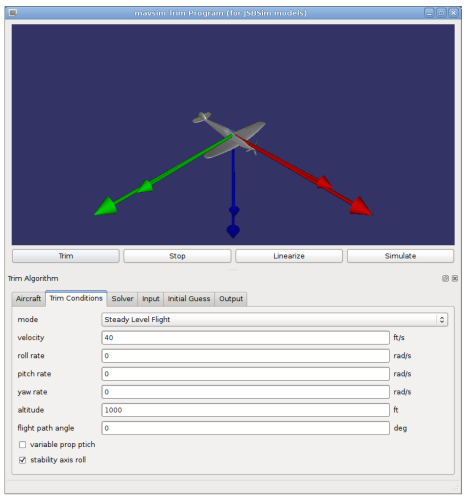


(b) Interval Method

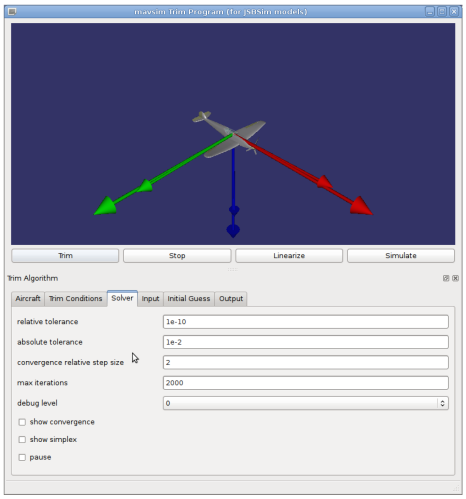
Fig. 8 Easy Star UAV Position History for Trimmed 10 deg Bank Turn @ 40 ft/s



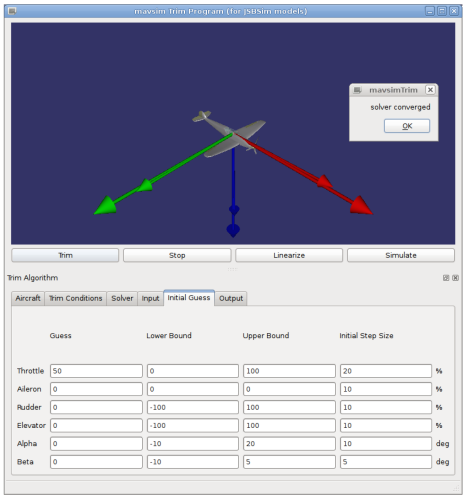
(a) The trim setup used for the Easy Star.



(b) The trim conditions used for the Easy Star.



(c) The solver conditions used to trim the Easy Star.



(d) An example of a converged trim solution in the MavsimTrim GUI.

Fig. 9 Screenshots of the *Mavsim Trim* GUI interface to JSBSim's new trimming capabilities.

from the control system block is fed back to the input of the JSBSim block. If visualiztion is required outside of Scicos, the JSBSim state can be forwarded to the FlightGear flight simulator. The user interface for the JSBSim block is shown in Figure 11.

When dynamics are modelled in ScicosLab, linearization can be done through specific commands. By default, ScicosLab employs a quasi-Newton method for finding the solution. In aircraft trim problems, the Nelder-Mead style, non derivative optimization routine (NDCost) typically is more robust, so a version using this latter method was created. The ability to select the optimization routine will be added to future ScicosLab releases.

When relying on ScicosLab's routines to trim an aircraft, it is important to remember that the routines have no knowledge of the aircraft dynamics and therefore cannot impose the rate of climb and turn coordination constraints to help reduce the dimension of the design vector. Therefore, it is strongly suggested to separate the slow time constants from the fast time constants in the diagram. For instance, in Figure 10, the motor lag block is separated from the aircraft dynamics so that the lag system can be linearized separately. After linearization, the motor lag and aircraft dy-

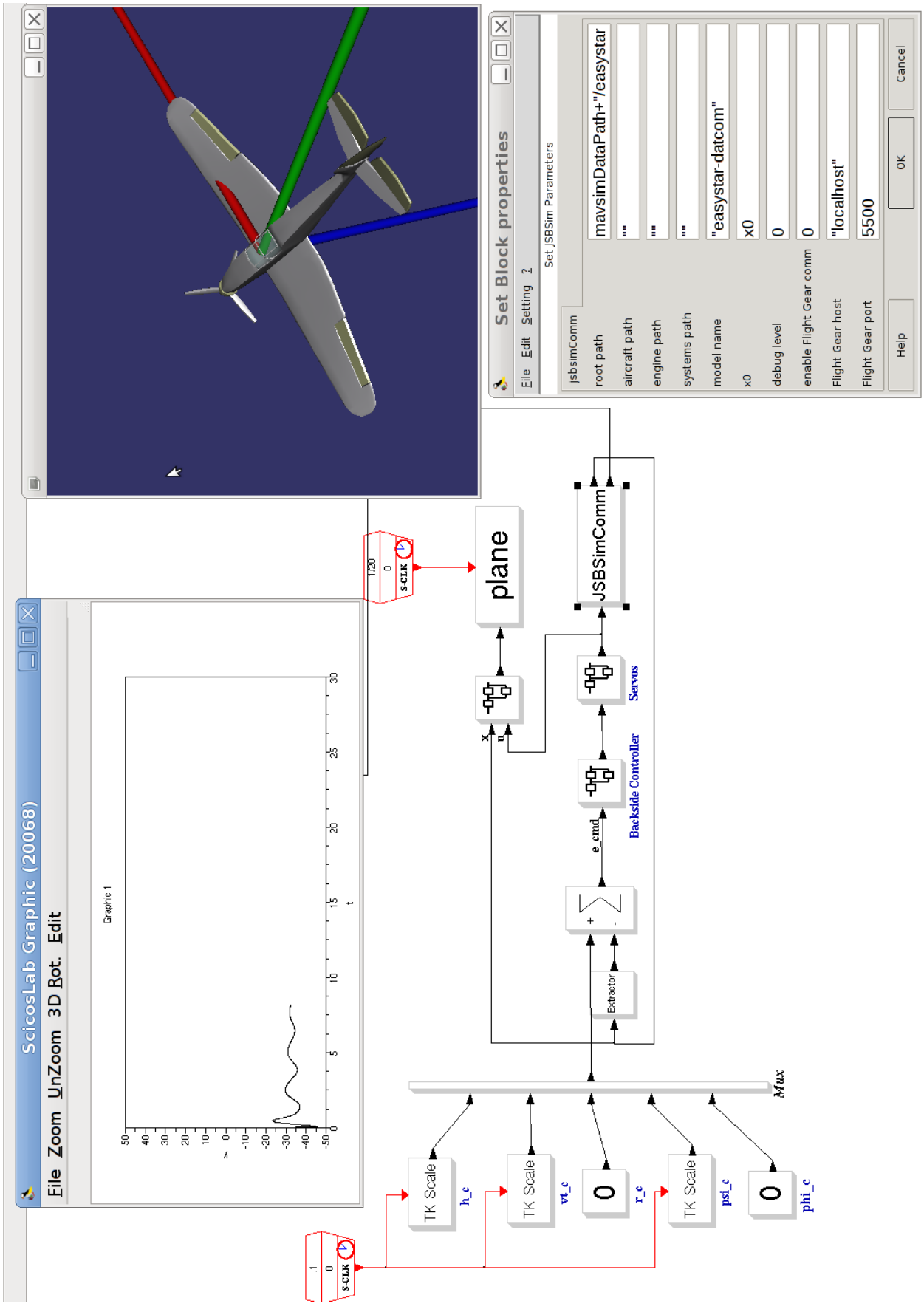


Fig. 10 A ScicosLab Diagram using the JSBSim Communications Block.

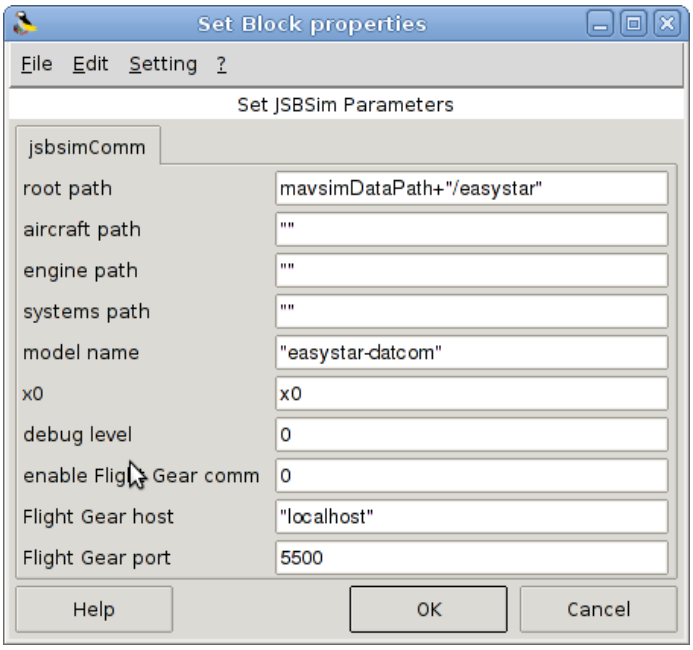


Fig. 11 Setting the properties of the JSBSim Communications Block for ScicosLab.

namics blocks are combined through series multiplication of state space systems.

5 Conclusion

We have implemented a new simplex based trimming method for the JSBSim flight dynamics library. It has been tested against the existing interval based method. This new trimming method is based upon the Nelder-Mead simplex method and randomization was added to the algorithm to avoid a deterministic behavior. At each step, the cost was allowed to converge to a steady state value. The new simplex based trimming method outperformed the interval method but required more computation.

We also created a generic state space interface to the JSBSim library. This interface allows users to create custom state space representations which are useful for control design and dynamics analysis.

Visualization of the trimming and linearization process can be informative and beneficial. Toward this end, we have developed a multiplatform GUI to interface to the trimming and linearization code. This user interface makes finding a trim condition and creating a linear model simple. The 3D model can be used for simulation and moves during trimming to help the user with more intuitive feedback.

Finally, we developed a Scicos block to interface to JSBSim. A linear model can be used to design a controller in Scicos/ScicosLab and the JSBSim aircraft model can be simulated with dynamics and control elements within the Scicos environment.

Acknowledgments

The authors would like to thank Bill Galbraith for his work with Datcom+ which was used to create the Multiplex Easy Star model.

References

[1] Berndt J., "JSBSim: An open source flight dynamics model in C++". *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Providence, Rhode Island, 2004.

[2] Berndt J. and De Marco A., "Progress on and usage of the open source flight dynamics model software library, JSBSim". *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Chicago, Illinois, 2009.

[3] Campbell S. L., *Modeling and Simulation in Scilab/Scicos with ScicosLab 4.4*. Springer Science, Business Media LLC, 2010.

[4] De Marco A., Duke E., and Berndt J., "A general solution to the aircraft trim problem". *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Hilton Head, South Carolina, 2007.

[5] Kvasnicka V. and Pospichal J., "A hybrid of simplex method and simulated annealing". *Chemometrics and Intelligent Laboratory Systems*, Vol. 39, No. 2, 1997, pp. 161–173.

[6] Nelder J. A. and Mead R., "Simplex method for function minimization". *Computer Journal*, 1965, pp. 308–313.

[7] Price W. L., "Global optimization by controlled random search". *Journal of Optimization Theory and Applications*, Vol. 40, 1983, pp. 333–348.

[8] Stevens B. L. and Lewis F. L., *Aircraft control and simulation*, Wiley, New York, 2003.

Flutter Certification of SKYCAR Aircraft: Rational Analysis and Flight Tests

R. Pecora[a], M. Pecora[b], L. Lecce[a]
[a]*Dipartimento di Ingegneria Aerospaziale - Università degli Studi di Napoli "Federico II"- ITALY*
[b]*Centro Italiano Ricerche Aerospaziali (C.I.R.A.) S. c. p. A. - ITALY*

G. Verde[c]
[c]*OMA SUD - Sky Technologies S. p. A. - ITALY*

Keywords: *SkyCar, Flutter Certification, Flight Flutter Tests, Aeroelastic Model Validation*

Abstract

SKYCAR is a CS-23 category aircraft designed and produced by OMA SUD to be the actual sole multi-utility/multipurpose aircraft in the General Aviation High Performance Piston marketplace with an opening bonnet on fly. The out-of-standards configuration -characterized by a double vertical tail surmounted by a fully movable horizontal tail and connected to the fuselage by means of two large booms- represented the key-reason of a quite challenging proof of compliance to EASA requirements in terms of flutter instability. Due to the inapplicability of consolidated simplified criteria, rational analyses supported by test evidence were considered mandatory, thus demanding for a wide set of numerical and experimental activities to be performed. In this paper a general overview of such activities has been presented by pointing out the used approaches, analysis methods as well as the numerical tools implemented on their base.

1 Introduction

SKYCAR is a CS-23 [1] category aircraft (Fig. 1) designed and produced by OMA SUD to be the actual sole multi-utility / multipurpose aircraft in the General Aviation High Performance Piston marketplace with an opening bonnet on fly.



Fig. 1 - SKYCAR Aircraft

The internal layout is flexible and can be changed in various configurations and for different uses such as aero taxi, light cargo as well as for special military and civil missions. A synergetic and innovative architecture has been conceived in order to assure an unequalled cabin volume thus providing a top class passenger comfort and the largest baggage compartment among same class aircraft, while preserving high aerodynamic efficiency and low fuel consumptions.

The resulting two tail-booms configuration, and the chosen solution of a fully movable horizontal tail mounted on a double fin structure, were the key-reasons of a quite challenging proof of compliance to EASA requirements, especially with reference to aircraft flutter clearance justification. As a consequence of the unconventional aircraft layout, simplified criteria resulted impracticable both for the estimation of stiffness and inertial distributions and for the evaluation of control

surfaces mass-balancing useful to avoid any aeroelastic instability within flight envelope. A rational approach -based on advanced modeling techniques and numerical tools in close interaction with experimental tests- was then followed in order to produce the required means of compliance to paragraph 629 of reference airworthiness regulations [1]. The approach was characterized by two different and subsequent levels of analysis (Fig. 2), the first level covering all the activities oriented to the release of the "permit to flight" certificate up to dive

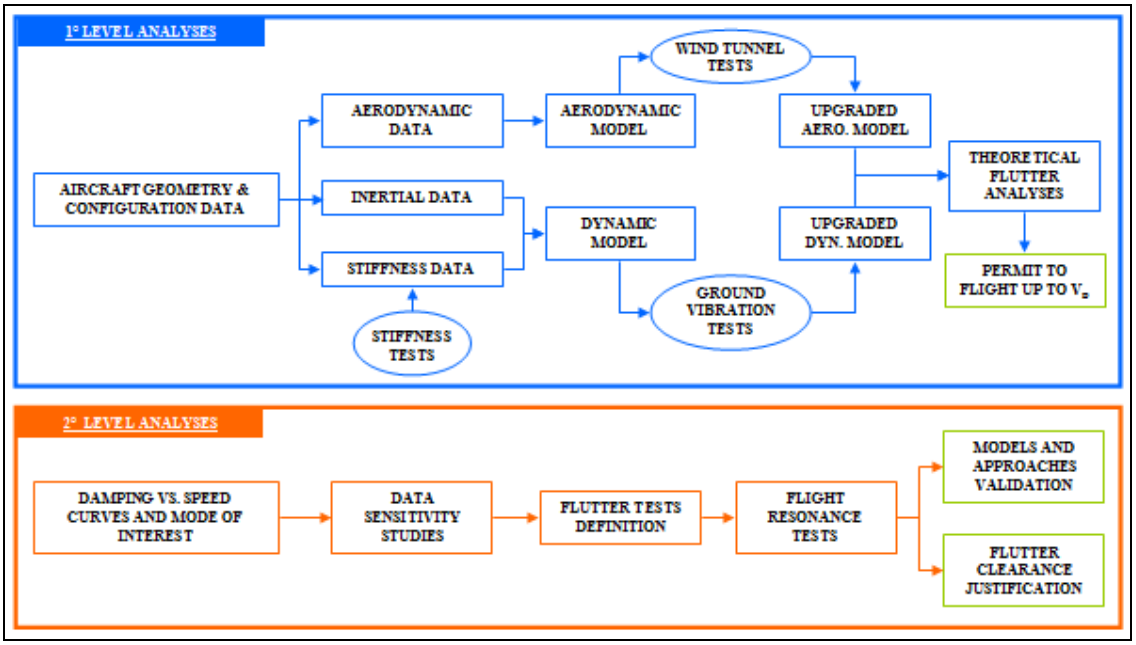


Fig. 2 - Flow chart of the approach used for flutter certification

speed (V_D = 184 kts) and the second properly referred to flight flutter tests and to the release of type certificate.

At the first level of analysis, the aircraft dynamic model was built and the related modes, frequencies and generalized masses were numerically determined. Theoretical modal parameters were then correlated to ground vibration tests results, thus allowing for dynamic model upgrading by means of slight corrective actions on assumed stiffness/inertial distributions.

Three-dimensional panel method for unsteady compressible flow simulation was used in order to generate aircraft aerodynamic model and to calculate modal pressure distributions. Relying upon wind tunnel test data, local corrections were made to pressure distributions for a proper simulation of control surfaces overall aerodynamics and hinge moments.

Upgraded dynamic model and generated aerodynamic lattice were finally matched and flutter speeds were evaluated in correspondence of several investigation cases characterized by different aircraft's inertial distributions and/or flight altitudes. On the base of obtained results, both in terms of flutter speed/frequencies and flutter mechanism, the permit to flight up to dive speed was released and the "go ahead" to the second level analyses and to flight flutter tests was given by airworthiness authorities.

Flight flutter tests were preceded by sensitivity analyses aimed to individuate the most promising approach for the in-flight excitation of the mode of interest (flutter mode), to predict aircraft time-responses under such excitation and to consequentially find out the optimal accelerometers disposition for data recording. All the analyses up to flight flutter test were performed by using the in-house developed code SANDY3.0 (whose general

description has been provided in Appendix) in combination with MSC-NASTRAN.

Experimental trends of modal frequency/ damping vs. flight speed showed the absence of any flutter instability up to $1.2V_D$; moreover, the excellent correlation level of experimental trends to theoretical expectations represented a means of validation for the adopted modeling strategies, investigation approaches and a further reliability proof of the used computational tools.

2 Aircraft Aeroelastic Model Generation

Aircraft (A/C) aeroelastic model was built up on the base of input data (mainly coming from OMA-SUD drawings and preliminary finite elements models) and following three consecutive phases:

Phase 1 generation of A/C dynamic (structural and inertial) model for theoretical modes evaluation;

Phase 2 generation of A/C aerodynamic lattice for unsteady aerodynamic forces evaluation trough DLM (Doublet Lattice Method);

Phase 3 generation of matching model (spline matrices at significant grids) for modes interpolation on aerodynamic lattice.

In the next two subparagraphs a general description of the approaches covering the above mentioned phases has been reported.

2.1 Dynamic Model

A/C dynamic model was obtained as the result of a procedure made up of two main steps:

- definition of the dynamic model for each A/C item (grids, stiffness and mass matrices at grids);
- assembly of items' dynamic models trough linkage elements.

Depending on item's characteristics and/or related input data formats, the dynamic model definition was carried out by referring to one of the two following approaches:

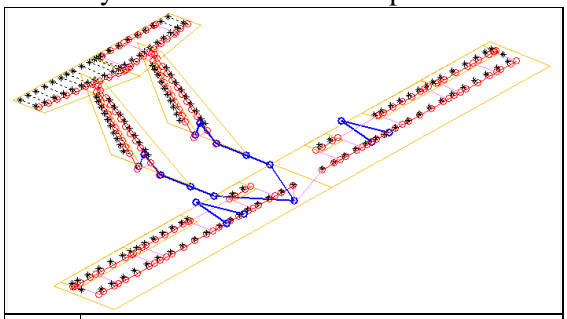
(a) stiffness and mass matrices definition by means of DMIG (Direct Matrix Input at Grids, [2],[3]) and/or set of DMIG related to significant item's grid nodes and derived by available item's FE (Finite Elements) models ;

(b) stiffness and mass matrices definition from an equivalent beam representation of the item.

Approach (a) was used for engine-mount and rear fuselage (here intended as the fuselage portion between wing front spar's joint up to tails connection frames) and control lines (all mechanical from the cockpit up to control surfaces).

Approach (b) was used for wing, fin and for all control surfaces. Front fuselage was taken in account only in terms of its inertial contribute and therefore modeled by means of lumped masses in offset to significant grids of rear fuselage model.

The linkage between item's structural models was performed just by means of rigid or beam elements; beam elements simulating the connection between main lifting surface and related control surface (i.e.: wing-aileron, fin-rudder, fin-stabilator) were characterized by suitable stiffness properties and pin flags [2,3] enabling control surface's rotation around its hinge axis. In Fig. 3 the A/C dynamic model has been sketched; for clearness sake, fuselage lumped masses, not-structural masses (fuel, payload, equipments) and DMIGs related to control systems have not been represented.



—	ITEM PLANT BORDER
—	ITEM BEAM ELEMENT
○	ITEM NODE
*	MASS CG
- - -	MASS OFFSET VECTOR
○	REDUCTION NODE OF A DMIG
—	DMIG ELEMENT
—	LINKING ELEMENT

Fig. 3 - Dynamic model layout

2.2 Aerodynamic and Matching Models

Aerodynamic lattice (Fig. 4) was characterized by 682 boxes obtained by meshing 10 macro-panels, three of which related to the wing, three to horizontal tail and two to each vertical tail unit.

For each panel, chordwise and spanwise mesh sizes were rationally chosen taking also in account the following criteria:

- higher boxes density all over control surfaces;
- avoidance of wake interference between parallel panels (i.e.: spanwise alignment of the boxes boundaries belonging to wing and horizontal tail);
- mesh continuity between panels related to vertical and horizontal tails (in order to properly simulate tails' mutual aerodynamic interference).

No panels were added to simulate the fuselage and the engine nacelle (their contribute was however taken in account by locally correcting aerodynamic model's pressure distributions on the base of wind tunnel tests, see paragraph 3).

Auxiliary grids were properly defined and rigidly connected to dynamic model's significant nodes in order to support the generation of spline matrices for modal displacements interpolation at the centroid of each aerodynamic box.

Spline matrices reliability was checked by performing (numerical) interpolation tests on imposed rotation modes of control surfaces around their hinge axes; more in detail, for each control surface:

1. an unitary rotation around control surface's hinge axis was imposed as a set of modal displacements at dynamic model grids;
2. modal displacements were interpolated along control surface boxes' and the pressure distribution due to the resulting boxes deflections was evaluated;
3. an unitary rotation around control surface's hinge axis was imposed by directly deflecting its aerodynamic boxes and the resulting pressure distribution was evaluated;
4. spanwise loads distributions coming from the integration of aerodynamic pressures obtained at point 2 and 3 were checked to be equal.

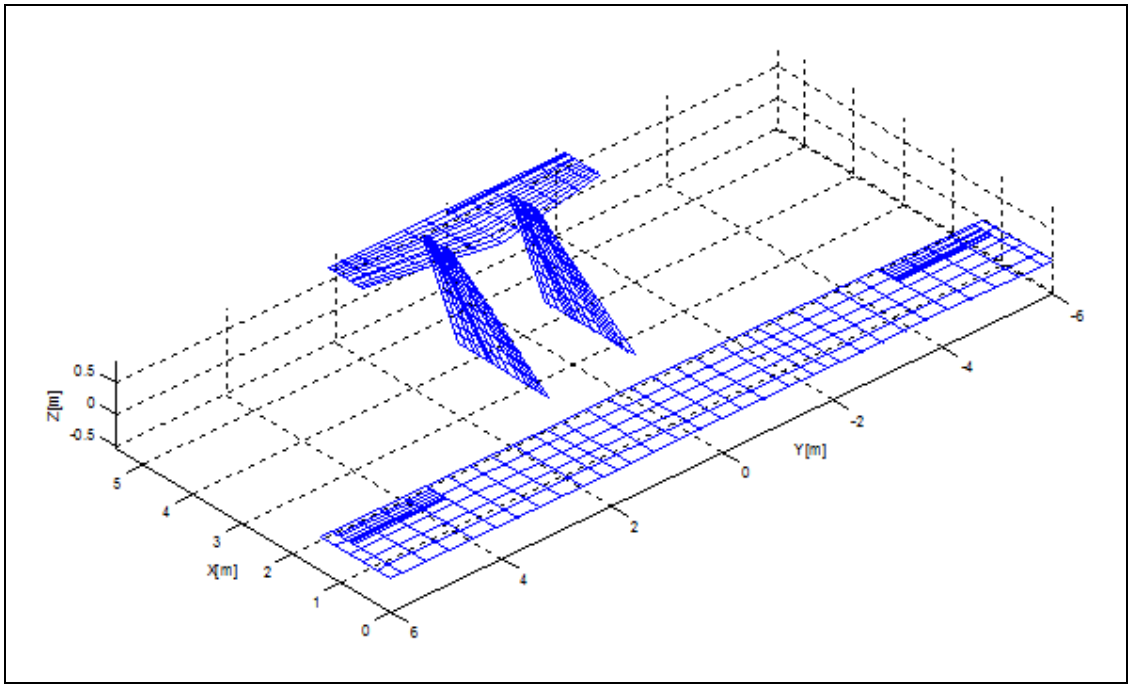


Fig. 4 - Aerodynamic Model

3 Models Validation and Flutter Analyses

Preliminary flutter analyses were carried out in order to fix a proper degree of mass-balancing (d.o.b.) for each control surface.
The robustness of adopted mass-balancing solution (see Table 1) was proved with reference to several A/C mass configurations, flight altitudes up to the maximum operative one, ±25 % variations of control lines harmonics with respect to their theoretically estimated value (both in stick-free and stick-fixed attitudes) as well as in correspondence of control lines failures.

CONTROL SURFACE	d.o.b. [%]
Aileron	80
Stabilator	80
Rudder	74

Table 1 – Control Surfaces d.o.b.

Ground vibration tests (GVT) were then performed in correspondence of the two A/C mass configurations believed most significant on the base of preliminary flutter analyses' results: MZFW (Max Zero Fuel weight) and MTOW@MF (Max Take Off Weight @ Max Fuel). In order to increase the level of correlation between theoretical and experimental modes, slight corrections were made on dynamic model with specific reference to:

- structural mass distributions of the items modeled according to approach (b) (see paragraph 2.1);
- control line stiffness.

Theoretical A/C modes, coming from upgraded dynamic model, resulted in extremely good correlation with the experimental ones both in terms of shapes and natural frequencies; same applied to control surfaces harmonics. For the sake of synthesis, in Fig. 5, the frequency correlation plot has been reported with reference to MTOW@MF configuration only.

Relying upon wind tunnel test data, local corrections were made to theoretical pressure distributions in order to refine the simulation of control surfaces overall aerodynamics and hinge moments values. The resulting upgraded aerodynamic model was then successfully validated on the base of flight test data.

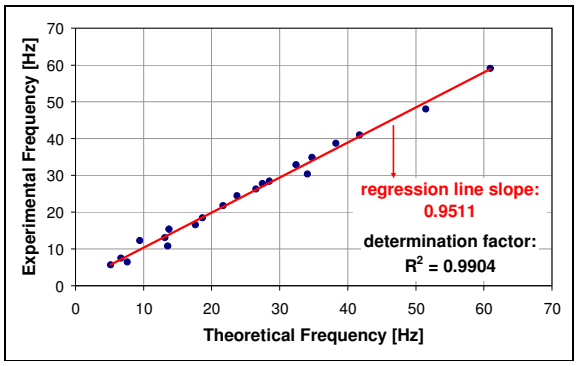


Fig. 5 - Modes Frequencies correlation plot (MTOW@MF)

Theoretical flutter analyses were then addressed by referring to upgraded models and in correspondence of the same A/C mass configurations and flight altitudes used in preliminary aeroelastic instability calculations. Control surface mass-balancing was set equal to the values reported in Table 1 and modal damping was conservatively assumed equal to 1% for all elastic modes. Flutter instabilities were searched in the speed range [0:1.8V_D] and no flutter was found within the certification envelope (i.e.: up to 1.2V_D for stick-free / stick-fixed control surfaces; up to V_D in case of control surfaces failures). The first (and only) flutter instability within investigation range was detected at 1.5V_D (at sea level; 1.4V_D at maximum operative altitude) and in correspondence of MTOW@MF configuration and stick-free condition; in Fig. 6 the related frequency/damping versus speed diagram (Vg plot) has been reported.

The flutter mechanism was deeply analyzed and discovered to be essentially due to the participation of three modes: fuselage lateral bending (flutter mode, curve 7 of Fig. 6), rudder symmetric harmonic and rudder fundamental; the first two modes (Fig. 7) being at quite the same frequency, around 7 Hz. The reason of the coupling was attributed to the presence of a symmetric rudder deflection in lateral fuselage bending mode.

Although outside the certification envelope, it was however decided to investigate the frequency/damping trends of the flutter mode by means of dedicated flight flutter tests.

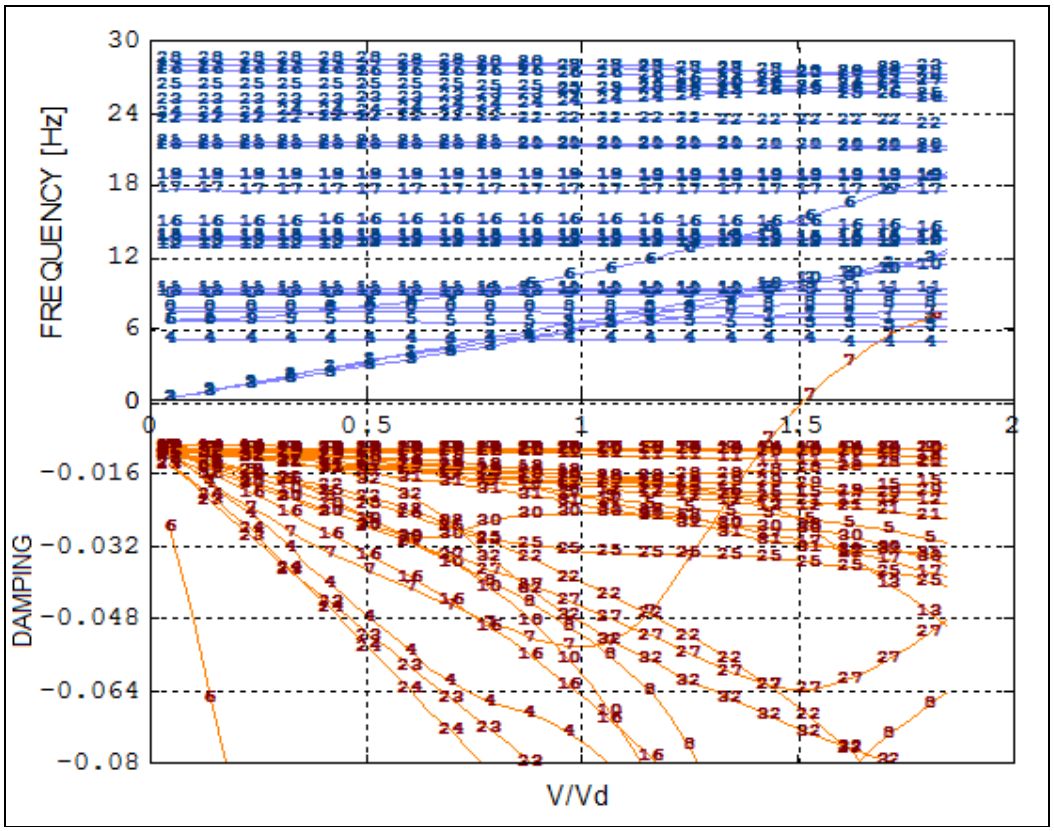


Fig. 6 - Vg plot, MTOW@MF, stick-free condition, sea-level

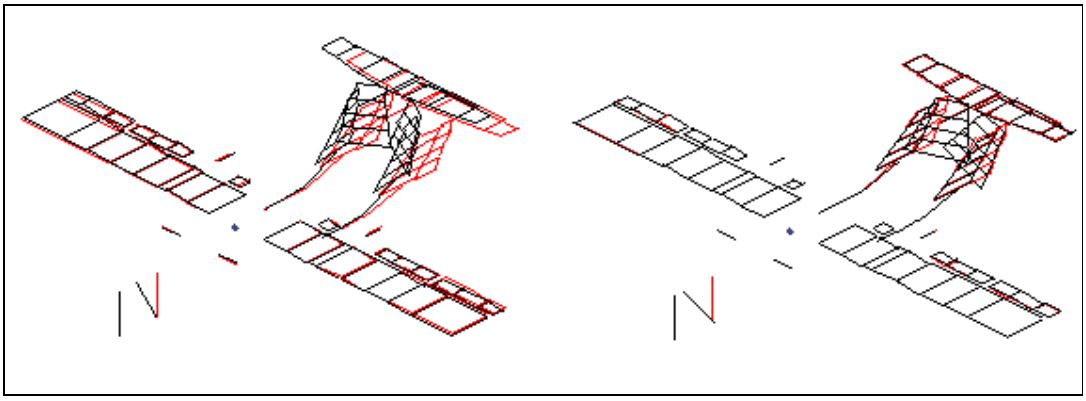


Fig. 7 - Fuselage lateral bending (left) and rudder symmetric harmonic (right)

4 Flight Flutter Tests

Flight Flutter Tests (FFT) were performed according to a methodology characterized by three main steps for each investigated flight speed:

1. in-flight excitation of flutter mode (fuselage lateral bending, see paragraph 3) by means of impulsive rudders deflection consequent to pilot's action on aircraft pedals (spike and/or doublet);
2. acquisition (at 128 samples per second) of the time histories related to side accelerations measured on left and right vertical tail units;
3. elaboration of accelerometric signals acquired at step 2 and identification of flutter frequency and damping values.

Investigated flight speeds ranged from 120 kts up to A/C V_D (184 kts). Flight speed range was covered by means of two different flights according to what reported in Table 2.

FLIGHT ID.	INVESTIGATED FLIGHT SPEEDS [KTS]
95	120,135,150,165
98	174,184

Table 2 - Flights Identification Numbers and Investigated Flight Speeds

In order to remove propeller induced noise (at 40 Hz) from accelerometric measurements, a 32 Hz low-pass digital filter was added to the acquisition chain.

For each test point (i.e.: investigated speed) the spectrum of the accelerations time histories was iteratively auto-correlated in order to isolate its main frequency component [4]. Autocorrelations transfer function were then identified in frequency domain by means of a polynomial ratio; by properly working on the poles of such polynomial ratio, frequency and damping values of flutter mode were finally evaluated (Table 3).

TEST POINT	SPEED [KTS]	DAMPING	FREQUENCY [HZ]
95_1	120	-0.0198	6.3492
95_2	135	-0.0246	6.3568
95_3	150	-0.0175	5.8792
95_4	165	-0.0186	6.2992
98_1	174	-0.0208	6.1886
98_2	184	-0.0156	6.0709

Table 3 - Flutter mode experimental frequency/damping vs. flight speed

Frequency and damping values reported in Table 3 resulted perfectly matching the theoretical trends (curve 1 of Fig. 8) thus further proving the high reliability of numerical models and adopted approaches as well as the robustness of aircraft's aeroelastic stability.

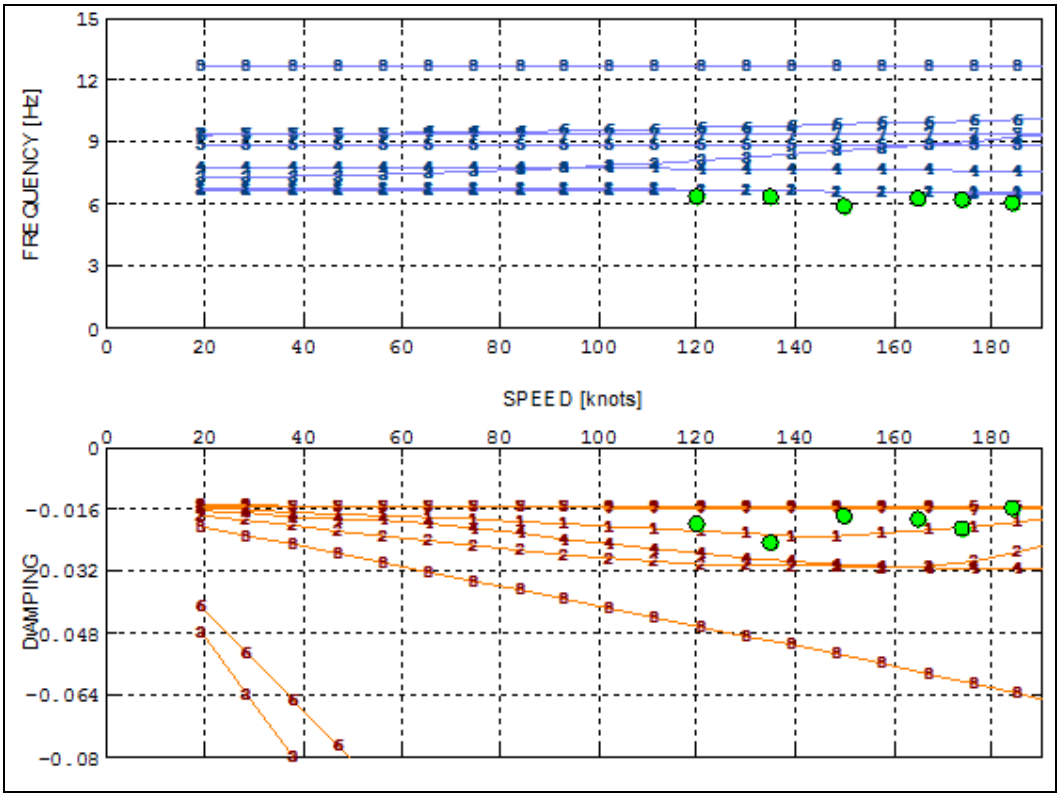


Fig. 8 - Flutter mode frequency/damping trends (curve 1: theoretical; green dots: experimental)

5 Conclusion

A rational approach based on advanced modeling techniques and numerical tools in close interaction with experimental tests was followed in order to prove SKYCAR aircraft compliance with paragraph 629 of EASA CS-23 airworthiness regulations. The approach was characterized by two different and subsequent levels of analysis: the first level covering all the activities oriented to the release of the "permit to flight" certificate up to dive speed and the second properly referred to flight flutter tests and to the release of type certificate.

The entire flutter certification program was carried out by mainly referring to in-house developed tools and resources. It covered a time frame of one year ending on January 2010; FAA certification was released only four month after.

The very short time in which such results were obtained not only represents a proof of reliability for the adopted methodologies and implemented tools but rather witnesses the great capabilities and commitment of a technical team which never stopped to power up the SKYCAR program with an extraordinary *spirit of achievement*.

Appendix: SANDY3.0 general description

SANDY code has been developed and upgraded within last twenty-five years with the intent of providing an excellent and reliable tool for static and dynamic aeroelastic and aero-servo-elastic analysis of aircraft.

Rational approaches and validated numerical methods, compliant with EASA standards CS-25 and CS-23, have been implemented in a multidisciplinary computational environment able to accomplish the following main tasks:

- generation of A/C dynamic model (structural model and inertial model);
- generation of A/C aerodynamic model;
- generation of accurate transfer matrices interfacing between dynamic and aerodynamic models;
- evaluation of A/C acceleration and loads response due to flight and ground manoeuvre and/or gust;
- evaluation of A/C static and dynamic acceleration and loads response to movable lifting surfaces deflections imposed by mechanical and/or electro-mechanical control circuits;
- evaluation of divergence, control reversal and flutter speeds.

The computational tool, and the numerical methods within implemented, assure fast analyses aimed also to investigate the influence of several design parameters on A/C aeroelastic behaviour; in other terms, it is provided the capability of fast sensitivity aeroelastic analysis in correspondence of variations in structural and dynamic properties pertaining to A/C components including also the integrated control circuits.

The high efficient computational structure is based on intercommunicating routines running in a common software environment (Visual Fortran / Matlab) and properly organized in three macro-modules:

- Dynamic module;
- Aerodynamic module;
- Aeroelastic module.

The dynamic module includes a generator of A/C structural and inertial models as well as an internal FEM solver for the evaluation of A/C modal parameters. In order to perform a reliable modelling in reasonable time, only two typologies of elements are available:

- simple mono-dimensional elements (beam, rod, rigid-link, spring);
- DMIG (Direct Matrix Input Generalized) elements for a correct evaluation of stiffness and inertial properties of all those parts for which a mono-dimensional elements model-ling leads to high approximations and/or is not applicable at all.

DMIG elements may also be imported by commercial FE software (e.g.: MSC NASTRAN); same applies to A/C modal parameters (modes shapes, frequencies, generalized masses, damping) which, moreover, may be imported directly from

experimental data obtained by means of ground vibration tests

In order to simulate loads alleviation systems based on movable surfaces in feed-back with A/C cinematic parameters (acceleration, velocity, displacements) and/or to perform fast instabilities analyses in presence of powered or fly-by-wire controls, the module is properly structured for the introduction of extra-modes representing movable surfaces deflections in addition to the normal modes related to A/C structure.

The aerodynamic module is characterized by a generator of aerodynamic model coupled to an internal solver for the evaluation of modal pressure distributions by means of consolida-ted numerical methods (e.g.: Doublet Lattice Method). More in detail it is provided the capability of generating 3D (three-dimensional) compressible aerodynamic models as well as the possibility of performing local corrections in pressure distributions in order to achieve a correct simulation of movable surfaces aerodynamics (as recommended by FAA AC 25.629-1A par. 6).

The aeroelastic module is finally constituted by all the routines devoted to:

- the interpolation of modal displacements on the aerodynamic lattice (matching routines implementing 3D spline methodologies);
- the evaluation of generalized aerodynamic pressures/forces and steady/unsteady aerodynamic influence coefficients;
- the solution of aeroelastic equations related to A/C aeroelastic response and aeroelastic instabilities identification.

Acknowledgments

The authors wish to acknowledge OMA-SUD technical office for the extraordinary and tireless support to all the activities which made possible the publication of this paper.

A special acknowledgement goes to ENAC Officer Ing. Corrado Maragò for his constant dedication to SKYCAR flutter program up to his premature demise.

This paper is dedicated to his memory.

References

[1] European Aviation Safety Agency (EASA), *"Certification Specifications for Normal, Utility, Aerobatic, and Commuter Category Aeroplanes – CS23"*, Edition: November 2003

[2] MSC-SOFTWARE, *"MD-NASTRAN 2006 Quick Reference Guide"*.

[3] MSC-SOFTWARE, *"MD-NASTRAN 2006 Reference Manual"*.

[4] Hubolt J.C., "On Identifying Frequencies and Damping in Subcritical Flutter Testing", *NASA SP415-Flutter Testing Techniques,* Dryden Flight Research Center, October 1975, pp. 10-41.

CEAS 2011

Materials 1

3rd CEAS Air&Space Conference
21st AIDAA Congress

Flow analysis during Resin Transfer Molding of thick structures

S. Laurenzi, M. Pinna, A. Grilli, M. Marchetti.
Sapienza University of Rome, Department of Astronautic, Electric and Energetic Engineering.

Keywords: *Resin Transfer Molding, Thick composite structures.*

Abstract

The Resin Transfer Molding (RTM) process is widely investigated as a powerful method to fabricate aerospace structure. So far RTM applications concern thin laminates and very little work on thick composites can be found in the literature.

Nevertheless, an investigation of thick composites is necessary in order to extend the RTM advantages to primary structure. To achieve fast, repeatable and reliable process one needs to acquire some predictive modeling capability to simulate the resin flow within the mold during the injection. This capability is readily available for RTM process modeling, but it has to be extended to the fabrication of thick structures.

We present experimental and numerical studies to determine a corrective factor of the preform permeability to take into account the effects of the transverse resin flow through the thickness in 2-D simulations. Numerical studies were compared with experimental cases.

1. Introduction

The Resin Transfer Molding technology is a competitive process to produce aeronautics composite structures. Nevertheless, the major limitation of RTM is the lack of confidence from the industry sector due to its preference towards more traditional processes, such as autoclave. Simulations are useful to expand the applications of RTM process, but predicting the flow front in thick preform is challenging and, in most cases, three-dimensional flow simulations are highly inefficient [1]. On the other hand, two-dimensional flow simulations of a thin arbitrary shape have a good correspondence with the real behavior of the filling phase, and it is computationally effective. It would be useful to take into account in the model the transverse flow features by means of anisotropic material properties in the plane.

Permeability is the most significant preform parameter that governs the resin flow. A corrective factor for the in-plane permeability involving transverse flow effects can reduce the three-dimensional equation system to a two-dimensional one [1-6].

In this paper, experimental and numerical investigations of the resin flow in thick sections of aeronautic structures are presented. An experimental set-up was developed to observe the resin flow during the filling phase.

In order to determine the 3-D permeability tensor, experiments were performed on carbon fiber laminates with a given lamination sequences and fixed fiber volume fraction.

The idea is that a thick laminate can be considered as formed by a finite number of

packages of layers. After determining the in-plane permeability of several groups of packages, it is possible to extrapolate the numerical factor related with the transverse permeability that influences flow front behavior.

Simulations based on finite element/modified control volume method were conducted considering the 2-D modified permeability. Simulations were compared with two real cases of thick plates realized with different resin.

2. Theory

The RTM process is generally modeled using the porous media flow approach. Figure 1 shows a control volume extrapolated from a fiber reinforcement, formed by fibers and porosity. The continuity equation is carried out from the mass balance of the flow rate in the control volume:

net flow of mass = inflow - outflow – losses

The loss term is due to a volume V inside which the fluid mass adsorbs.

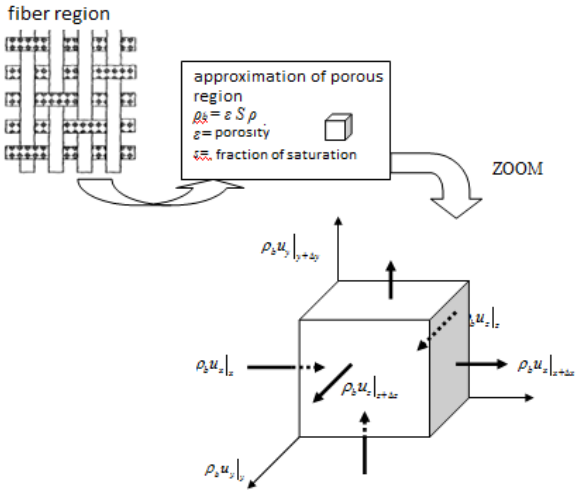


Figure 1: concept of control volume for fiber region.

The resin interstitial velocity inside the composite is given by Eq. (1).

$$U = u_x \boldsymbol{i} + u_y \boldsymbol{j} + u_z \boldsymbol{k} \tag{1}$$

The resin density ρ_b is the mass of resin divided by the control volume containing both resin and fiber. Within a region partially filled fibrous, ρ_b is given by Eq. (2),

$$\rho_b = \varepsilon S \rho \tag{2}$$

where ε is the porosity of the volume control that is occupied by the resin, which can be expressed as $\varepsilon = 1\text{-}v_f$. Here v_f is the volume fraction of fiber, S is the fraction of saturation of pore space filled with a resin, and ρ is the actual density of the resin (mass of resin/volume of resin). The pore volume within the volume control is completely empty when S = 0, and completely full when S = 1.

Therefore, the mass balance can be expressed as:

$$\frac{\partial(\varepsilon S\rho)}{\partial t} = -u_x\frac{\partial(\varepsilon S\rho)}{\partial x} - u_y\frac{\partial(\varepsilon S\rho)}{\partial y} - u_z\frac{\partial(\varepsilon S\rho)}{\partial z} - \varepsilon S\rho\frac{\partial u_x}{\partial x} - \varepsilon S\rho\frac{\partial u_y}{\partial y} - \varepsilon S\rho\frac{\partial u_z}{\partial z} - s \tag{3}$$

The general governing equation that describes the resin flow into the open channels is the Navier-Stocks for viscous and non-compressible fluid. Nevertheless, considering that the Reynolds's number is less than the unity, the resin flow throughout the preform can be well described by the empirical Darcy's law given by Eq. (4):

$$\begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = -\frac{1}{\mu(1-v_f)} \begin{pmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{pmatrix} \begin{pmatrix} \frac{\partial P}{\partial x} \\ \frac{\partial P}{\partial y} \\ \frac{\partial P}{\partial z} \end{pmatrix} \tag{4}$$

The Darcy's law describes the flow of a viscous fluid through an anisotropic, homogenous, porous medium. Eq. (4) is written in components along the three principal axes as indicated in Figure 2. Eq. (4) gives the medium velocity profile that is directly proportional to the tensor of permeability $\overline{\overline{K}}$ and the pressure gradient $\overline{\nabla P}$, then μ is the resin viscosity.

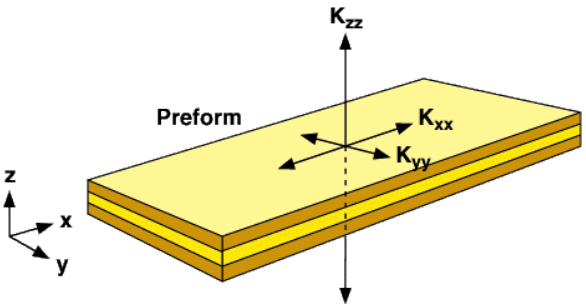


Figure 2: rectangular preform with the indication of the principal directions x y z. K_{xx}, K_{yy}, K_{zz} are the permeability components along x,y,z.

The tensor permeability is symmetric, so if the directions of the reference frame are selected along the main directions of the preform, as it appears in Figure 2, the matrix becomes diagonal.

The permeability value depends on the preform and needs to be experimentally determined. In fact, a great number of the authors tried to calculate it by numerical and theoretical approaches, but the results do not well fit with the experimental data.

Darcy's law and continuity equation were combined to model the flow process.

RTM parts typically have a thickness substantially less than the other dimensions [1-6], so the changes in the flow through the thickness can be neglected assuming a two-dimensional model and the in-plane permeability. For thick sections, this assumption falls. The transverse flow becomes very significant.

Figure 3 shows the typical behaviour of the resin flow for such case.

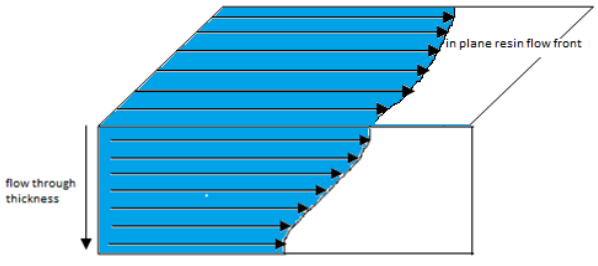


Figure 3: typical trend of in-plane and through thickness resin flow in a thick perform.

The transverse flow is affected mainly from the permeability in the z direction, that results generally one order of magnitude lower than in-plane permeability [2].

In order to establish a relationship between the in-plane and transverse permeability, we make the assumption that the flow encounters various simple packages as shown in Figure 4.

The packages are made of few layers of given orientation. The different orientations change the local porosity, and so the local permeability. The global flow can be split into two fronts: the main flow direction front or "between packages flow" and the transverse direction front or "through packages flow".

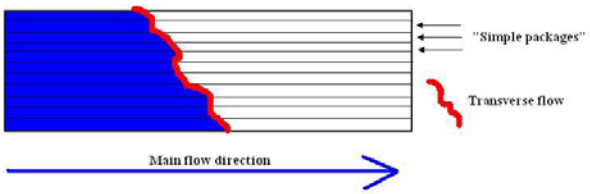


Figure 4: package subdivision of a generic thick plate.

The most significant front is obviously the main flow direction front; the transverse flow acts as a "brake" to the main one.

From these considerations, we investigated the corrective factor, that considers the "trough packages flow" effects in the in-plane permeability values.

3. Experimental

Carbon fiber reinforcements, with commercial name of Hexcel G0926 5H Satin, was used for the experimental tests. The fiber volume fraction was fixed at 0.55 and a quasi-isotropic sequence was adopted for each package. Mono-component systems such as the epoxy Hexcel RTM6 and the benzoxazine Henkel Epsilon II 99110 were used to fabricate the reference plates.

The in-plane permeability was determined adopting a one-dimensional method, carrying out the flow front position as function of the time. Eq. (5) shows the one-dimensional version of the Darcy,s law, that integrated gives Eq. (6).

$$\frac{\partial x}{\partial t} = -\frac{K_x}{\mu(1-v_f)}\frac{\partial P}{\partial x} \qquad (5)$$

$$x_f^2 = \frac{2K_x \Delta P}{\mu(1-v_f)} t_f \qquad (6)$$

Plotting Eq. (6) by the values determined experimentally, the in-plane permeability value along the x direction is calculated as

$$K_x = \frac{\mu(1-v_f)\text{slope of}\left[x_f^2(t_f)\right]}{2\Delta P} \qquad (7)$$

The thick composite plates of 300 mm x 260 mm and thickness of 30 mm were manufactured by RTM, with 0.55 fiber volume fraction and 75 layers. Figure 5 shows the manufactured plate.

The filling time of the plate was measured.



Figure 5: 30-mm thick carbon reinforced plate manufactured by RTM. The picture shows a detail of the thickness.

4. Numerical analysis

Figure 6 shows the individuated numerical correlation between the in-plane permeability and the number of packages. The permeability decreases with the increasing of packages.

Numerical analyses were performed to simulate the filling phase of the manufactured plates. In order to do that, the simulations considered the process parameters and the material features of the real cases. The permeability values were instead determined by the curve in Figure 6.

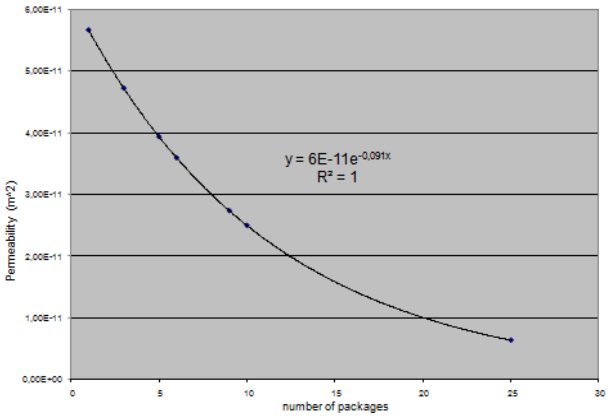


Figure 6: correlation between the in-plane permeability and the number of packages.

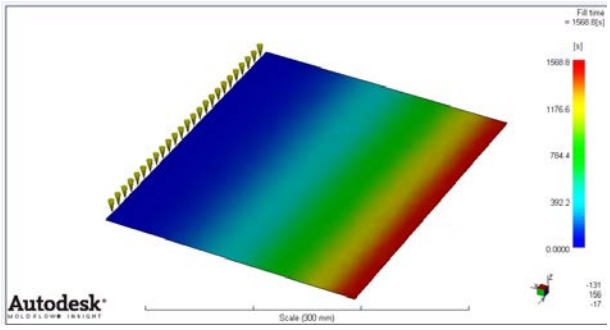


Figure 7: trend of the filling time. In such case, it was taken into account the viscosity curve of the Henkel Epsilon II 99110 resin and the in plane modified permeability of 6e-12 m^2.

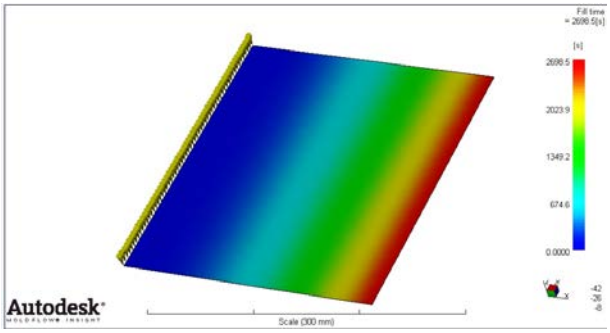


Figure 8: trend of the filling time. In such case, it was taken into account the viscosity curve of the Hexcel RTM6 resin and the in plane modified permeability of 6e-12 m^2.

Table 1 summarizes the numerical and experimental results obtained for the filling phase of the plate. A good approximation was found comparing the results.

Table 1: numerical and experimental results of filling time.

Resin	Numerical filling time (min)	Experimental filling time (min)
RTM6	~ 45	~ 45
Epsilon II 99110	~ 26	~ 27

5. Conclusion

The comparison between the experimental flow and the simulations shows that the corrective factor for the in-plane permeability can be successfully used to model the resin flow front for a thick structure by two-dimensional model. This result is important to design and manufacture primary structures while avoiding and/or limiting expensive trial and errors approach and computational costs.

Limitations are given by the strong dependence of the corrective factor on the preform characteristics, such as fiber volume fraction, type of reinforcement and lamination sequences.

These results are current applied to design the mold of an aeronautic beam, which is a complex large structure characterized by different thick sections.

References

[1] T. Kruckenbrg, R. Paton, *Resin Transfer Moulding for Aerospace Structures*, 2nd ed., Kluwer Academic Publishers, 2000.

[2] S. G. Advani, E. M. Sozer, *Process Modeling in Composites Manufacturing*, Marcel Dekker press, 2002.

[3] D.S. Cairns, D.R. Humbert, J.F. Mandell, "Modeling of resin transfer molding of composite materials with oriented unidirectional plies", *Composites Part A*, Vol. 30, 1999, pp 375 – 383.

[4] S. Amico, C. Lekakou, "An experimental study of the permeability and capillary pressure in resin-transfer moulding", *Composites Science and Technology*, Vol. 61, 2001, pp 1945 – 1959.

[5] Y. Luo, I. Verpoest, K Hoes, M. Vanheule, H. Sol, A. Cardon, "Permeability measurement of textile reinforcements with several test fluids", *Composites Part A*, Vol. 32, 2001, pp 1497 – 1504.

[6] S. Bickerton, "Modeling and control of flow during impregnation of heterogeneous porous media, with application to composite mold filling processes", Ph. D. dissertation, Department of Mechanical Engineering University of Delaware, 1999.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Reinforcements of CFRP laminates through polymer nanofibers produced by means of electrospinning process.

R. Palazzetti, A. Zucchelli L. Donati, G.Minak, E. Troiani
DIEM Department, University of Bologna, Italy

C. Gualandi, M. L. Focarete
Department of Chemistry 'Giacomo Ciamician', University of Bologna, Italy

M. Parente
Impregnatex S.r.l.,Italy

Keywords: *polymeric nanofibers, electrospinning,CFRP composite laminates, delamination*

Abstract

In the work polymeric electrospun nanofibers are used as interleaving filler placed between adjacent plies of a CFRP laminate in order to study the behavior of the composite material under Mode II ENF mechanical tests (End Notched Flexure, ASTM 7246). Influence of nanofibrous interlayer is investigated by comparing absorbed energy and bending stiffness respect to conventional CFRP laminates. Laminates are made of epoxy resin/carbon fiber 0/90 woven prepreg; nanofiber mats are manufactured using electrospun Nylon 6,6. For considered test (quasi-static one), one nanofibrous layer is placed in specimen's midplane. ENF tests reveal that the nanofibrous mats strongly contribute to improve bending stiffness before the material crisis (13,3% of increment) and a measurable increment of the maximum mechanical energy that can be absorbed by the material during the crack propagation (27,4% of increment) is registered.

1 Introduction

Fiber reinforced resin composites are typically formed by laying up a plurality of plies formed of reinforcing fibers and resin, cured by applying heat and pressure or by using a catalyst-promoter system. Due to the ply-by-ply nature of composite laminates, the susceptibility to delamination along interlaminar planes is an intrinsic and severe problem of these materials that can ultimately undergo structural failure. Interlaminar stresses concentration occurs at free edges, joints, matrix cracks, and under out-of-plane loading. Delamination is often the dominating failure mode in composite laminates subjected to impact and fatigue loading. Several methods have been developed over the years to enhance the delamination resistance of the advanced composite laminates, such as matrix toughening [1,2], optimization of stacking sequence [3-5], laminate stitching [6,7] braiding [8] edge cap reinforcement [9] critical ply termination [10] and so on. One approach initially proposed by Dzenis and Reneker in their patent issued in 1999 [11] involves the use of a nano-interlayer in the form of polymeric nanofibrous non-woven mat fabricated trough the electrospinning process. Electrospun nanofibers possess interesting properties mainly due to their extremely high surface area-to-volume ratio and, compared to conventional fibrous membranes, nanofibrous ones show higher porosity and smaller pores.

In their patent, Dzenis and Reneker demonstrated that a fiber reinforced composite laminate with polymeric electrospun

nanofibrous mats placed at one or more ply interfaces is characterized by an improved delamination resistance [11]. Indeed, composite laminates incorporating nanofibrous reinforcements are expected to have an improved interlaminar fracture toughness, strength, and delamination resistance towards static, fatigue, and impact load. In particular, nanofibrous reinforcement can be useful in suppressing and arresting delamination from any sources, including matrix cracks, free edges, notches or holes, ply drops, bonded joints, bolted joints, out-of-plane loading, and so on. The main function of nanofibers is to reduce the stress concentration due to mismatch of ply properties and to bond adjacent plies, without a reduction of in-plane properties and without increasing either the composite weight or the laminate thickness. Nanofibers can be small enough to reinforce regions of the matrix located in between adjacent plies of the laminate [12]. In addition, incorporating in laminated composites an entangled nanofiber layer, characterized by a combination of interlocked in-plane and out-of-plane nanofibers, might improve interlaminar fracture resistance, much like the "hooks and loops" in Velcro [13].

Following the pioneer idea of Dzenis and Reneker, other researchers have investigated the use of polymeric nanofibers to enhance the mechanical performances of laminated composites. However at present only few papers are available in the literature [14-19] and more research efforts are needed in order to investigate the benefits of electrospun nanofibers in reinforcing laminated composite materials. In reference [20] a comprehensive review of the recent works can be find.

This work investigates the use of an electrospun polymeric nanofibrous mat as the interleaved material to enhance composite laminate mechanical performances. Nylon 6,6 electrospun mats are produced directly on epoxy matrix/carbon fibers composite fabrics then inserted in the middle layer of a 20 plies stacking sequence and cured in autoclave. Final laminate is tested under End Notched Flexure (ENF) Mode II test following ASTM 7246 standard in order to analyze bending stiffness and absorbed energy proprieties.

2 Materials and Methods

2.1 Materials

The prepreg (GG205PIMP503) used for specimens fabrication is a 0/90 woven epoxy resin/carbon fiber prepreg (Table 1) and it was kindly provided by Impregnatex Compositi S.r.l. (Milan, Italy). The epoxy resin formulation used was based on a diglycidyl ether of bisphenol A prepolymer.

Table 1: Prepreg GG205IMP50 main properties (Source: Impregnatex Compositi S.r.l.)

Property	Value
Density	1.5 g/cm^3
Weight Resin	353 g/cm^3
Amount Resin	42%
Amount Thick	0.2 mm
PostCured T_g (DSC)	125 °C

Mechanical properties	Value
Bending Resistance	850 MPa
Bending Modulus	59 GPa
Tensile Resistance	650 MPa
Tensile Modulus	59 GPa
Short beam Shear	65 MPa

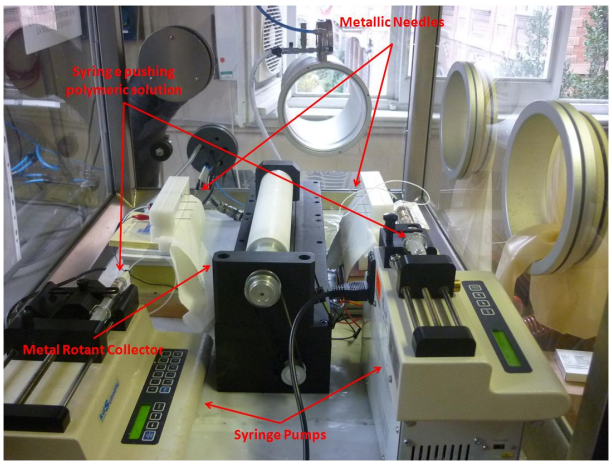
Nylon 6,6 Zytel E53 NC010 (Table 2) was kindly provided by DuPont. Formic Acid and Chloroform were purchased by Sigma Aldrich and were used without further purification.

Table 2: Nylon 6,6 properties (Source: DuPont)

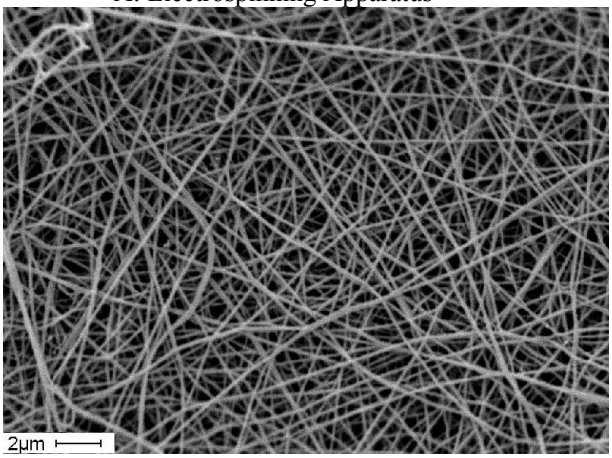
Properties	Value
Yield Stress	84 MPa
Nominal Strain at Break	>50%
Yield Strain	4.4%
Tensile Modulus	3 GPa
Notched Charpy Impact Strength	7 kJ/m^2

Electrospun non-woven mats were fabricated by using an in-house electrospinning apparatus (Figure 1A) composed of: (i) a high voltage

power supply (Spellman, SL 50 P 10/CE/230), (ii) a syringe pump (KDScientific 200 series), (iii) two chambers containing the polymeric solution, each one equipped with four stainless-steel blunt-ended needles and connected with the power supply electrode, and (iv) a grounded rotating collector (length=30 cm, diameter = 6 cm) positioned 10 cm away from the tip of the needles. The polymer solution located in two syringes, was dispensed, through a Teflon tube, to the two four-needle chambers that were positioned at the opposite sides of the collecting drum. Nylon 6,6 was dissolved in a Formic Acid/Chloroform solvent (50:50 v/v) at a concentration of 14% w/v. The electrospinning process was carried out under the following conditions: applied voltage 22-26 kV, feed rate 0.3 mL/h per nozzle, at room temperature and relative humidity RH = 40 ÷ 50%. Electrospun non woven mats (Figure 1B), 25 ± 8 μm thick, were kept under vacuum over P2O5 at room temperature overnight to remove residual solvents. Fiber morphology was observed with a Philips 515 scanning electron microscope (SEM) at an accelerating voltage of 15 kV. Prior to SEM analysis, samples were sputter-coated with gold. Fiber diameter distribution (150 ± 20 nm) was determined by measuring 200 fiber diameters with an image acquisition software (EDAX Genesis). Thermal properties of the Nylon 6,6 electrospun mat were investigated by means of differential scanning calorimetry (DSC) using a TA Instruments Q100 DSC equipped with the LNCS low-temperature accessory. The Nylon 6,6 nanofibrous mat was found to be particularly suitable for the purpose of the present work due to its chemical and physical characteristics: it shows excellent adhesion to epoxy matrix and it is characterized by a high-melting crystal phase (T_m = 262°C, ΔHm 65 J/g, by DSC). Therefore it was observed that the prepreg curing treatment at a temperature (150°C) well below the Nylon 6,6 melting temperature, did not cause any modification in the Nylon 6,6 electrospun mat shape or in fiber morphology. Electrospinning process and chemical tests were performed at the G.Ciamician laboratory of Chemistry Faculty of University of Bologna.



A: Electrospinning Apparatus



B: SEM micrograph of nylon 6,6 nanofibers

Figure 1: Nylon 6,6 electrospinning

Many parameters influence the final morphology and the quality of the fibers [21,22]: the most important ones are the amount of each solvent, concentration of the polymer, flow rate, room temperature and humidity, applied tension and distance between nozzle and collector. Such effects are linked each other, and it causes the optimizing process to be a long "trial by error" iterative process. In general it is observed that the uniformity and the cylindrical shape of the fibers increase with the polymer concentration in the solution, as well as the fibers diameter. At lower concentration, the fibers become thinner but the amount of beads and defects increase. The formation of beads can be prevented by increasing the electrical conductivity of the solution (which itself influences the choice of the solvents), or the intensity of the electrical field; increasing the voltage needs the increasing of the flow rate [23-25]

2.2 Specimen fabrication

Laminates were fabricated by hand lay-up process, stacking twenty plies with fiber directions parallel to the laminate borders. During such process, nanomodified layer is located as the 10th layer in order to place the nanofibers in the specimen middle interface. On the border, a Teflon sheet is introduced in the 10th layer too, to induce the initial specimen delamination during ENF testing. After the lay-up, panels were cured by using a vacuum bag in autoclave according to process specifications provided by the prepreg supplier. Laminates curing process was realized at the MasterLab laboratory of the Second Facoluty of Engeneering of University of Bologna in Forlì. Vacuum, pressure and temperature cycles are shown in Figure 2. Beam specimens were cut from the panel using an high-speed rotating diamond disc. From each panel, 5 specimens were obtained.

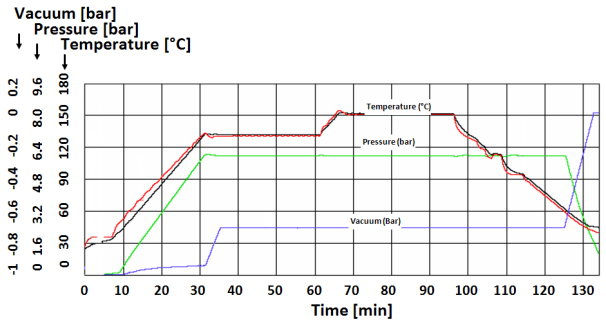


Figure 2: Prepreg curing process

2.3 Test methods

In order to investigate the behavior of nanomodified laminates under Mode II fracture condition, End Notched Flexure (ENF) tests were performed following ASTM 7246 standard. Results of nanomodified specimens are compared to virgin ones. Four specimens were tested for each configuration.

End Notched Flexure test (Figure 3) is widely used for experimental evaluation of resistance to in-plane shear stress of delamination and it is known as Mode II fracture mechanism. This is a Three Point Bending test where the specimens are characterized by an initial delamination. The delamination is 25mm long, provided by a 15µm thick teflon sheet placed in between the 10th and 11th ply, i.e. the layer where the nanofibers might be placed and the shear stress reaches its maximum. Specimens are 160mm long, 20mm wide, 4mm thick, the support spam is 100mm. Specimens were tested with and without nanofibrous layer. Tests were performed under controlled displacement condition with 1.5mm/min constant crosshead rate; acquisition frequency was 10Hz. Maximum tension, Flexural Modulus and Absorbed Energy are determined for each specimen, and the two conditions, virgin and nanomodified, were compared.



Figure 3: ENF Test

Stress (σ), strain (ε) and Flexural Modulus (E^{flex}) are calculated using the international standard ASTM 7246 [26]:

$$\sigma = \frac{3 \cdot P \cdot L}{2 \cdot b \cdot h^2} \quad (1)$$

$$\varepsilon = \frac{6 \cdot \delta \cdot h}{L^2} \quad (2)$$

$$E^{flex} = \frac{L \cdot m^2}{4 \cdot b \cdot h^3} \quad (3)$$

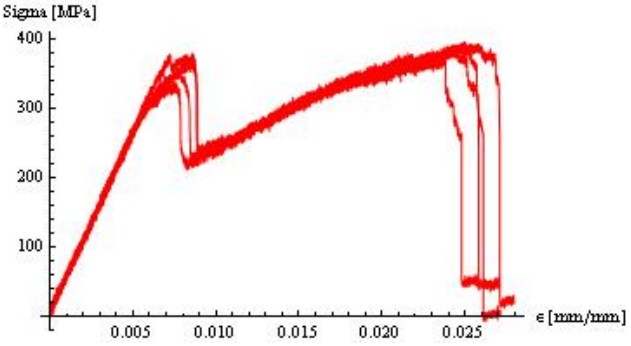
where:

- P and δ are applied load and opening displacement respectively;
- L, b and h are the specimen length, width and thickness respectively;
- m is the slope of the secant of the force-deflection curve.

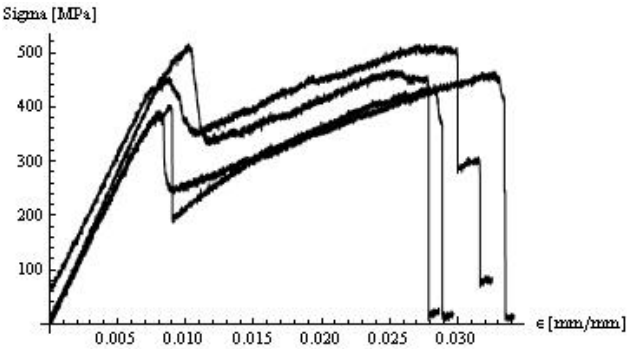
Absorbed Energy is determined by integrating Stress-Strain curve.

3. Results and discussion

Mechanical results for the repetitions of ENF tests are reported in Figures 4:



A: ENF mechanical results for virgin specimens



B: ENF mechanical results for nanomodifiedspecimens

Figure 4: ENF Mechanical Test Results

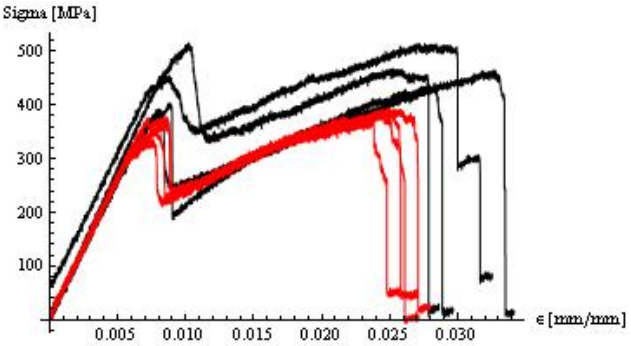
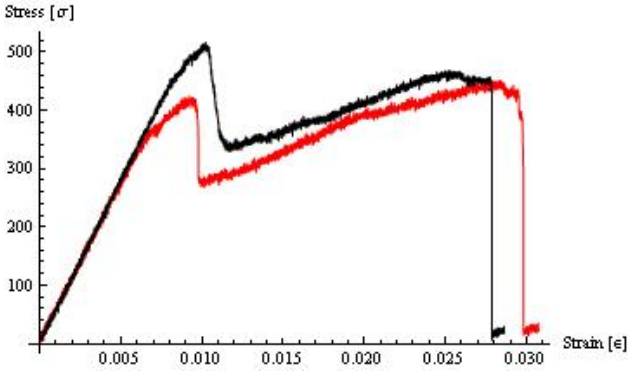


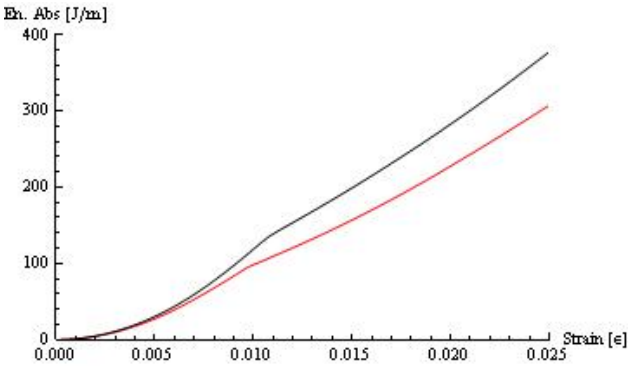
Figure 5: Mechanical Test Comparison

In the following Figure 5 comparison between the two conditions is presented.

Virgin specimens (figure 4-A) showed a very good repeatability of the stress-strain curves: initially a peak stress of around 360 Mpa is required in order to start the crack propagation then lower stresses are required to propagate the crack but with an increasing trend. Indeed at specimen break (0,026 in strain) the ultimate stress is comparable to peak stress. In nanomodified specimens a greater data scattering is evidenced (fig. 4-B and fig. 5): two conditions presented a peak stress value slightly higher than virgin ones and a very similar propagation trend. The others two conditions instead showed much greater peak values (440 and 500 MPa respectively) and also higher propagation stresses. It is also very important to note that in all the nanomodified specimens higher strain at fracture are also obtained thus being in correlation with a greater absorbed energy at break. In Figure 6 the mean curves are represented for both the virgin and nanomodified conditions:



A: Stress-Strain Curves



B: Energy Absorbed curves

Figure 6: Mean curves

In Table 3, maximum stress, flexural stiffness and mechanical energy resulting for both virgin and nanomodified configurations are presented. Energy Absorbed is normalized respect to the specimen's width. Symbols μ and σ indicate respectively the mean value and the standard deviation of all the performed tests.

From Figure 6 and from Table 3 it can be observed that the presence of nanofiber strongly influences the behavior of the laminates.

Table 3: ENF Mechanical Test Results summary

	σ_{max} [MPa]		
	μ	σ	Δ%
Virgin	393	2.45	+22.8
Nanomodified	483	27.7	

	E^{flex} [MPa]		
	μ	σ	Δ%
Virgin	48.1	0.56	+13.4
Nanomodified	54.5	4.1	

	Energy Absorbed [J/m]		
	μ	σ	Δ%
Virgin	258	19	+33.8
Nanomodified	345	47	

Flexural Modulus is increased by nanofibers, even if they are placed in the mid-plane of the laminate and the mat thickness does not significantly affect the total thickness of the laminate. In the mid-interface, maximum shear stress and crack tip are located: nanofibers tough the interface, and it directly influences flexural stiffness. Presence of the nanofibrous mat at the mid-interface contributes to increase both the maximum value of the stress and the energy absorbing capability. In particular, considering the curves in Figure 6(A) and the values in Table 3, it can be observed that the nanomodified specimen has a maximum stress value that is about 22.3% higher than the maximum value of the virgin specimen. Likewise it can be noted that in the case of the nanomodified laminate the strain value, when the maximum stress is reached, is higher than the strain value at the maximum stress of the virgin specimen. In summary, in the case of a end notched flexure laminate, the nanofibrous mat contributes to increase the maximum sustainable stress and, at the same time, it contributes to delay the first material crisis. Moreover, it can be observed that the general trend of the stress for the nanomodified specimen overcomes that of the virgin laminate. After the first stress drop it can be observed that the stress values of the nanomodified and of the virgin laminate increase following trends with a similar slope but different absolute values. In particular the stress values of the nanomodified laminate always overcome the stress values of the virgin laminate. Such a behavior can be also evaluated by the mechanical energy diagram reported in Figure 6(B). From this diagram it can be noted that at strain values higher than those at which the laminate crisis takes place, after the two trend break points, the mechanical energy stored into the nanomodified laminate overcome the energy stored into the virgin one. Mechanical energy data reported in Table 3 confirm these data: the maximum mechanical energy that can be stored by a nanomodified laminate is about 33.8% higher than in the case of the virgin laminate.

3 Conclusion

In the present work electrospun nanofibrous Nylon 6,6 random mats were used as interleaving in composite laminates. Specimens were composed of twenty plies of woven carbon fibers - epoxy resin prepregs. Nanofibrous mat was placed in the mid-plane of beam shaped specimens that were tested under Mode II of fracture investigated by means of End Notched Flexure tests. Both virgin and nanomodified laminates were considered. The results analysis and discussion can be summarized as follow:

- flexural modulus is increased by nanofibrous mats in the particular case investigated, i.e. when only one mat is placed at the mid-interface of specimens;
- nanofibrous mat enhances the maximum value of the force that has to be applied to initiate the longitudinal crack propagation;
- nanomodified composite laminates, when subject to the interlaminar crack propagation, can absorb an higher amount of mechanical energy with respect to the virgin ones.

ENF results demonstrate that the interleaved nanofibrous mat can enhance composite delamination strength and damage tolerance.
Obtained results indicate that several engineering areas, where composite laminate are applied, can benefit from the usage of electrospun nanofibrous mats, in particular in

the case of structural components. Due to the reduced damage progression sensitivity under Mode II of fracture it is expected that nanofibrous mats can enhance reliability of structural components used in aeronautic vehicles.

References

[1] Ishai O Banks-Sills L Sela N. "The effect of adhesive thickness on interlaminar fracture toughness of interleaved CFRP specimens." *Composites*, 20(3):257–264, 1989.

[2] Ishai O Banks-Sills L Sela N. "Toughness improved high performance torayca prepreg T800H/3900 series." *33rd Int. SAMPE Symp, Society for the Advanceent of Material and process Engeneering*, pages 272–283, 1988.

[3] de Morais A B Marques A T de Castro P T Pereira A.B. "Mode II interlaminar fracture of carbon/epoxy multidirectional laminates." *Composites Science and Technology*, 64(10-11):1653–1659, 2004.

[4] A. M. Pereira. "Mode II interlaminar fracture of carbon/epoxy multidirectional laminates." *Composites Science and Technology*, 64(10-11):1653–1659, August 2004.

[5] Pipes N.J., PaganoR.B. "Influence of Stacking sequence on laminate strength." *Journal of Composite Materials*, 5:50–57, 1971.

[6] Chen J.-W. Tsai G.-C." Effect of stitching on Mode I strain energy release rate." *Composite Structures*, 69(1):1–9, 2005.

[7] Mignery L.A., Tan T. M., and Sun C.T. ASTM Special Technical Publication, pages 371–385, 1985.

[8] M B Dow and H B Dexter. "Development of Stitched, Braided and Woven Composite Structures" in the ACT Program and at Langley Research Center (1985 to 1997) Summary and Bibliography. Technical report, 1997.

[9] W E Howard, T Gossard, and M J Robert. "Reinforcement of composite laminate free edges with U-Shaped caps." *AIAA Paper*, volume 1, 1986.

[10] Ochoa O O Chan W.S. "Edge delamination resistance by a critical ply termination." *Mechanical and corrosion properties. Series A, Key engineering materials*, 37:285–303, 1989.

[11] Y.A. Dzenis and D.H. Reneker. United States Patent 6265333: "Delamination resistant composites prepared by small diameter fiber reinforcement at ply interfaces", 2001.

[12] Reneker D H Kim J.-S. "Mechanical properties of composites using ultrafine electrospun fibers." *Polymer Composites*, 20(1):124–131, 1999.

[13] Y. A. Dzenis. "Structural Nanocomposites." *Material Science*, 319(January):419–420, 2008.

[14] K. Shivakumar, L. Shivalingappa, Huanchun Chen, Paul Akangah, Gowthaman Swaminathan, and Larry Russell. "Polymer Nanofabric Interleaved Composite Laminates." *AIAA Journal*, 47(7):1723–1729, July 2009.

[15] S Sihn, R Kim, W Huh, K Lee, and a Roy. "Improvement of damage resistance in laminated composites with electrospun nano-interlayers." *Composites Science and Technology*, 68(3-4):673–683, March 2008.

[16] S. Sihn, Y.R. Kim, W Huh, K-H Lee, and A.K. Roy. "Improvement of damage resistance in laminated composites with electrospun nano-interlayers." *Composites Science and Technology*, 68(3-4):673–683, 2008.

[17] L Liu, Z Huang, C He, and X Han. "Mechanical performance of laminated composites incorporated with nanofibrous membranes." *Materials Science and Engineering: A*, 435-436:309–317, November 2006.

[18] J. Zhang, T. Lin, and X. Wang. "Electrospun nanofibre toughened carbon/epoxy composites: Effects of polyetherketone cardo (PEK-C) nanofibre diameter and interlayer thickness." *Composites Science and Technology*, 70(11):1660–1666, October 2010.

[19] S Lin, Q Cai, J Ji, G Sui, Y Yu, X Yang, Q Ma, YWei, and X Deng. "Electrospun nanofiber reinforced and toughened composites through in situ nanointerface formation." *Composites Science and Technology*, 68(15-16):3322-3329, December 2008.

[20] A. Zucchelli, M.L. Focarete, C. Gualandi, and S. Ramakrishna. "Electrospun nanofibers for enhancing structural performance of composite materials." *Polymers for Advanced Technologies*, 22(3):339–349, March 2011.

[21] S. Ramakrishna. An introduction to Electrospinning and Nanofibers

[22] A. Greiner and J. H. Wendorff. "Electrospinning: a fascinating method for the preparation of ultrathin fibers". *Angewandte Chemie* (International ed. in English), 46(30):5670–703, January 2007.

[23] A. Baji, Y-W Mai, S-C Wong, M. Abtahi, and P. Chen. "Electrospinning of polymer nanofibers: Effects on oriented morphology, structures and tensile properties." *Composites Science and Technology*, 70(5):703–718, May 2010.

[24] S. Koombhongse, W. Liu, and D.H. Reneker. "Flat Polymer Ribbons and Other Shapes by Electrospinning". Journal of Polymer Science. Part B, 39:2598–2606, 2001.

[25] S-C Moon and R.J. Farris. "The Morphology , Mechanical Properties , and Flammability of Aligned Electrospun Polycarbonate (PC) Nanofibers Preparation of Polymer Solution."

Polymer Engineering and Science, pages 1848–1854, 2008.456
[26] ASTM D7246/D. “Standard Test Method for Flexural Properties of Polymer Matrix Composite Materials”. Annual Book of ASTM

Infrared Thermography in Aeronautics

Giovanni Maria Carlomagno
Department of Aerospace Engineering – University of Naples Federico II

Keywords: *Infrared Thermography*

Abstract

Infrared Thermography (IRT) is a methodology which allows for: remote detection of thermal energy that is radiated from objects in the infrared band, conversion of such energy into a video signal, and representation of the object surface temperature map. This methodology constitutes a great potentiality to be exploited in aeronautics, in a lot of application fields and for several different purposes.

In fact, an infrared imaging system is the right solution for maintenance purposes and non-destructive evaluation (NDE) of materials, which are practically based of heat transfer by conduction. On the other side, IRT has also proved its suitability for measurements of convective heat fluxes in complex fluid flows and for investigation, through the Reynolds analogy, of the flow field behaviour also over complicated body shapes. This latter is for the comprehension of fluid dynamics phenomena such as flow instability including transition from laminar to turbulent flow, flow separation and reattachment.

The scope of this presentation is to review some of the most relevant experimental results obtained by the author, and the research team he belongs to, over the past several years, in the study of the application of infrared thermography to both thermofluid-dynamic purposes and non-destructive testing.

The thermo-fluid-dynamic aspect will cover several complex fluid flows, which range:

- *from natural convection, up to hypersonic flows;*
- *from still, to moving channels;*
- *from steady, to unsteady flows.*

Within this context, the attention will be particularly focused on the comprehension of the complex phenomena which may arise when considering different fluid flow configurations involving thermofluid-dynamic aspects, with variation of Reynolds and Mach numbers, flows which strike on a surface, jets in a cross-flow, aerodynamic bodies at an angle of attack, or flows confined inside a duct such as those for turbine blades cooling.

Instead, the non-destructive testing aspect will analyze:

- *several different materials (metals, plastics, composites, hybrid composites)*

- *many structures and components (sandwich panels, honeycombs, turbine blades)*
- *different types of anomalies (slag inclusions, voids, impact damage, delamination).*

This part of the presentation will be devoted to the temperature mapping over a solid body and the related useful information which such a map may supply referring to different NDE application fields. One main field of applications relates to ascertain materials conditions, within which the use of an infrared camera is of remarkable value and for certain problems represents the unique solution. In fact, an infrared imaging device allows for a fast and precise acquisition of the temperature map, so it consents to discovering buried anomalies in the material. As a matter of fact, nowadays, infrared thermography is gaining ever more importance amongst NDE techniques.

CEAS 2011

Advancements in Fluid-Dynamics 1

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Kinetic Dynamics for Reactive Wave Propagation Problems

Gilberto M. Kremer
Departamento de Física, Universidade Federal do Paraná, Curitiba, Brazil

Miriam Pandolfi Bianchi
Dipartimento di Matematica, Politecnico di Torino, Torino, Italy

Ana Jacinta Soares
Departamento de Matemática e Aplicações, Universidade do Minho, Braga, Portugal

Keywords: Boltzmann equation, Chemically reactive flows, Wave propagation, Transport processes

Abstract

The dynamics of different wave problems, as the steady detonation, the acoustic perturbation and the light scattering, can be satisfactorily described by means of pertinent models of the Boltzmann equation. The present study deals with the extension of the exact and approximate models of the Boltzmann equation to a gas mixture of four constituents undergoing a reversible bimolecular reaction. Such models constitute the starting point for a consistent derivation of reactive Euler equations and reactive equations with Navier-Stokes, Fourier and Fick constitutive equations. Aim of this paper is to enlighten the mathematical aspects of the methodology which allows to formulate and solve the above said wave propagation problems on the basis of a shared kinetic framework. At the same time, some relevant results recently obtained by the authors in previous papers for the mentioned wave problems are also reviewed.

1 Introduction

Physical processes where gas phase chemical reactions occur, are described at various levels in literature, as documented in the books [1, 2, 3]. The investigation of chemically reacting flows is in fact fundamental in order to enlarge the knowledge of significant fields ranging from plasma chemistry to atmosphere physics at high altitudes, from ionization processes accompanying the reentry of hypersonic vehicles to chemical technology, and several others.

In particular, the kinetic study of non-equilibrium effects and their influence on the main aspects of reactive gas systems constitute the object of a large production concerned with the chemical kinetics Boltzmann equation (BE).

A chemical reaction occurring in a gas causes a perturbation of the local equilibrium, disturbing the molecular velocity distribution from its Maxwellian form and the reaction rate from its equilibrium value, whereas elastic collisions contribute to restore the equilibrium. These deviations appear to be more relevant at higher atmosphere altitudes where the elastic collisions are not sufficient to sustain the equilibrium of the gas flow, as discussed in Ref. [4]. A departure from the Maxwellian velocity distribution, due to the proceeding of the reaction itself, is in general responsible of the reaction rate decrease and may also induce qualitative changes of the

system properties.

The non-equilibrium effects arising in a dilute gas system whose constituents undergo a bimolecular chemical reaction, were recognized in the pioneering study by Prigogine and coworkers [5, 6] for the simple reaction $A_1 + A_1 \rightarrow A_2 + A_3$. They first generalized the Chapman-Enskog solution of the BE to the case of reacting gases with slow chemical reactions, for which the chemical relaxation time is longer than the mechanical one, and showed that the rate coefficient decreases from early stages of the reaction for which the role of products can be neglected.

As known, see the book [7], the most relevant model which describes the evolution of a rarefied gas is the integro-differential BE. The integral form of the collision operator represents a mathematical complexity for both theoretical studies and numerical simulations. On the other hand, the Bhatnagar Gross and Krook (BGK)--type models, replacing the integral operator by a relaxation linear term which retains the main properties of the exact BE operator [3], can be more easily handled.

The present study deals with the extension of both exact BE and approximate BGK modeling to a gas mixture of four constituents undergoing a reversible bimolecular reaction. The former is assumed to derive the Euler equations in Subsection 3.1, in view of the steady detonation wave problem of Section 4. The latter is adopted to deduce the reactive equations with Navier-Stokes, Fourier and Fick constitutive equations and transport properties in Subsections 3.2 and 3.3, in view of both the sound wave propagation of Section 5 and the light scattering problem of Section 6. On the other hand, the fundamental mathematical features of these modellings are summarized in the following Section 2.

2 Kinetic modelling

The main aspects of the exact BE and approximate BGK kinetic modellings, at microscopic level, are here briefly recalled in order to state the hydrodynamic picture, at the macroscopic level, which constitutes the common basis for the wave problems of Sections 4, 5 and 6.

The present modellings refer to a gas mixture of four constituents with a reversible reaction of type $A_1+A_2 \rightleftharpoons A_3+A_4$. For each constituent A_α, $\alpha = 1, 2, 3, 4$, let m_α denote the molecular mass, with $m_1 + m_2 = m_3 + m_4$, ε_α the formation energy of a molecule and $f_\alpha(t, \boldsymbol{x}, \boldsymbol{c})$ the one-particle velocity distribution function.

At the microscopic scale, in the phase space, the evolution of the mixture is described by the kinetic equations

$$\frac{\partial f_\alpha}{\partial t} + c_i^\alpha \frac{\partial f_\alpha}{\partial x_i} = \sum_{\beta=1}^{4} \mathcal{Q}_{\alpha\beta}^E + \mathcal{Q}_\alpha^R, \qquad (1)$$

where $\mathcal{Q}_{\alpha\beta}^E$ and $\mathcal{Q}_{\alpha\beta}^R$ are elastic and reactive collision operators. Their explicit form is of integral type when the exact BE modelling is assumed, whereas it is homogeneous of degree one when the approximate BGK modelling is considered. The detailed expressions of the exact BE collision operators are well known in literature, see for example [7] and [3]. Moreover the considered approximate BGK collision operators have been derived in paper [8] by imposing that the production terms are the same in both BE and BGK models. Consequently, both modellings share the same conservation laws of mass, momentum and total energy of the mixture, and the common macroscopic picture, which is relevant for the wave propagation problems considered here, is provided by the hydrodynamic system

$$\frac{\partial n_\alpha}{\partial t} + \nabla\cdot(n_\alpha \boldsymbol{u}_\alpha) = \tau_\alpha, \ \alpha = 1, \ldots, 4, \qquad (2a)$$

$$\frac{\partial}{\partial t}(\varrho \boldsymbol{v}) + \nabla\cdot(\varrho \boldsymbol{v} \otimes \boldsymbol{v} + \mathbb{P}) = 0, \qquad (2b)$$

$$\frac{\partial}{\partial t}\Big(\frac{3}{2} n k_B T + \sum_{\alpha=1}^{4} n_\alpha E_\alpha + \frac{1}{2}\varrho \boldsymbol{v}^2\Big) + \nabla\cdot\Big[\boldsymbol{q} + \mathbb{P}\boldsymbol{v} + \Big(\frac{3}{2} n k_B T + \sum_{\alpha=1}^{4} n_\alpha E_\alpha + \frac{1}{2}\varrho \boldsymbol{v}^2\Big)\boldsymbol{v}\Big] = 0. \qquad (2c)$$

In the above system, k_B is the Bolzmann constant, the main independent macroscopic fields are the constituent number densities n_α, mean velocity $\boldsymbol{v}$ and temperature T of the whole mixture, with $\varrho = \sum_{\alpha=1}^{4} m_\alpha n_\alpha$. Moreover τ_α and $\boldsymbol{u}_\alpha$ are the constituent reaction rate and diffusion velocity, $\mathbb{P}$ and $\boldsymbol{q}$ are the mixture pressure tensor and heat flux vector, whose kinetic defi-

nitions are

$$\tau_\alpha = \int_{\mathbb{R}^3} \mathcal{Q}_\alpha^R(f_\alpha)\, d\mathbf{c}_\alpha, \tag{3}$$

$$u_i^\alpha = \frac{1}{\varrho_\alpha}\int_{\mathbb{R}^3} m_\alpha \xi_i^\alpha f_\alpha\, d\mathbf{c}_\alpha, \tag{4}$$

$$\mathbb{P}_{ij} = \sum_{\alpha=1}^{4}\int_{\mathbb{R}^3} m_\alpha \xi_i^\alpha \xi_j^\alpha f_\alpha d\mathbf{c}_\alpha, \tag{5}$$

$$q_i = \sum_{\alpha=1}^{4}\left(\int_{\mathbb{R}^3} \frac{1}{2} m_\alpha \xi_\alpha^2 \xi_i^\alpha f_\alpha d\mathbf{c}_\alpha + n_\alpha \varepsilon_\alpha u_i^\alpha\right), \tag{6}$$

where $\xi^\alpha = \boldsymbol{c}^\alpha - \boldsymbol{v}$ is the peculiar velocity of the α-constituent. The presence of these last quantities renders the system non-closed so that a closure procedure must be applied for a specified flow regime.

3 Hydrodynamic closures

The closure of the hydrodynamic system ($2a$-$2c$) is achieved once τ_α and the further fields $\boldsymbol{u}_\alpha$, $\mathbb{P}$ and $\boldsymbol{q}$ are expressed in dependence on the main macroscopic fields.

The Chapman-Enskog methodology is used in a chemical regime close to equilibrium for which the elastic and reactive frequencies are of the same order. An approximate solution of the kinetic Eqs. (1) is determined as an expansion around the Maxwellian equilibrium distribution

$$f_\alpha^{(0)} = n_\alpha\left(\frac{m_\alpha}{2\pi kT}\right)^{\frac{3}{2}} \exp\left(-\frac{m_\alpha (c_i^\alpha - v_i)^2}{2kT}\right) \tag{7}$$

of the form

$$\widetilde{f}_\alpha = f_\alpha^{(0)}\left[1 + \phi_\alpha\right], \tag{8}$$

where ϕ_α represents the deviation from the mechanical equilibrium and turns out to be explicitly depending on the main macroscopic fields ϱ_α, $\boldsymbol{v}$ and T. Thus the approximation $\widetilde{f}_\alpha$ in turn depends on such main fields and is then introduced in the kinetic definitions (3-6). Such dependence is strictly related to the choice of the kinetic modelling, which can be either the exact BE or the approximate BGK equations, together with pertinent assumptions of elastic and reactive cross sections. Consequently, the structure of the elastic and reactive collision operators $\mathcal{Q}_{\alpha\beta}^E$ and $\mathcal{Q}_{\alpha\beta}^R$ is specified.

At each level of the Chapman-Enskog procedure, the constitutive laws for τ_α, $\boldsymbol{u}_\alpha$, $\mathbb{P}$ and $\boldsymbol{q}$ result from the actual computation of the integrals in definitions (3-6), once f_α is replaced by its corresponding approximation $\widetilde{f}_\alpha$.

3.1 Reactive Euler equations

At the first-order of the Chapman-Enskog method, the approximate solution $\widetilde{f}_\alpha$ is the Maxwellian $f_\alpha^{(0)}$ which includes the non-equilibrium effects induced by the chemical reaction only, when the number densities n_α are not constrained by the mass action law.

The corresponding constitutive laws are the reaction rate law

$$\tau_\alpha = \int_{\mathbb{R}^3} \mathcal{Q}_\alpha^R(f_\alpha^{(0)})\, d\mathbf{c}_\alpha \tag{9}$$

with conditions

$$\tau_1 = \tau_2 = -\tau_3 = -\tau_4, \tag{10}$$

which assure the correct chemical exchanges predicted by the reaction, and the further laws for $\boldsymbol{u}_\alpha$, $\mathbb{P}$ and $\boldsymbol{q}$

$$\boldsymbol{u}_\alpha = 0, \quad \mathbb{P} = pI, \quad \boldsymbol{q} = 0, \tag{11}$$

where I represents the identity matrix and p the mixture pressure.

The hydrodynamic system closed at this level consists in the reactive equations given by ($2a$-$2c$) together with the constitutive laws (9-11.)

Such system, which is appropriate when the transport effects are absent, is the mathematical tool for the kinetic description of the steady detonation wave and its linear stability.

3.2 Reactive Navier-Stokes equations

At the second-order of the expansion procedure, the approximate solution $\widetilde{f}_\alpha$ in the form (8) is characterized in terms of the deviation ϕ_α containing the effects of the transport fluxes. The corresponding constitutive laws are then the reaction rate law

$$\tau_\alpha = \int_{\mathbb{R}^3} \mathcal{Q}_\alpha^R(f_\alpha^{(0)}\left[1 + \phi_\alpha\right])\, d\mathbf{c}_\alpha \tag{12}$$

and the generalized Fick, Navier-Stokes and

Fourier laws, respectively given by

$$u_i^\alpha = \frac{1}{\varrho_\alpha}\int_{\mathbb{R}^3} m_\alpha \xi_i^\alpha f_\alpha^{(0)} \phi_\alpha \, d\mathbf{c}_\alpha, \qquad (13)$$

$$\mathbb{P}_{ij} = p\,\delta_{ij} + \sum_{\alpha=1}^{4}\int_{\mathbb{R}^3} m_\alpha \xi_i^\alpha \xi_j^\alpha f_\alpha^{(0)} \phi_\alpha d\mathbf{c}_\alpha, \qquad (14)$$

$$q_i = \sum_{\alpha=1}^{4}\Bigg(\int_{\mathbb{R}^3}\frac{1}{2} m_\alpha \xi_\alpha^2 \xi_i^\alpha f_\alpha^{(0)} \phi_\alpha d\mathbf{c}_\alpha + \varepsilon_\alpha \int_{\mathbb{R}^3} \xi_i^\alpha f_\alpha^{(0)}[1+\phi_\alpha] d\mathbf{c}_\alpha\Bigg). \qquad (15)$$

The hydrodynamic system closed at this level consists then in the reactive equations given by (2*a*-2*c*), together with the reaction rate (12) and Navier-Stokes, Fourier and Fick constitutive equations (13-15).

In particular, when the kinetic modelling is chosen, the deviation ϕ_α is explicitly determined and the constitutive laws (13-15) permit to detail the transport picture through the actual computation of the transport coefficients.

Thus, the system closed at this level is suitable for a flow regime where the transport properties are relevant. They constitute the basis for the kinetic description of the sound wave propagation and light scattering spectra.

3.3 Transport properties

The transport picture is here represented starting from the kinetic modelling of paper [8], which is based on the BGK collision operators, hard-spheres elastic cross sections and Present's reactive cross sections. Therefore, the generalized Fick, Navier-Stokes and Fourier laws are deduced respectively in the form

$$d_i^\alpha = -\sum_{\beta=1}^{4}\frac{x_\alpha^{\rm eq} x_\beta^{\rm eq}}{D_{\alpha\beta}}\left(u_i^\alpha - u_i^\beta\right), \qquad (16)$$

$$\mathbb{P}_{ij} = p\delta_{ij} - \mu\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\frac{\partial v_r}{\partial x_r}\delta_{ij}\right), \qquad (17)$$

$$q_i = -\lambda\frac{\partial T}{\partial x_i} + \sum_{\alpha=1}^{4}\left(\frac{5}{2}kT + \varepsilon_\alpha\right) n_\alpha^{\rm eq} u_i^\alpha, \qquad (18)$$

where $x_\alpha^{\rm eq} = n_\alpha^{\rm eq}/n^{\rm eq}$ denote equilibrium molar fractions, d_i^α the constituent diffusion forces and $D_{\alpha\beta}$, μ and λ are the diffusion, shear viscosity and thermal conductivity coefficients. They turn out to be known functions of T, n_α and both elastic and reactive collision frequencies.

As an application, the behavior of the transport coefficients and the influence of the chemical reaction are investigated for the elementary reaction H_2+ Cl $\rightleftharpoons$ HCl+H. For brevity only some diffusion coefficients are reported here, the others being quite similar.

The results of Fig. 1 show an increasing behavior with T which is more pronounced for the inert mixture.

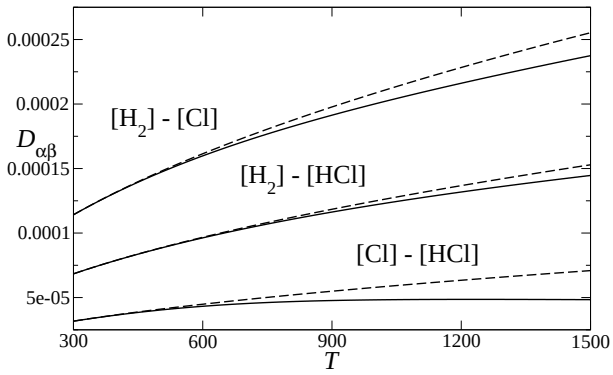


Figure 1: Diffusion coefficients $D_{\alpha\beta}$ as function of T. Dashed line: inert mixture; solid line: reactive mixture.

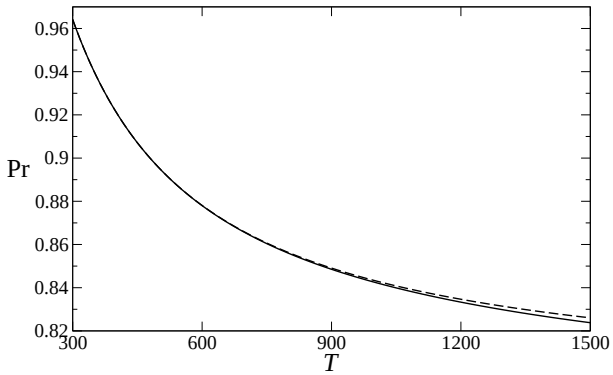


Figure 2: Prandtl number as function of T. Dashed line: inert mixture; solid line: reactive mixture.

Moreover the Prandtl number of the considered mixture, whose kinetic definition is given by

$$\mathrm{Pr}=\frac{5\mu}{2\lambda}\left[\frac{4k(n_1^{\mathrm{eq}}+n_2^{\mathrm{eq}}+n_3^{\mathrm{eq}}+n_4^{\mathrm{eq}})}{m_1n_1^{\mathrm{eq}}+m_2n_2^{\mathrm{eq}}+m_3n_3^{\mathrm{eq}}+m_4n_4^{\mathrm{eq}}}\right], \quad (19)$$

is represented in Fig. 2.

The results show a decreasing behavior with T, as expected, and an acceptable range of values of Pr, lying inside the difference between the BGK approximation and the exact BE evaluation for the Prandtl number of a single gas.

4 Steady detonation waves

In this section the kinetic dynamics of a planar detonation wave arising in an explosive gas is studied for an Eulerian reactive mixture, according to Subsection 3.1, in a flow regime for which mechanical equilibrium and strong chemical disequilibrium hold, and diffusion, shear viscosity and heat flux are absent.

The kinetic modelling is based on the exact BE collision operators, and on a particular choice of reactive cross sections with threshold energy, which satisfies the micro reversibility principle and extends the Present model, see Refs. [3] and [9]. The hydrodynamic system is formed by Eqs. (2*a*-2*c*) reduced to one space dimension and closed with the constitutive laws (11) and rate law (9) in the form

$$\tau_\alpha=\left[-n_1n_2\left(\frac{m_3m_4}{m_1m_2}\right)^{3/2}\exp\left(-\frac{E}{k_BT}\right)+n_3n_4\right]S_\alpha, \quad (20)$$

where S_α is a weight function depending on temperature and threshold velocity.

With reference to the classical ZND (Zeldovich, Von Neumann and Doering) structure of a detonation wave [10], a non-reactive shock wave propagating in a quiescient mixture is followed by a finite reaction zone sustained by the energy release of an exothermic chemical reaction. The Rankine-Hugoniot (RH) conditions connect the fluxes of the main fields between the initial state, ahead of the shock, and an arbitrary von Neumann (VN), intermediate (I) or final state (F) behind it. Figure 3 shows a detonation wave profile for the total number density versus the wave thickness, for an elementary reaction of the explosive hydrogen-oxygen chain [11].

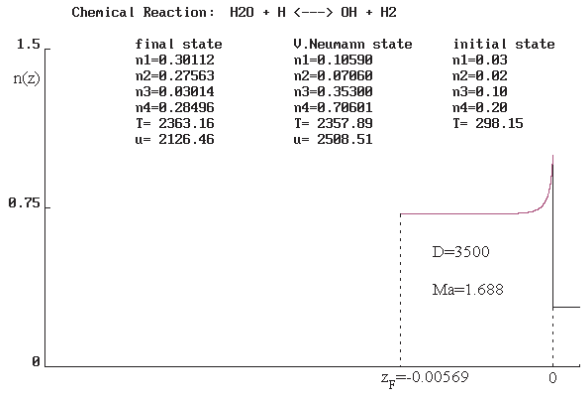


Figure 3: Detonation wave profile of number density n versus the wave thickness, for Mach number $Ma = 1.6888$ and detonation velocity $\mathcal{D} = 3500$.

The linear stability of the steady detonation wave can also be studied, analysing the response of the steady solution to small rear boundary perturbations, which induce a distortion $\psi(t)$ in the position of the planar shock wave. Such distortion may either decay with time or grow larger affecting the steady character of the detonation wave solution.

The mathematical approach of the linear stability problem requires the linearization of the reactive Euler equations and RH conditions through a normal mode expansion of the main fields around the steady solution, in the form

$$\begin{aligned}\boldsymbol{z}(x,t)&=\boldsymbol{z}^*(x)+\exp(at)\,\overline{\boldsymbol{z}}(x), \quad (21)\\ \psi(t)\ &=\overline{\psi}\exp(at),\quad a,\overline{\psi}\in\mathbb{C},\quad a=\alpha+i\beta,\end{aligned}$$

where $\boldsymbol{z}(x,t)$ represents the vector of the main fields, $\boldsymbol{z}^*(x)$ steady detonation solution, $\overline{\boldsymbol{z}}(x)$ unknown space disturbances, $\overline{\psi}$ complex amplitude of the shock distortion, α and β perturbation growth rate and frequency.

Consequently, the sign of α determines the stability behaviour of the steady solution $\boldsymbol{z}^*(x)$. The resulting stability equations with related RH initial conditions provide the evolution of the disturbances $\overline{\boldsymbol{z}}(x)$ in the reaction zone.

As discussed in paper [12], Fig. 4 shows a typical instability behaviour of the steady solution since the perturbation increases in time, whereas the further Fig. 5 describes a typical stability behaviour.

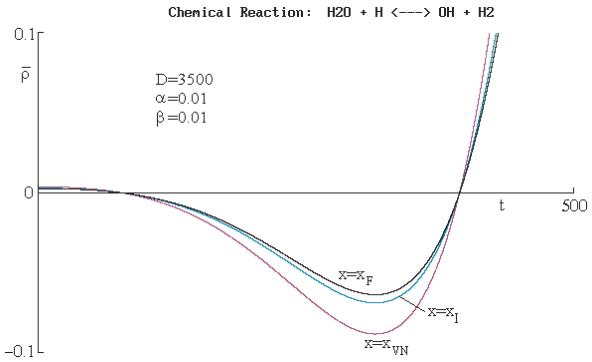


Figure 4: Instability behaviour of the mass density perturbation versus time at different states, for growth rate $\alpha = 0.01$ and disturbance frequency $\beta = 0.01$.

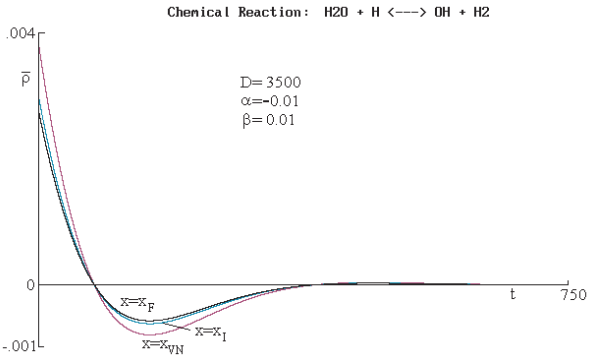


Figure 5: Stability behaviour of the mass density perturbation versus time at different states, for growth rate $\alpha = -0.01$ and disturbance frequency $\beta = 0.01$.

5 Sound wave propagation

In this section the propagation of sound waves in a reactive mixture where the transport effects are relevant is studied, according to Subsections 3.2 and 3.3, by a normal mode analysis at the Navier-Stokes, Fourier and Fick level. The kinetic modelling is based on the approximate BGK collision operators, and Present's reactive cross sections. The hydrodynamic system is formed by Eqs. (2*a*-2*c*) reduced to one space dimension and closed with the constitutive laws (16-18) and rate law in the form

$$\tau_\alpha = \nu_\alpha n_\alpha n_\gamma k_\sigma^{(0)} \frac{A}{\imath \mathcal{T}}, \qquad (22)$$

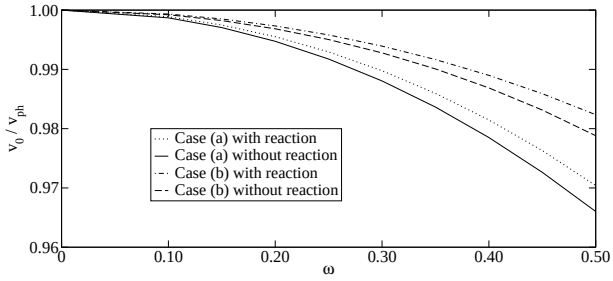


Figure 6: Phase velocity for two mixtures of the H_2-Cl system.

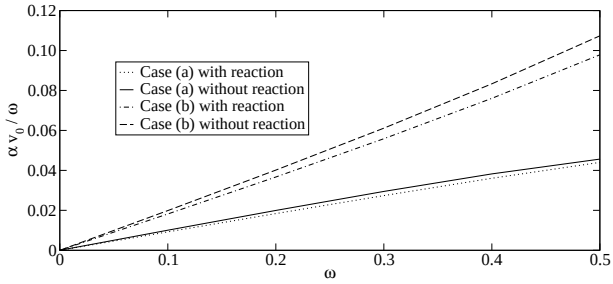


Figure 7: Attenuation coefficient for two mixtures of the H_2-Cl system.

where $k_\sigma^{(0)}$ is the first approximation to the forward ($\sigma = 1$) and backward ($\sigma = -1$) rate constants [8], $\mathcal{A}$ is the affinity of the forward chemical reaction, and $\gamma = 2, 1, 4, 3$ for $\alpha = 1, 2, 3, 4$, respectively. The kinetic dynamics of the longitudinal sound waves is studied assuming small complex amplitudes and exponential factors depending on the complex wave number κ and real wave angular frequency ω,

$$\begin{aligned} \widetilde{n}_\alpha &= n_\alpha^{(\mathrm{eq})} + \overline{n}_\alpha \exp[\imath(\kappa x - \omega t)], \\ \widetilde{v} &= \overline{v} \exp[\imath(\kappa x - \omega t)], \\ \widetilde{T} &= T_0 + \overline{T} \exp[\imath(\kappa x - \omega t)]. \end{aligned} \qquad (23)$$

The behavior of the wave phase velocity $v_{ph} = \omega/\mathrm{Re}\,\kappa$ and attenuation coefficient $\alpha = \mathrm{Im}\,\kappa$ is shown in Figs. 6 and 7, for two mixtures of the Hydrogen-Chlorine system, where the elementary bimolecular chemical reaction H_2+Cl $\rightleftharpoons$ HCl+H takes place, in both reactive and inert conditions. Two cases are considered, namely *case* (a) $x_1^{\mathrm{eq}} = 0.1$, $x_2^{\mathrm{eq}} = 0.618$, $x_3^{\mathrm{eq}} = 0.082$, $x_4^{\mathrm{eq}} = 0.2$ and *case* (b) $x_1^{\mathrm{eq}} = 0.2$, $x_2^{\mathrm{eq}} = 0.424$, $x_3^{\mathrm{eq}} = 0.076$, $x_4^{\mathrm{eq}} = 0.3$.

6 Light scattering spectra

The spectrum of the light scattered by the reactive mixture at thermodynamic equilibrium returns a wide information on non-equilibrium thermodynamics and transport properties of the gas. In particular, a relevant aspect of such a phenomenon is the decay of the thermal density fluctuations, which is mainly ruled by the transport coefficients, see Ref. [13]

In this section the dynamics of light-scattering phenomenon is studied starting from the hydrodynamic picture of Subsections 3.2 and 3.3, in order to describe the response of the reactive mixture in thermodynamical equilibrium to the incident light field.

When a monochromatic linearly polarized plane light wave impinges upon a gas mixture, the intensity of the scattered light is related to the so called dynamic structure factor $S(\mathbf{q},\omega)$, $\mathbf{q}$ and ω being the wave vector and shift in the angular frequency.

As known, see Ref. [13], the intensity of the light scattered by the mixture is classically evaluated by means of the time correlation functions of the constituent number density fluctuations due to the interaction of the light field with the gas. The decay of such fluctuations is mainly ruled by the transport coefficients of the considered mixture and is strictly depending on the dynamic structure factor $S(\mathbf{q},\omega)$. This factor constitutes the main mathematical tool within the light-scattering dynamics. It is defined in terms of the Laplace-Fourier transforms of the cross and auto correlation functions of the constituent number density fluctuations, namely

$$S(\mathbf{q},\omega)=\sum_{\alpha,\gamma=1}^{4} a_\alpha a_\gamma \mathrm{Re}\big[\langle \delta n_\alpha(\mathbf{q},i\omega)\delta n_\gamma(\mathbf{q},0)\rangle\big],$$

Above, a_α, $\alpha=1,\dots,4$, denote the atomic polarizability of the α-constituent.

The factor $S(\mathbf{q},\omega)$ can be expressed in an equivalent form, in terms of the macroscopic perturbations $\widetilde{n}_\alpha$ of the number densities, namely

$$S(\mathbf{q},\omega)=\sum_{\alpha,\gamma=1}^{4} a_\alpha a_\gamma \mathrm{Re}\left[\langle \widetilde{n}_\alpha(\mathbf{q},i\omega)\,\widetilde{n}_\gamma(\mathbf{q},0)\rangle\right]. \quad (24)$$

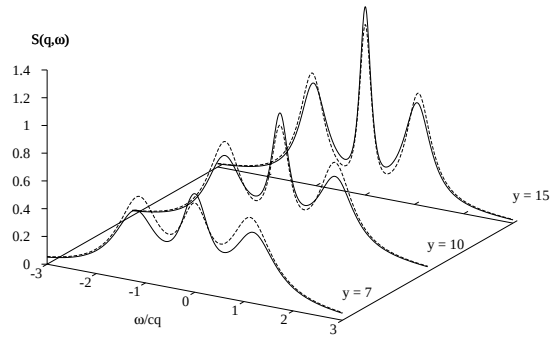


Figure 8: Case (a). Dynamic structure factor versus the reduced angular frequency, for different values of the uniformity parameter y. Reactive mixture (solid line) and inert mixture (dashed line).

In this equation the presence of $\widetilde{n}_\alpha$ and $\widetilde{n}_\gamma$ in place of their corresponding fluctuations, is justified thanks to the Onsager reciprocity regression hypothesis, which states that the fluctuations δn_α regress to equilibrium according to the same hydrodynamic equations which govern the relaxation to equilibrium of the macroscopic perturbations $\widetilde{n}_\alpha$. Therefore, the factor S can be explicitly computed from the hydrodynamic Eqs. ($2a$-$2c$) closed with the constitutive laws (16-18) and rate law (22). More specifically, the hydrodynamic system and the closure conditions must be first linearized around a thermodynamical equilibrium state, namely

$$n_\alpha = n_\alpha^{\mathrm{eq}} + \widetilde{n}_\alpha, \quad v_i = \widetilde{v}_i, \quad T = T_0 + \widetilde{T}, \quad (25)$$

and successively Laplace-Fourier transformed. In Eqs. (8), n_α^{eq} and T_0 represent constant number densities and mixture temperature of an equilibrium state with vanishing mean velocity, whereas $\widetilde{n}_\alpha$, $\widetilde{v}$ and $\widetilde{T}$ are small unknown perturbations of the corresponding main fields.

The procedure finally consists in solving the resulting algebraic system for the perturbations $\widetilde{n}_\alpha$, which involves all the transport coefficients, and performing ensemble average over the initial state in order to compute the dynamic structure factor through Eq. (25).

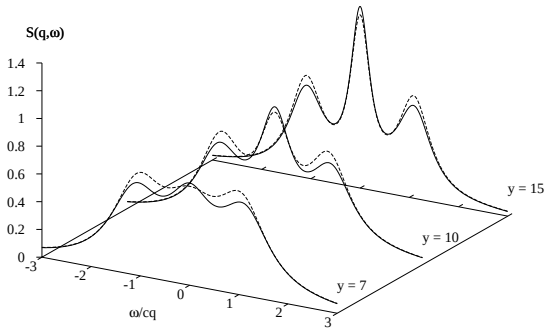


Figure 9: Case (b). Dynamic structure factor versus the reduced angular frequency, for different values of the uniformity parameter y. Reactive mixture (solid line) and inert mixture (dashed line).

Some numerical simulations have been performed for two different mixtures of the Hydrogen-Chlorine system where the chemical reaction H_2+Cl $\rightleftharpoons$ HCl+H occurs. The light scattering spectrum, through the Rayleigh and Brillouin peaks, is represented in Figs. 8 and 9 as function of the reduced angular frequency, for different values of the uniformity parameter y. See the book [13] for the definition of the parameters. Two cases are analyzed, namely *case* (a) $x_1^{\rm eq} = x_2^{\rm eq} = 0.33$, $x_3^{\rm eq} = x_4^{\rm eq} = 0.17$ and *case* (b) $x_1^{\rm eq} = 0.1$, $x_2^{\rm eq} = 0.618$, $x_3^{\rm eq} = 0.082$, $x_4^{\rm eq} = 0.2$.

7 Conclusions

The methodology applied to formulate and solve the different wave problems, as steady detonation wave, sound waves propagation and light scattering which have been treated by the authors in Refs. [8, 12, 14] respectively, are here focused at the end of outlining the common features of the kinetic approaches which are based on either the exact or the approximate extended models. The results obtained on the basis of the proposed kinetic modellings for each of these applications show an appreciable agreement with the ones known in classical literature.

The following improvements and extensions however can be considered.

The study of the steady detonation wave and its linear stability could include the heat of reaction and the dissipative contribution due to an endothermic reaction as well as the transport effects. In addition, this analysis could also be extended to the bi-dimensional case.

At last, the dynamics of acoustic waves and light scattering could be widened starting from the exact BE rather than resorting to the approximate BGK equation and considering a detailed description of the transport picture, where also cross effects would be taken into account.

Acknowledgments

The paper is partially supported by Brazilian Research Council (CNPq), by Italian National Project GNFM 2010/11 and by Minho University Mathematics Centre (CMAT), FEDER–COMPETE Funds, FCT–Project Est-C/MAT/UI0013/2011.

References

[1] Brun R., *Introduction to Reactive Gas Dynamics*, Oxford University Press, NewYork, 2009.

[2] Nagnibeda E. and Kustova E., *Non-Equilibrium Reacting Gas Flows: Kinetic Theory of Transport and Relaxation Processes*, Springer Verlag, Berlin, 2009.

[3] Kremer G. M., *An introduction to the Boltzmann equation and transport processes in gases*, Springer, Berlin, 2010.

[4] Shizgal B. D. and Chikhaoui A., "On the use temperature parameterized rate coefficients in the estimation of non-equilibrium reaction rates", Physica A, Vol. 365, 2006, pp. 317-332.

[5] Prigogine I. and Xhrouet E., "On the perturbation of maxwell distribution function by chemical reactions in gases", *Physica*, Vol. 15, 1949, pp. 913-932.

[6] Prigogine I. and Mahieu M., "Sur la perturbation de la distribution de Maxwell par des réactions chimiques en phase gazeuse", *Physica*, Vol. 16, 1950, pp. 51-64.

[7] Cercignani C., *Theory and application of the Boltzmann equation*, Scottish Academic Press, Edinburgh, 1975.

[8] Kremer G. M., Pandolfi Bianchi M. and Soares A. J., "A relaxation kinetic model for transport phenomena in a reactive flow", *Phys. Fluids*, Vol. 18, 037104, 2006, pp. 1-15.

[9] Conforto F., Monaco R., Schürrer F.and Ziegler I., "Steady detonation waves via the Boltzmann equation for a reacting mixture", *J. Phys. A: Math. Gen.*, Vol. 36, 2003, pp. 5381-5398.

[10] Fickett W., *Introduction to Detonation Theory*, University of California, Berkeley 1986.

[11] Monaco R., Pandolfi Bianchi M. and Soares A. J., "A Kinetic Approach to Propagation and Stability of Detonation Waves", *Rarefied Gas Dynamics*, AIP Proceedings, Ed. Abe T., Vol. 1084, 2009, pp. 45-50.

[12] Pandolfi Bianchi M. and Soares A. J., "Modelling and solutions to the linear stability of a detonation wave in the kinetic frame", *J. Difference Eqs. App.*, Vol. 17, 2011, pp. 1169-1184.

[13] Berne B. and Pecora R., *Dynamic Light Scattering*, Wiley, New York, 1976.

[14] Kremer G. M., Marques Jr. W., Pandolfi Bianchi M. and Soares A. J., "Spectral Distribution of Scattered Light from a Chemical Relaxation System", *Rarefied Gas Dynamics*, AIP Proceedings, Eds. Levin D.A., Wysong I. J. and Garcia A. L., Vol. 1333, 2011, pp. 673-678.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Numerical experiments on the polygonalization of converging shock waves by means of aerodynamic obstacles

F. Vignati, A. Guardone
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano
via La Masa 34, 20156 Milano, Italy

Keywords: *Implosions, Polygonal Shocks, Obstacles*

Abstract

Numerical simulations are performed to observe some particular cases of cylindrical converging shocks being reshaped into a polygonal form by the interaction with aerodynamic obstacles.

Since lenticular obstacles introduce a weaker perturbation with respect to cylindrical ones, the number of obstacles must be increased to guarantee that the whole shock surface is affected. Therefore, the final polygonal shock shape is usually made of more edges than that obtained by using cylindrical obstacles.

We observed a different shock reshaping process if the initial shock is already polygonal. In fact the simultaneous symmetry in the obstacles position and the concavity of the convergent shock results in two reflections (the first one of the shock against the obstacle; the second one of the reflected shock wave at the symmetry surface), which interact with the shock itself, eventually modifying its geometry into an almost linear one.

All the simulations have been performed for standard air using an unstructured-grid adaptive code.

1 Introduction

The problem of cylindrical converging shocks has been investigated in several studies, because of its relevance to diverse applications including Inertial Confinement Fusion and Sonoluminescence. In the latter cases, the desired effect is to attain high pressure and temperature at the center of the implosion.

Previous both numerical and experimental studies [1] show the intrinsic instability of these shocks fronts. This implies that the shock effectiveness may decrease, as a consequence of a loss of regularity of the surface.

Fortunately, a second aspect of the unstable behavior of converging shocks is the possibility of modifying their shape in relatively easy ways. One of the most effective one consists in placing a number of obstacles along the path of the shock, which cause a modification of its shape. In particular, if the obstacles are properly shaped and placed, the shock fronts turns from cylindrical into polygonal, which is a stable configuration.

Most of available references concern converging shocks meeting cylindrical obstacles [2, 3], and demonstrate the effectiveness of use of obstacles in reshaping the shock waves. Unfortunately, cylindrical obstacles cause also not negligible losses [4]. As a consequence,

recent studies focused upon the possibility of employing lenticular obstacles, whose aerodynamic shape is more suitable in order to reduce these losses [1].

This paper illustrates the results from numerical simulations performed on cylindrical converging shocks with different configurations of aerodynamic obstacles in a symmetrical arrangement. It consists of four sections: the first one presents a summary of the characteristics of the problem, and the adopted guidelines in the research of results. The second section describes the various conditions of the domain and the flow field of each simulations, with special regards to the different highlighted phenomena. The third part illustrates the results of the simulations and reports some considerations about the quality of the generated shocks. The final section contains conclusions, remarks and future developments of the work.

2 Physical problem characteristics

2.1 Description of the phenomenon

The reference framework is that illustrated in Fig. 1, concerning the case of eight symmetrical aerodynamic obstacles. A cylindrical shock is converging towards the origin, and aerodynamic obstacles are symmetrically placed.

Thanks to the symmetry of the physical problem, we refer to a slice of the whole domain, whose boundaries represent symmetry lines [5].

With both cylindrical and lenticular obstacles, the cylindrical shock turns approximately into a polygon, with its vertexes placed in correspondence to either the obstacle axis or the symmetry lines. A narrower observation reveals that the final shape of the shock front is closer to an astroid: the sides concavity is not zero, even though very small, and it is turned outwards.

With reference to Fig. 2, the shock reshaping process takes place in at least three steps.

Firstly, as soon as the shock reaches the obstacle leading edge, it is reflected. This first reflection, following identified as “wave A”, is the *Regular reflection* and immediately affects the shock, reversing its local concavity outwards. Like the rest of the phenomenon, also wave A is unsteady and it moves away from the obstacle in time, while the perturbed shock continues propagating inwards.

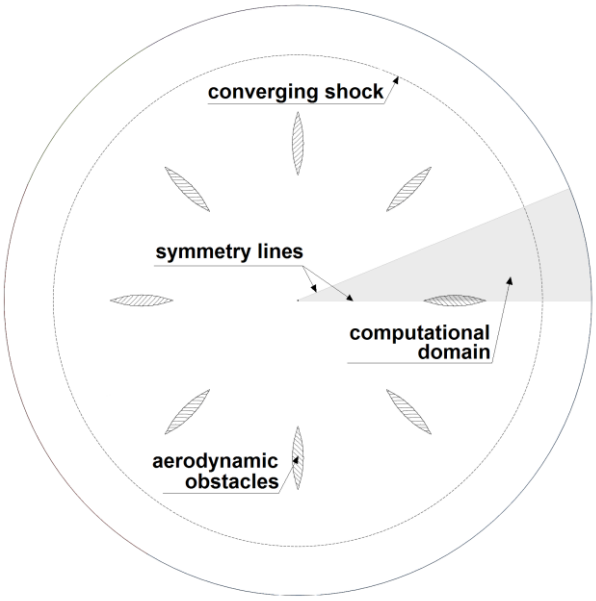


Fig. 1 Sketch of the investigated problem. The highlighted part is the computational domain: this reduction is possible thank to the geometrical symmetry.

The second reflection occurs when wave A reaches the symmetry surface. The resulting wave, named “wave B” in Fig. 2, propagates towards the center of the implosion.

The third step of the polygonalization is the interaction between wave B and the shock and, consequently, which starts as soon as wave B is formed. The reshaped shock concavity is turned outwards, see Fig. 2.

Immediately after this step, the shock is approximately polygonal, with $2n$ sides, where n is the number of obstacles. Nevertheless, the propagation velocity of the shock portions that were formed after the interaction with wave A is lower than Mach number of the sides generated by the wave B. As a consequence, the first ones interacts with the latter and eventually disappear. Afterwards, the shock consists of a polygonal front, whose vertexes are in correspondence of the symmetry axis of the obstacles, as represented in Fig. 2 [6]. The final configuration of the shock depends on both the

shock Mach number and the obstacles geometry.

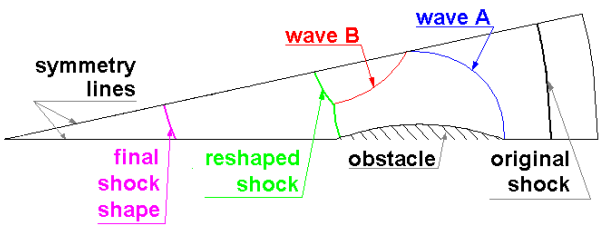


Fig. 2 Sketch of the evolution of the shock front shape, highlighting waves A and B, and the symmetry boundaries

In fact, if the shock reflection occurs as a strong *Mach reflection*, a further modification in the shock shape occurs, starting from the trailing edges of the obstacles. In this case, the shock transforms again into a *2n*-polygon, with the new sides converging faster than the previous ones. Eventually, the shock turns again into an *n*-polygon, but now its vertexes are placed along the symmetry lines.

Otherwise, only a negligible shock deformation is observed at the vertexes, due to the weak reflection. Therefore, the polygonal shock approaches its center with the same shape as defined after the third step. This double possibility is confirmed also by [4] in case of cylindrical obstacles.

2.2 Difference between cylindrical and aerodynamic obstacles

Previous works show that the interaction between cylindrical obstacles and converging shocks is usually sufficient to produce the desired reshaping effect. Nevertheless, because of high viscous losses, this kind of obstacles appears to be not optimal. Recently, the attention moved towards aerodynamic obstacles, which produce a weaker interaction with the shock. A drawback of e.g. lenticular obstacles is that also the size of the shock affected region is smaller and consequently the number of obstacle is to be increased for the reshaping process to occur effectively.

This works moves from the observation that, by increasing the number of obstacles, the relevant arc of the shock approaches a straight line. The consequent question is whether it is possible to obtain a reshaping not only of cylindrical, but also of polygonal or planar shock fronts by means of disturbances. The above question is also supported by the results in [6], which show a "self-reshaping" of polygonal shock, due mainly to Mach reflections, downstream the obstacles and, consequently, when the shock is already polygonal.

3 Numerical simulations setting

The observations presented in this paper make reference to simulations performed using standard air as test fluid, threated as a perfect gas.

3.1 Explored shock-obstacle configurations

Three kinds of numerical simulations were considered in order to highlight different behaviors. A first set of simulations deals with simple planar shocks propagating in a rectangular domain and meeting one obstacle. This case is the most simple one, since the shock wave is not interacting with the symmetry plane and it is not concave. The interest towards such a simplified condition is twofold. On the one hand, these simulations allowed us to quantify the disturbance on the shock shape introduced by an aerodynamic obstacle, mainly in terms of the size of the affected region. On the other hand, this simple configuration allowed to validate the capability of the adopted software to capture the shock wave with the required accuracy. In particular the simulations were performed using the *Flowmesh* code from the Aerospace Department, Politecnico di Milano [7].

Following this preliminary assessment, the second condition that was investigated is that of a cylindrical implosion meeting forty obstacles (see Fig. 3). In this case, the angle in the origin, corresponding to a configuration with forty symmetrical obstacle, is $\alpha = 4.5°$.

The difference with respect to the cases with few obstacles [3] is both practical and theoretical. Indeed, the higher is the number of obstacles, the more simultaneous is the convergence of the shock and, consequently, the

higher are pressure and temperature at the center of the implosion, that is the desired effect, i.e., in Inertial Confinement Fusion application.

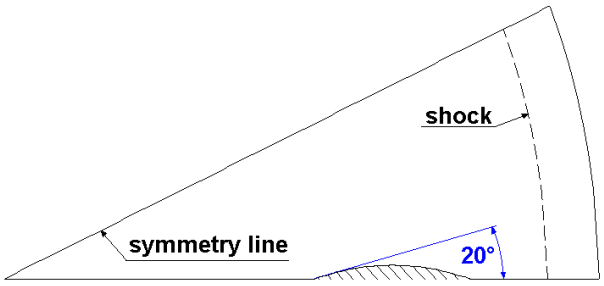


Fig. 3 A sketch of the computational domain for the implosion simulations

Moreover, focusing on a very small section of the circumference allows to compare the results of the cylindrical shock to those of a plane one converging in the same conditions, and to investigate the sole effect of the shock convexity. Eventually, the number of obstacles influences also secondary reflections, for example of wave B on the obstacle. Nevertheless, this kind of considerations have not been the object of this paper, where secondary reflections have been considered as negligible, and may be the topic of a future work.

The third set of simulation is a regular forty-sided-polygonal shock propagating towards the center an encountering the forty aerodynamic obstacles considered above. This problem is interesting because it links the first two cases. Indeed, the high number of sides makes this shock front very close to the cylindrical one but still with zero concavity. Therefore, the comparison between the second and this third set of simulations is useful in order to observe the influence of the curvature of the shock in the interaction first with the obstacle, and later with the wave B. The very small difference of concavity between these two cases will highlight also small effects of reshaping due to the shock curvature. Moreover, the angle between each side and the symmetry line is so small that only a weak reflection occurs, that is to be considered negligible, as shown in the results section. This is a further element that assimilates the two last considered geometries, since the cylindrical shock is everywhere normal to the infinite symmetry surfaces.

All the three sets of simulations have been performed using obstacles of the same shape. They are made of a circular arc; the angle formed by the two tangent lines at the leading and trailing edge is 40°, see figure 3. The ratio between the obstacle length and its distance from the origin is not constant for the different configurations. This is not expected to affect the comparisons, because the simulations have been performed using a software based upon Euler equations, and consequently any effect of boundary layer or recirculation has been neglected.

Another feature common to all the simulations is the Mach number of the shock: three different numbers have been investigated, respectively of 1.2, 1.4 and 1.6.

3.2 Numerical parameters

The used code is the in-house-developed *Flowmesh*, which allows to perform mesh adaptation. In all simulations, the dimensionless ratio between the minimal element dimension and the unit length has been set to 10^{-6}. This quantity turned out to be a good compromise between the speed of calculation and the accuracy. All the simulations have been ran with a Courant – Friedrichs – Lewy number evolving from 1 to 100.

3.3 Initial conditions

For the two sets of simulations concerning planar shocks, initial conditions have been set thanks to Rankine-Hugoniot conditions. External pressure, density and velocity components have been calculated from the unperturbed state and the shock Mach number; the discontinuity has been placed at a distance of R = 9.5 from the origin.

On the other hand, the initial condition for the cylindrical implosion simulation was chosen in order to mimic the effect of a “cylindrical diaphragm” [8]: a pressure discontinuity was placed at R = 21.5, with a ratio of 10 between internal and external pressure. Everywhere density is homogeneous and both velocity

components are equal to zero. The solution of the cylindrical Riemann problem is made to avoid a complex initialization procedure because the converging shock produces a non-uniform post shock condition due to its cylindrical shape. Unfortunately, such a choice forced us to place the initial discontinuity further away, to ensure that the contact discontinuity following the shock did not affect the shock polygonalization process. In the latter case, the boundary condition at the right boundary is a non-reflective Riemann condition, to avoid the reflection of the rarefaction wave moving outwards.

4 Results

Numerical simulations of the planar shock problem (see Fig. 4 for the initial mesh) show that the influence of the obstacle remains limited, and its most important effect is the generation of wave A. This is particularly evident by the observation of Fig. 5, which reports the density maps of the flow field generated by the interaction of the planar shock with the obstacle at two different Mach numbers. Notice that both the size and the intensity of the disturbance in the shock shape are very small.

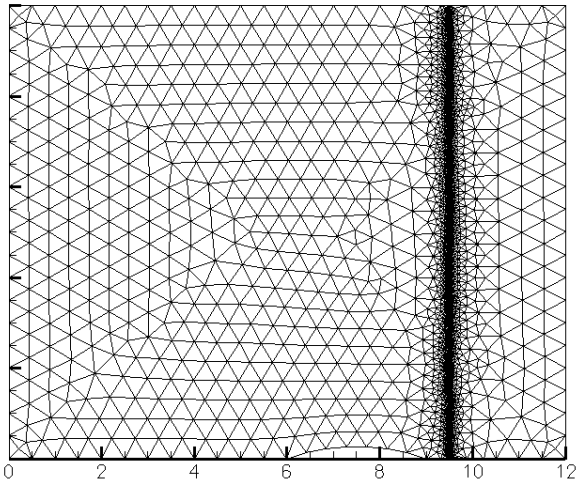


Fig. 4 Domain and adapted mesh used in the simulation of the simple planar shock meeting one obstacle. The grid is finer in radius R = 9.5 where the shock has initially been placed.

Moreover, as shown in Fig. 6, about one chord downstream the obstacle, both the reflected wave and the disturbance faded almost completely. This means that the only wave A is not sufficient to completely reshape the shock.

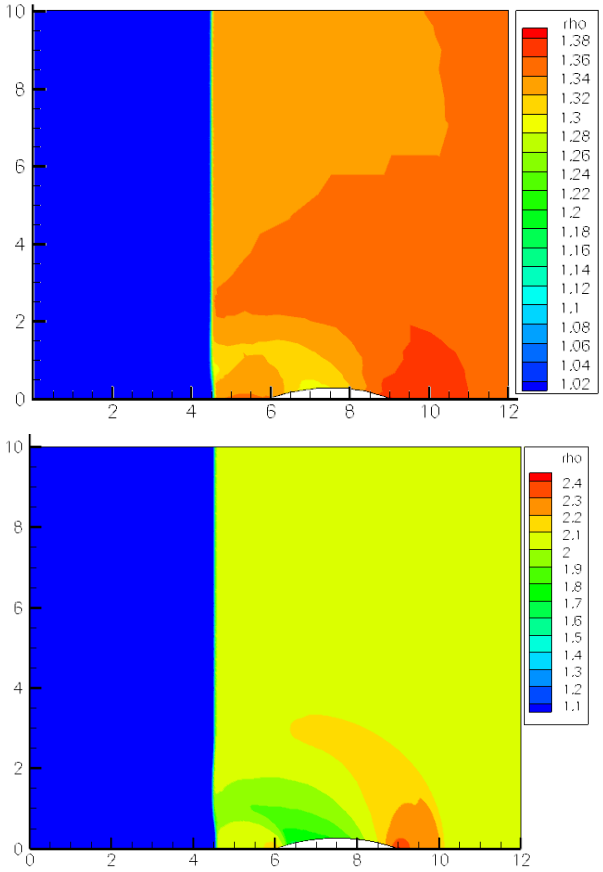


Fig. 5 Density colormap of the flow field after the interaction of the simple planar shock with one obstacle at half chord downstream the obstacle. Mach number is 1.2 and 1.6 respectively

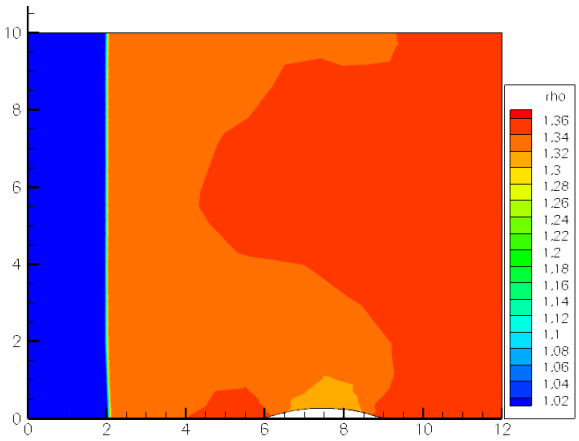


Fig. 6 Density color map of the case with Mach number 1.2. The shock is 1.3 chords downstream the obstacle.

Observations of the simulation performed on the real implosion are now presented.

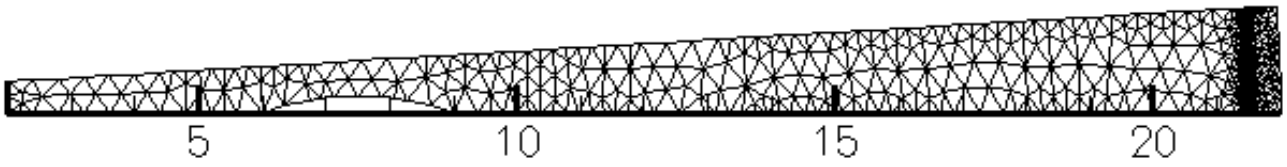
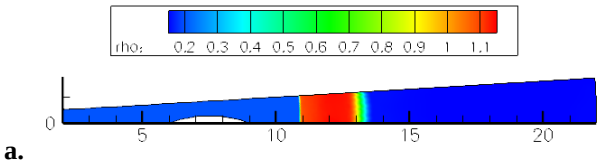


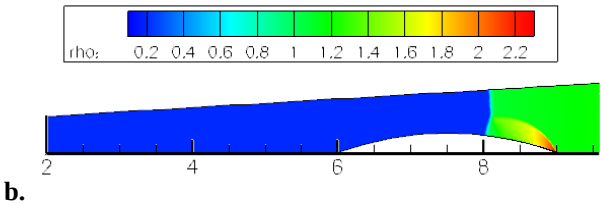
Fig. 7 Domain and adapted mesh used in the simulation of converging cylindrical shock meeting forty obstacles. In radius R = 21.5 there is the pressure discontinuity, in correspondence of the denser region of the grid

The computational mesh, defined at time t = 0 after a grid adaptation process, is reported in Fig. 7.

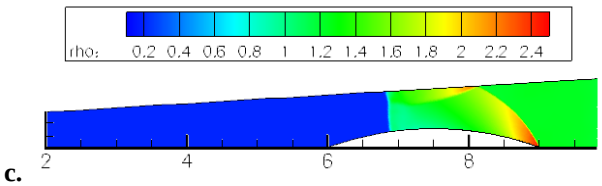
Figure 8a illustrates the propagation of the unperturbed shock. Figures 8b and 8c show the formation of waves A and B. Figure 8d shows the reflected shock as it approaches the trailing edge of the obstacle. Finally, in Fig. 8e the shape of the shock front leaving the obstacle is represented, where the Mach reflection at the symmetry plane is clearly visible.



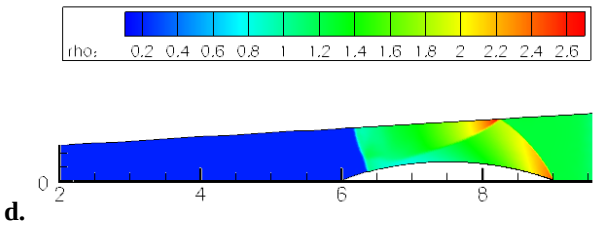
a.



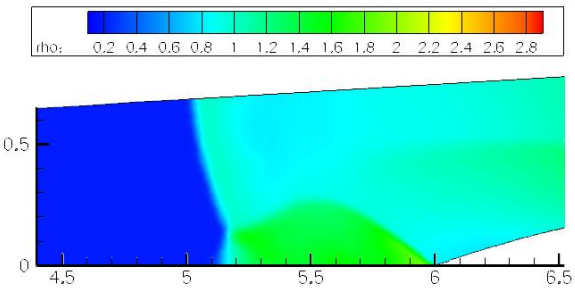
b.



c.



d.



e.

Fig. 8 Density color maps of the evolution of the cylindrical shock. a: unperturbed shock. b: wave A formation and local inversion in the shock concavity. c: wave B formation and influence on the shock. d: new sides extension, against the previously formed. e: final shock shape.

Finally, the case of the polygonal shock with forty obstacles is addressed. In all the explored cases there is an interaction between the incident shock and the obstacle, even though weak. Because of the proximity of the obstacle to the symmetry surface, wave A reaches the symmetry surface and it is reflected, but its influence on the reshaping is very limited.

Figure 9 illustrates the initial mesh of the simulated slice, with the grid refinement at the shock initial position. Since a reflection behind the shock is expected, due to the symmetry condition at the upper symmetry boundary, the grid is refined not only in correspondence of the shock, but also close to the upper boundary.

Figure 10 shows the propagation of the shock at Mach number 1.6: the low value of density gradients across wave A and wave B is evident (Fig. 10a and 10b). Moreover, as shown in Fig. 10d, the final effect on the reshaping is weak, since the reflection at the trailing edge of the obstacle compensates previous modifications to the front shape.

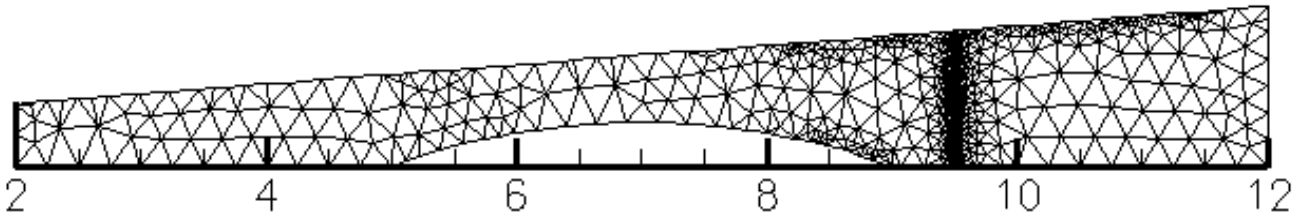


Fig. 9 Domain and adapted mesh used in the simulation of the converging planar shock meeting forty obstacles. The grid is finer at R = 9.5 where the shock has initially been placed, and behind the shock, where its reflection on the symmetry surface occurs.

The same effect has been observed also in the other two cases at different Mach numbers.

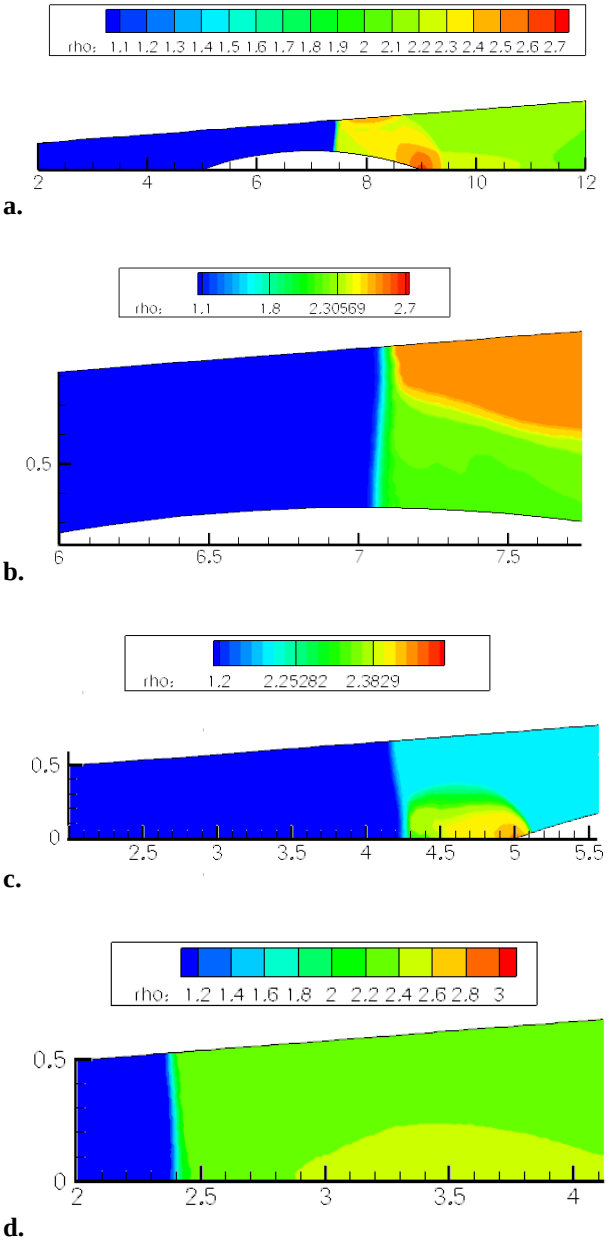


Fig. 10 Different steps in the propagation of the forty sides polygonal shock. a: formation of waves A and B. b: zoom in correspondence of the interaction between wave B and the shock. c-d: final reshaping effect.

5 Conclusion

Numerical simulations were performed in order to explore the phenomenon of the reshaping of converging shocks.

Our simulations confirmed that aerodynamic obstacles can provide the reshaping, but for the considered cases reshaping occurs only if the incident shock is curved. A further consideration concerns the shock Mach number: the higher it is, the more evident is the effect of the interactions, at least when Mach remains between 1 and 2. The interaction among waves is also a very important factor.

To maximize both the temperature and the pressure at the center of the implosion, the optimal combination of Mach number of the shock, number of obstacles, chord-distance ratio and gas is to be chosen. In this respect it may be interesting to extend the present research to shocks propagating in real gases, to vary the position of the obstacles to produce a non-symmetric shock configuration, and to study suitable obstacle shapes that reduce the intensity of the shock reflection at the trailing edge.

References

[1] Kjellander M., Tillmark N and Apazidis N, "Thermal radiation from a converging shock implosion", *Physics of Fluids*, Vol. 22, 2010, pp.1-12.

[2] Eliasson V., "On focusing of shock waves", Ph.D. Dissertation, Department of Mechanics, KTH, Sweden, 2007.

[3] Eliasson V., Apazidis N., Tillmark N., "Controlling the form of strong converging

shocks by means of disturbances", *Shock Waves*, Vol 17, 2007, pp. 29-42.

[4] Shi H., Yamamura K., "The interaction between shock waves and solid spheres arrays in a shock tube", *Acta Mechanica Sinica*, Vol. 20, No. 3, 2004, pp. 219-227.

[5] De Santis D., Geraci G., Guardone A., "Equivalence condition for finite volume/element discretizations in cylindrical coordinates", *V European Conference on Computational Fluid Dynamics ECCOMAS CFD*, Lisbon, Portugal, 2010, pp. 1-17.

[6] Eliasson V., Kjellander M., Apazidis N., "Regular versus Mach reflection for converging polygonal shocks", *Shock Waves*, Vol. 17, 2007, pp. 43-50.

[7] Isola D., Guardone A., Quaranta G., "Arbitrary Lagrangian Eulerian formulation for two-dimensional flows using dynamic meshes with edge-swapping", *Journal of Computational Physics* 230, 2011, pp- 7706-7722.

[8] Zel'dovich Ya. B., Raizer Yu. P., *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, Academic Press Inc., New York, 1967, Vol. 2, Chap 12.

Experimental and Numerical Investigation on the MHD Interaction in a Hypersonic Plasma Flow.

Andrea Cristofolini, Carlo A. Borghi, Gabriele Neretti, Fabio Roveda
Department of Electrical Engineering, University of Bologna, Italy.

Andrea Passaro, Damiano Baccarella
Alta SpA, Pisa, Italy.

Keywords: *Hypersonic flight control, hypersonic flows, magnetohydrodynamics, MHD interaction*

Abstract

The MHD interaction in a hypersonic argon flow was experimentally investigated around a conical test body. The aim of the experiment was to produce a database to be used for the validation of numerical codes for the analysis and simulation of the magnetofluiddynamics in hypersonic flows. This paper summarize the work carried out in the High-Enthalpy Arc-heated hypersonic wind Tunnel of Alta, Pisa, Italy, in cooperation with The Department of Electrical Engineering of the University of Bologna, under the support of The European Space Agency. In the experiments reported, the MHD interaction has been investigated in a Mach 6 ionized argon flow. The experiments has shown a large effect of the MHD interaction on the values of the measured quantities. The experimental results has been numerically investigated. A discussion of the results is presented here.

1 Introduction

The utilization of the MHD-interaction in the hypersonic flight has been investigated since the late '50 [1], and is receiving an increasing interest. Many possible applications of the MHD interaction for hypersonic flight has been proposed and are currently investigated, such as thermal flux mitigation, flight control and shock control at the inlet of an air breathing thruster [2-7]. In AJAX project MHD techniques are utilized to by-pass kinetic energy of the working fluid from the supersonic diffuser to the nozzle. By doing this, the flow velocity in the combustion chamber is reduced to acceptable values, even for high Mach numbers.

The knowledge of the physics of the MHD-hypersonic flow interaction is essential for the application of this technology to the hypersonic flight. Various experiments has been carried out in Russia [8-9], in USA [10], in Japan [11,12] and in European countries [13]. In a recent work [14], the experimental results of the MHD interaction in hypersonic regime on a ramp with a Hall electrodynamic configuration has been presented. In order to maximize the effect of the MHD interaction, a magnetic field perpendicular to the test body surface were used, and a set of electrodes had been posed on the test body surface to shorten the Faraday field (i.e. the electric field component in the U×B direction). However, despite of the electrodes configuration, the measured electric potential on the test body surface revealed that the Faraday component of the electric field was not shorten. The proposed explanation is that the boundary layer, constituted by a low

electrical conductivity plasma, introduces a high electric resistance for the current between the test body and the core flow and prevents the shortening of the Faraday field, hence causing a reduction of the effect of the MHD interaction.

A particular attention has been then devoted in designing a 2D test body to operate in a low density plasma. In this condition, the plasma is characterized by relatively high values of the electron mobility and, thus, of the Hall parameter. This means that the electric current density will exhibit some tendency to flow in a perpendicular direction to the magnetic flux density and the electric force per charge unit E + U×B. As a result, for high values of Hall parameter, it would be advisable to exploit the Hall component (i.e. the component parallel to the flow velocity) of the electric field to sustain the current density which generates MHD interaction. Experimental results as well as a discussion on the numerical rebuilding are here presented.

2 Experimental Facility

Alta High-Enthalpy Arc-heated Tunnel (HEAT) is a pulsed hypersonic wind tunnel operative since 1996 [15, 16]. It can produce Mach 6 flows in a low to medium Reynolds number range (10^4-10^6). In the wind tunnel settling chamber the gas is heated by means of an arc discharge powered by a 260 kW DC power supply and delivering arc currents up to 630 A during running times of 20-300 ms. In the settling chamber a total specific enthalpy up to 6 MJ/kg with stagnation pressures up to 9 bar can be obtained. The gas heater can be operated with air, nitrogen, argon, CO_2 in a pulsed, quasi-steady mode. At the exit of the settling chamber a nozzle is mounted. The effective test section at the exit of the nozzle has a diameter of 55 mm. The arc heater scheme is shown in Fig. 1. The wind tunnel (gas heater and nozzle) is installed on a vacuum test section of 600 mm diameter. In the test section there are optical accesses from all sides. This section is connected to a vacuum chamber of a volume of 4.1 m^3. The chamber is evacuated by four rotary pumps to reach an ultimate pressure of 10 Pa before each run. The flow characteristics are obtained by means of pitot probes [17] provided with fast miniaturized piezo-resistive pressure transducers and stagnation temperature probes provided with fast coaxial thermocouples or thin-film gauges. In the settling chamber the pressure is measured by fast pressure transducers. Accuracies of about 3% are obtained in all pressure measurements. The total enthalpy is evaluated from the arc discharge power and compared with the values obtained from the stagnation temperature measured in the test section. The total enthalpy error is ±5-10%.

The test gas used in this experiment is argon. The nozzle utilized for the experiments reported in this paper is a Mach 6 conical nozzle, with an effective exit test section of a diameter of about 50 mm and is shown in Fig. 2.

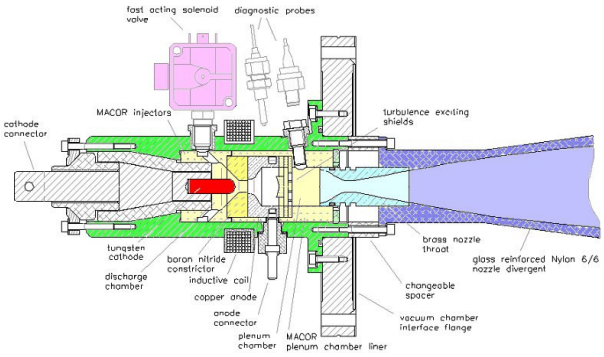


Fig. 1. Arc heater scheme.

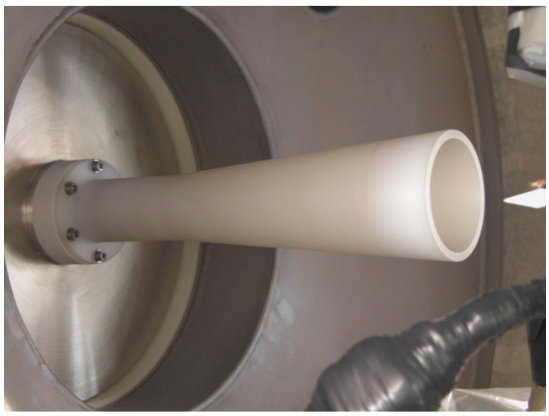


Fig. 2. Mach 6 nozzle mounted at the exit of the settling chamber

Properties such as electron temperature and electron density are required for the evaluation

of the electric transport properties of the plasma (i.e. electron mobility and electrical conductivity), which are essential data for the numerical rebuilding of the experiment and for the interpretation of the experimental results. Thus, a crucial point is the evaluation of the plasma parameter in the free stream flux at the nozzle exit. In order to do this, optical diagnostics and microwave measurements have been utilized.

The diagnostic results reported in this paper are from pressure measurements and electric voltage measurements in the shock layer of the test body, and imaging of the plasma under the MHD interacting effect. A PCO SensiCam SVGA interline CCD color camera, with a minimum exposure time of 100 ns, has been utilized for the imaging of the MHD interacting plasma. The CCD quantum efficiency peak is nearly 53% at a wavelength of 380 nm. The electric measurement are done by means of sets of electrical probes mounted on the test body wall. In order to perform the pressure measurements miniaturized Kulite fast pressure transducer (model XCS-062-5A) was mounted flush with the wall, along the lateral surface of the test body. Due to the expected low level of the pressure to be measured, the pressure signal was sent to an amplifier circuit based on the Analog Devices SSM2019 chips, that allows x10, x100, and x1000 DC gain. The signals were acquired through the HEAT data acquisition system, based on the Labview 7 environment and a National Instruments 6071 board.

3 Test Body Design

At the low plasma densities, typical of the re-entry of hypersonic vehicles into the atmosphere, the Hall parameter is high. The test body design is thus aimed to exploit the Hall field to sustain the Faraday current which generate the MHD interaction. In order to do this, a conical geometry with the axis of the cone directed along the gas free stream flow has been chosen. The magnetic flux density is generated by means of permanent magnets. The configuration adopted appears to be particularly suitable as it does not require any conductor carrying current for the generation of the magnetic flux density.

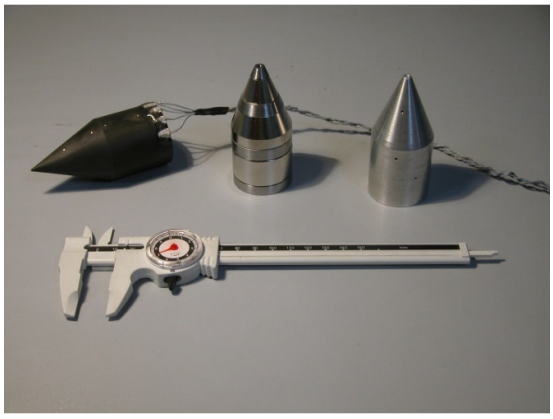


Fig. 3.. Test bodies

The test body is shown in Fig. 3. It is made of two sections. The first one is a cone with a half-vertex angle of 22.5◦ and a base diameter of 40 mm. The second section is a cylinder, starting from the cone base with a length (height of the cylinder) of 48.3 mm. The cylindrical section was added in order to allow the placement of a third magnet. The Magnetic field necessary for the MHD interaction is produced by an array of three permanent magnets in the cone central part, at the cone base, and in the cylinder central part. Four sections of ARMCO iron alternate the magnets, and they are placed at the cone vertex, the central parts of the cone and cylinder, and the base of it. The B-field distribution generated by the permanent magnets as a result from the analysis performed by means of the finite element magnetostatic numerical code is shown in Fig. 4. A non magnetic test body has also been made, in order to produce a reference measurements of the flow without MHD interaction.

In fig. 3, the magnetic and non magnetic test bodies are shown, along with the nylon covering. In the figure, the test body on the left-hand side is the magnetic test body used for the MHD-interaction experiments. The non magnetic test body on the right-hand side is completely made of aluminum. The reinforced nylon shell can also be observed. Four electrostatic probes are placed on the shell, to provide a measure of the Hall field produced by

the MHD interaction. Pressure sensors has been placed in the iron elements on the magnetic test body, and in the corresponding position on the non magnetic one.

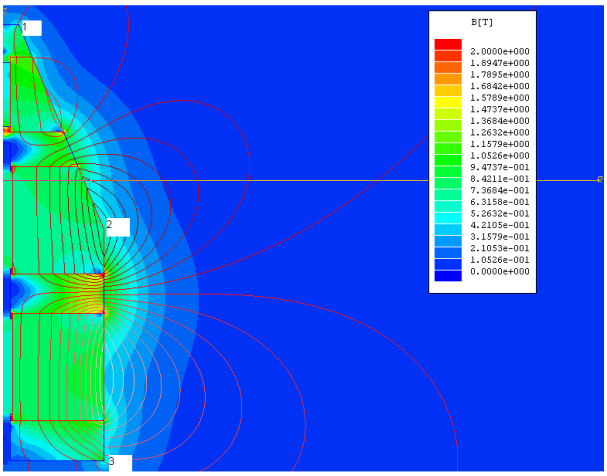


Fig. 4. Magnetic field around the small cone test body

4 Flow Characterization

In order to perform the aerothermodynamic characterization of the produced flow, the following set of instruments was used:

–two absolute Kulite XCS-062 sensors in the tunnel settling chamber (one with a 250 psi FSO and one 100 psi),

–a supersonic pitot probe equipped with a Kulite XCS-062-5A miniaturized pressure sensor for high frequency Mach number retrieval

–a hemispheric stagnation point temperature probe equipped with a MEDTHERM coaxial thermocouple for the stagnation enthalpy and total enthalpy retrieval through the Sutton&Graves relations [18].

A complete mapping of the nozzle flow was performed both for low and high-enthalpy conditions. The characterization activity allowed to determine the core flow dimensions, actual axial symmetry degree of the flow, fluctuations both for what concerns Mach and temperature, repeatability of the test conditions. The picture in Fig. 5 shows the typical flow profile at the nozzle exit, and provides an estimation of the deviation from the axial symmetry. In fig. 6, the Mach number map of the test section area is reported.

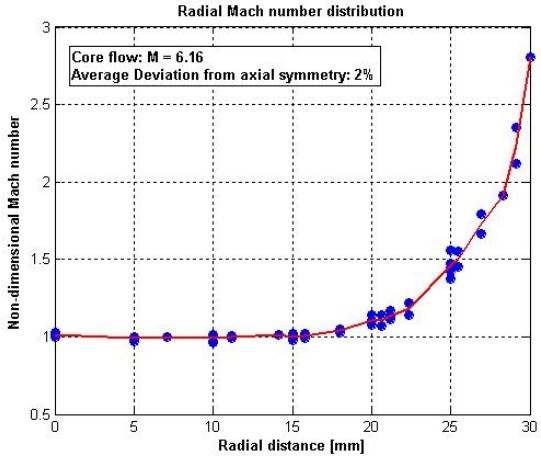


Fig. 5. Radial profile of the flow Mach number: mean profile (red) superimposed on the actual values in different sections.

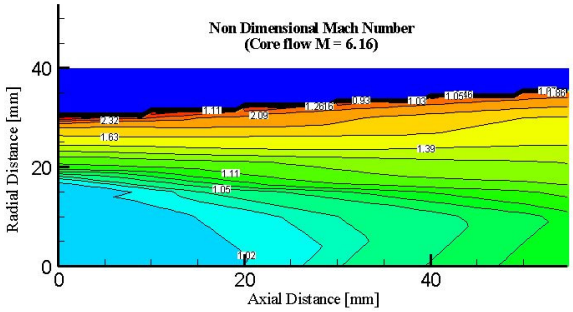


Fig. 6. Mach number profile in the test section

In general, for all of the considered conditions, the flow can be considered axial symmetrical and uniform for a diameter of about 40 mm (the Mach number at 20 mm is always within 9-10% higher than the value of the core flow), with a slight non-symmetry starting at 18 mm (maximum value of the asymmetry 4%). The flow expansion continues in the test section with an increase of 5 % in the first 50 mm. Relatively high differences in total enthalpy level can be individuated comparing different runs, probably due to the complicate flow pattern that is created in the settling chamber at this low pressure levels: this leads to the need of several repetition (at least 5) in each test condition in order to gather a statistically significant sample. The flow

condition utilized in the experiment are listed in Table 1

Table 1. Flow Characteristics in the test conditions

Condition	Mach Number	Mass flow rate	Average pressure in the heating chamber	Gas temperature in the heating chamber
	[-]	[g/s]	[bar]	[K]
A	6.16	5.034	0.537	4455
B	6.14	5.851	0.608	3850
C	6.20	7.027	0.671	3344

In order to gain information about the plasma characteristic of the flow, spectroscopic and microwave diagnostic has been utilized.

The radiation intensity emitted by the plasma in the stagnation region before expansion in the nozzle has been measured. In order to do this, an F/4 mm collimation lens of a focal length of 40 mm has been placed in the heating chamber and was focused on the central region of it. The lens was connected through a 600 μm optical fiber to a Jobin-Yvon Hr 460 spectrometer, with a 460 mm focal length, an F/5.3 focal ratio, and JY 1200 lines/mm holographic grating. An SVGA fast shutter PCO CCD camera with a resolution and a pixel size of 6 μm × 6 μm was connected to the spectrometer.

Line emission, continuum recombination radiation and Stark broadening has been observed. The electron number density n_e in the heating region has been determined by measuring the broadening effect of Hydrogen α and Hydrogen β lines of wave lengths of 656.28 nm and 486.50 nm respectively. Hydrogen was present as an impurity due to the diffusion pumps utilized by the vacuum system. These lines are broadened due to the Coulomb interactions and are characterized by a Lorentz shape (Stark broadening). The Stark broadening shape depends on the charged particles surrounding the emitters and hence on n_e. Data table in [19] has been utilized to obtain the final results

As at partial local thermal equilibrium (PLTE) condition may be assumed, the free electrons have a Maxwellian electron energy distribution determined by the electron temperature and the radiation energy distribution due to the recombination electron-ion delivering a particle excited at the l level is obeys the Boltzmann equation. The values of the radiation intensity at two different wave frequencies allow determining T_e and, when T_e is known, the radiation intensity at a wave length allows deriving n_e. In the test considered both T_e and n_e have been derived by the experimental values of the continuum recombination radiation [20]. The electron number density is also obtained by the Saha equation at the value of the electron temperature obtained by the measurements.

The electron number density and the collision frequency of the plasma in the free stream at the exit region of the nozzle have been determined by the microwave absorption technique [21]. In order to do this, the measurements have been carried out before the test body placing. The results obtained are summarized in Table 2, where the values of the electron density obtained utilized the different approaches described so far are reported.

Table 2 Electron number density in the heating chamber at conditions A, B and C of Table I

	H_α	H_β	Griem H_β	Contin.	ArI-430 nm/ Contin.	Saha
	[m^{-3}]	[m^{-3}]	[m^{-3}]	[m^{-3}]	[m^{-3}]	[m^{-3}]
A	8.5E+21	2.8E+21	3.0E+21	3.2E+21	1.55E+22	3.3E+21
B	8.0E+21	2.4E+21	2.7E+21	2.8E+21	1.39E+22	1.9E+21
C	6.1E+21	1.8E+21	1.9E+21	1.3E+21	1.20E+22	8.0E+20

The microwave diagnostic set up consisted of a microwave generator, two horn antennas and a crystal detector. A Gigatronics 1018-009 microwave generator, able to produce a non-modulated microwave signal from 50 MHz up to 18.5 GHz with a power exit of 10 dBm, has been used. Two 15 dB horn antennas designed for an optimal working frequency between 12 and 18 GHz have been placed inside the vacuum vessel, at opposite sides of the plasma flow, at about 50 mm from the nozzle exit. In this configuration the microwave beam crossed about 60mm with the Mach 6 nozzle and 80mm with the Mach 15 one**.** An HP 8472A crystal detector with a bandwidth of 10 MHz - 18.5G Hz, was connected to the receiving horn. In order to enhance the signal to noise ratio, at

each condition considered seven different signals have been detected and the transmission factor was determined from their average. Microwaves diagnostic has confirmed tha, during the expansion, the plasma kinetics is essentially frozen, and the ionization degree does not vary significantly.

5 Experimental Results on MHD Interaction

A picture of the interaction around the small cone model at test condition A is shown in fig. 7. Three bright rings are clearly observed in the regions where the B-field intensity is maximal and its direction is perpendicular to the flow velocity. This corresponds to the region facing the iron sections of the magnetic test body.



Fig. 7. MHD interaction around the small cone (test condition 1)

The pressure time behaviors on the cone surface are shown in figs. 8-10 for flow conditions A, B, and 3C, respectively. The plots were taken with and without MHD interaction (red and blue lines, respectively). In these experiments, the intensity of the fluctuations appears to be reduced. This is due to the better insulation used in these experiments.

In table 3, the summary of the test results with and without MHD interaction, pressure increases, and pressure coefficient C_p with and without MHD interaction and its increase at the three flow conditions considered in this experiment is reported. For all values given in

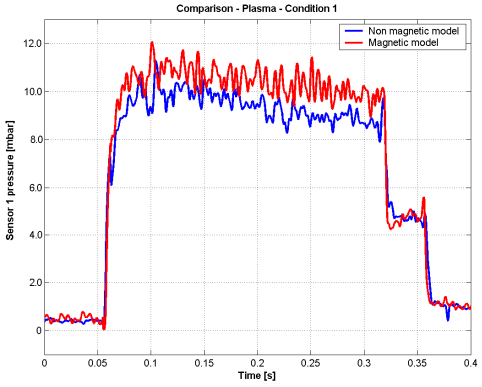


Fig. 8. Pressure measurements with and without MHD interaction at flow condition A: average pressure for the non-magnetic case is 9.01 mbar - MHD interaction case is 10.34 mbar (14 % increase)

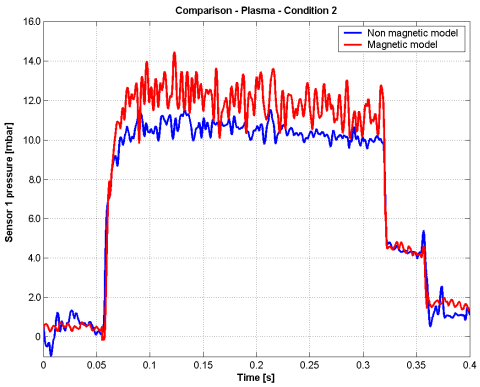


Fig. 9. Pressure measurements with and without MHD interaction at flow condition B: average pressure for the non-magnetic case is 10.30 mbar - MHD interaction case is 11.78 mbar (13 % increase).

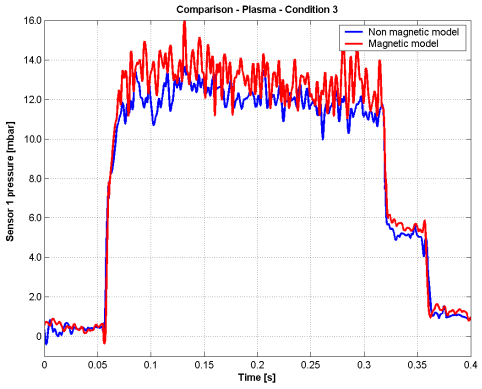


Figure Fig. 10. Pressure measurements with and without MHD interaction at flow condition C: average pressure for the non-magnetic case is 11.77 mbar - MHD interaction case is 12.97 mbar (10% increase).

the table, the averages between 100 and 200 ms from the flow-pulse starting time are reported. Table 3 refers to the measurements done by means of the probe placed on the conical surface. An increase between 10% and nearly 15% of the pressure and an increase between 7% and 13% of pressure coefficients were observed. These confirm the results obtained in the runs with the large test body. The measurements of the pressure sensors placed on the cylindrical surface are reported in Table 4. The accuracy in this region is low due to the low values of the pressure in this location and the vicinity to the sensor minimum threshold. Nevertheless, the increases of the pressure and the pressure coefficient at conditions A and B were recorded.

Table 3. Time averages of the pressure measurements on the cone sensor.

Cond.	Press. No MHD [mbar]	Press. MHD [mbar]	Pressure increase [%]	Cp no MHD	Cp MHD	Cp increase [%]
A	9.01	10.34	14.78	0.344	0.390	13.44
B	10.31	11.68	13.31	0.343	0.388	13.27
C	11.77	12.97	10.16	0.343	0.366	6.88

Table 4. Time averages of the pressure measurements on the cylinder sensor.

Cond.	Press. No MHD [mbar]	Press. MHD [mbar]	Pressure increase [%]	Cp no MHD	Cp MHD	Cp increase [%]
A	1.01	1.06	5.4	0.010	0.012	10.5
B	1.26	1.34	6.5	0.014	0.017	17.6
C	1.43	1.43	~0.0	0.013	0.013	~0.0

The Hall voltages measured during the experiments at flow conditions A, B, and C and at a stagnation pressure of 0.455 mbar are shown in Fig. 11, corresponding to the blue, green, light blue, and red lines, respectively. The dotted line indicates the Hall potential differences between the cone vertex and a position at 41 mm from the vertex and along the flow direction on the cone lateral surface. The full line corresponds to the Hall potential difference measured between the cone vertex and a probe at 17 mm from the vertex. These measurements indicate the Hall fields ranging between 400 and 500 V/m.

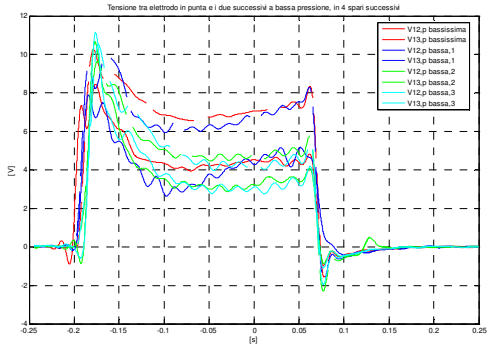


Fig. 11. Hall voltage distribution along the test body wall .

6 Numerical investigation on the MHD Interaction around the Test Body

A two dimensional model has been developed for the solution of the magnetohydrodynamic problem [22]. The model has been developed under the low magnetic Reynolds number assumption [23].

A cell-centered finite volume formulation has been adopted for the spatial discretization of the fluid dynamics. Fluxes are evaluated by means of a centered scheme, corrected by an Osher scheme. A second order accuracy is obtained, by a combination of upwind and centered discretization of the gradient. Electrodynamics is considered to vary as a sequence of steady state solutions driven by the evolution in time of the fluid dynamics and is solved by means of a finite element approach. Time integration is performed utilizing an explicit scheme, based on a fourth-order Runge-Kutta method.

A sensitivity analysis has been performed assigning three different values of the electron temperature. On the basis of the assumed values of electron temperature and ionization degree, the electric transport properties has been evaluated, utilizing the results of a model based on a two terms approximation of the Boltzamann equation, in the assumption of a maxwellian distribution of the electrons. In Fig. 12 are reported the quantity $n_{Ar} \times \mu_e$ as a function of electron temperature and ionization degree. From the results reported in Fig. 12, the electrical conductivity and the electron mobility are evaluated in the assumption of constant

electron temperature and ionization degree on the calculation domain. In the given assumption, the term $n_{Ar} \times \mu_e$ is constant on the domain. The electrical conductivity expressed as:

$$\sigma = e n_e \mu_e = e \alpha (n_{Ar} \mu_e) \qquad (1)$$

is also constant, whereas the electron mobility is inversely proportional to the plasma density. Three values for the electron temperature has been chosen in order to evaluate the electric transport properties. The values of the free stream electrical conductivity and electron mobility corresponding to the condition chosen for calculation are shown in Table 5

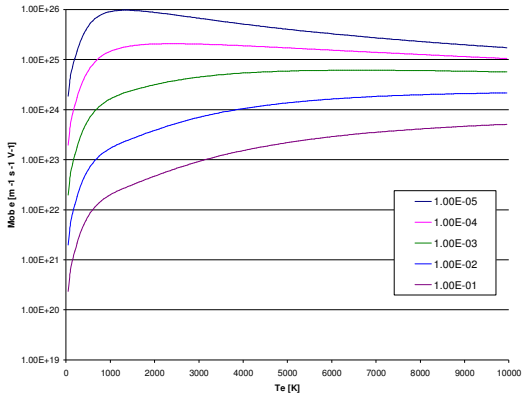


Fig. 12. $n_{Ar} \times \mu_e$ as a function of electron temperature and ionization degree

Calculations have been initially carried out on the domain shown in Fig. . The domain is discretized by means of a 26032 points mesh. In Fig. 4, the computed pressure distribution around the test body is reported. The numerical rebuilding reproduced qualitatively the flow field (Fig. 13), and some of the flow patterns revealed by fast imaging appear to be correctly replicated (Fig. 14). In Fig. 15 the Faraday component (i.e. the azimuthal component) of the current density is reported.

Table 5. Free stream electric transport properties

Case	T_e	σ	$\mu_{e,\infty}$
1	[eV]	[S m^{-1}]	[m^2 s^{-1} V-1]
2	0.221	493	120
3	0.086	262	52
	0.052	135	24

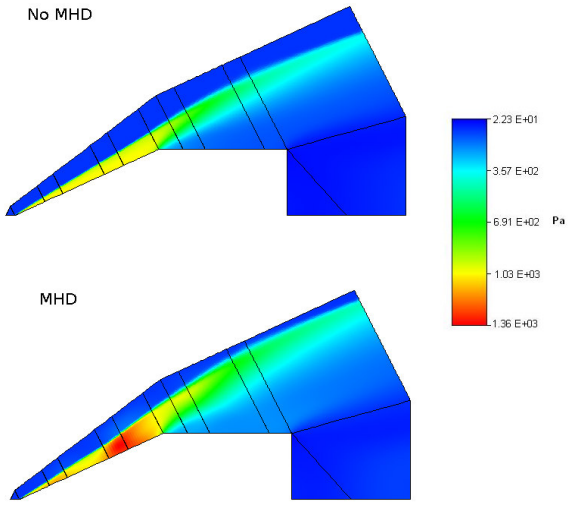


Fig. 13. Computed pressure field without and with MHD interaction

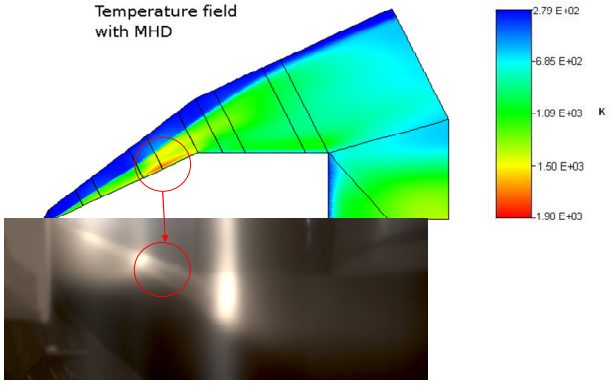


Fig. 14. Comparison of flow structure details between simulation and fast imaging (right).

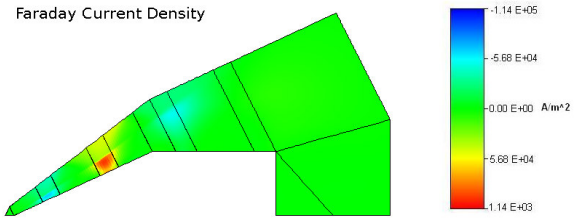


Fig. 15. Azimuthal current density

In Fig. 16, a very good agreement between the measured and calculated pressure value can be observed at position 1 for the set C of electric transport properties. Also the calculated voltage

(Fig. 16) gets close to the measurements; however, a comparable agreement is not reached. This points out the need of a refinement in the modelling of the phenomenon. A plasma kinetics model would take into account the electron behaviour and the ionization processes, providing a more realistic description of the electric transport properties of the plasma. A comparison between the measured and calculated values of pressure on the test body wall is reported in Table 6

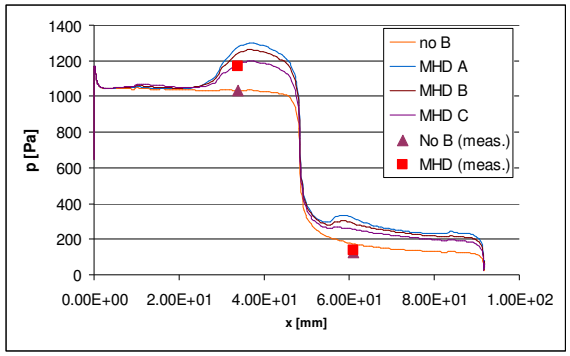


Fig. 16. Pressure on the body wall

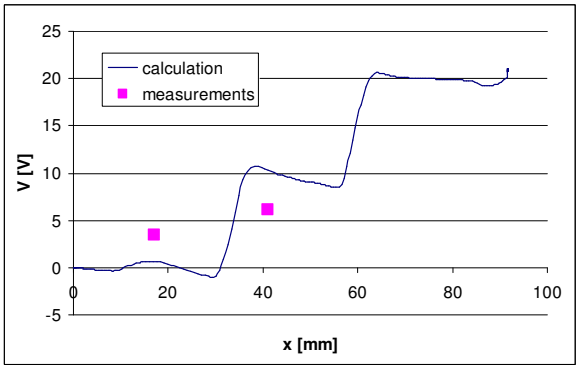


Fig. 17. Voltage on the body w all

Table 6. Comparison between calculated and measured values of the pressure

	Meas.	Calc.	Diff. %	Exp. Uncert. %
P pos. 1 [Pa] (no MHD)	1031	1030	0.10	0.3
P pos. 1[Pa] (MHD)	1168	1178	0.86	4.3
P pos. 2 (no MHD) [Pa]	125.5	168	33.86	26.4
P pos. 2 (MHD) [Pa]	133	229	91.62	47.7

7 Conclusion

The MHD interaction in a hypersonic flow was experimentally investigated around a conical test body. The electrical-connection scheme allowed short-circuiting of the Faraday current inside the plasma, enhancing the effect of the MHD interaction. Pressure, imaging, and electrical observations were done. A pressure increases on the conical section of the test body caused by the MHD interaction has been observed. The pressure increase in this experiment was between 7% and 15%.

The images around the test body showed a rather different flow field in the tests without and with MHD interaction. The Hall field was measured by means of electrical probes placed along the test body. The measured values of the Hall field were in the range between 400 and 500 V/m.

A rebuilding of a MHD experiment has been also presented. The analysis has been performed utilizing a low Re_m approximation. The result obtained seem to be in a reasonably agreement with the experimental data available. The Hall parameter substantially affects the electrodynamics in the plasma and, as a result, the whole MHD interaction around the test body. Evidences has been found of a significant influence of the calculation domain on the results obtained. An improvement of the analysis is expected including the plasma kinetics modeling for a more accurate determination of the electric transport parameter in the flow.

Acknowledgments

The work described in this paper was carried out thank to the support of the European Space Agency (ESA) under the HPF contracts.

References

[1] Ziemer, R.W., "Experimental Investigation in Magneto Aerodynamics", *American Rocket Society Journal*, Vol. 29, No. 9, 1959, pag. 642.

[2] V.A. Bityurin, et al., "Assessment of Hypersonic MHD Concepts", *28th Plasmadynamics and Lasers Conference*, Paper AIAA-97-2393, 1997.

[3] J.T. Lineberry, et al., "Prospects for MHD Flow Control for Hypersonics", *31st Plasmadynamics and Lasers Conference*, Paper AIAA-2000-3057, 2000.
[4] S.O. Macheret, et al., "MHD Power Extraction from Cold Hypersonic AirFlow with External Ionization", *30th Plasmadynamics and Lasers Conference*, Paper AIAA-99-4800, 1999.
[5] D.J. Moorhouse and C.F. Suchomel, "Exergy Methods Applied to Hypersonic Vehicle Challenge", *32sd Plasmadynamics and Lasers Conference*, Paper AIAA-2001-3063, 2001.
[6] J.S. Shang, et al., "Magneto-Aerodynamic Interactions in Weakly Ionized Hypersonic Flow", *40th Aerospace Sciences Meeting & Exibit*, paper AIAA 2002-0349, Reno, NV, 2002.
[7] J.S. Shang, "Recent Research in Magneto-Aerodynamics", *Progress in Aerospace Sciences*, Vol. 37, pp. 1-20, 2001.
[8] E.P. Gurianov and P.T. Harsha, "AJAX: "New Directions in Hypersonic Technology", *27th Plasmadynamics and Lasers Conference*, Paper AIAA-96-4609, 1996.
[9] V. Bityurin, A. Bocharov, D. Baranov, and J. Lineberry, "Experimental and Theoretical Study of MHD Interaction in Hypersonic Ionized Air Over a Wedge", 6- 9 Janaury 2004, Reno, Nevada, paper AIAA 2004-1195
[10] J.S. Shang , J.Hayes,J. Miller, J. Menart, "Magneto Hydrodynamic Interaction in Weakly Ionized Hypersonic Flow", *40th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, Jan. 14-17, 2002, paper AIAA 2002-349
[11] Y. Takizawa, K. Matsushita, S. Sato, K. Fujita, T. Abe, and D. Konigorski, "Electro–Magnetic Effect on Reentry-related High-enthalpy Flow", *34th AIAA Plasmadynamic and Laser Conference*, 23–26 June 2003, Orlando, Florida, paper AIAA 2003-4168.
[12] Matsuda, A., Wakatsuki, K., Takizawa, Y., Kawamura, M., Otsu, H., Konigorski, D., Sato, S., and Abe, T., "Shock Layer Enhancement by Electro-Magnetic Effect for Spherically Blunt Body", *37th Plasmadynamics and Lasers Conference*, AIAA Paper 2006-3573, San Francisco, June 2006.
[13] A. Gülhan, B. Esser, U. Koch, et al., "Experimental Verification of Heat-Flux Mitigation by Electromagnetic Fields in Partially-Ionized-Argon Flows", *Journal of Spacecraft and Rockets,* Vol. 46, No. 2, March–April 2009
[14] L. Biagioni, C.A. Borghi, M.R. Carraro, A. Cristofolini, A. Fantoni, and A. Passaro, "Magneto-Hydrodynamic Interaction in the Shock Layer of a Ramp in Hypersonic Flow", *43rd AIAA Aerospace Science meeting and Exhibit*, 10-13 Janaury 2004, Reno, Nevada, paper AIAA 2004-2265.
[15] F.Scortecci, F. Paganucci, L. d'Agostino, M. Andrenucci, "A New Hypersonic High Enthalpy Wind Tunnel", AIAA Paper 97-3017, July 1997.
[16] F. Scortecci, F. Paganucci, L. d'Agostino, Experimental Investigation of Shock-"Wave/Boundary-Layer Interactions Over an Artificially Heated Model in Hypersonic Flow", AIAA Paper 98-1571, April 1998.
[17] A. Passaro, L. Biagioni, and L. d'Agostino, "Compressibility effects in weakly turbulent hypersonic flow fields", in Proc. *4th Eur. Symp. Aerothermodynamics Space Vehicles*, Caserta, Italy, Oct. 2001, pp. 709–716.
[18] K. Sutton and R. A. Graves, "A general stagnation-point convective heating equation for arbitrary gas mixtures", NASA, Washington DC, NASA TR R-376, Nov. 1971.
[19] Marco A. Gigos et al., "New plasma diagnosis tables of hydrogen stark broadening including ion dynamics", J. *Phys. B: At. Mol. Opt. Phys.,* 29:4795-.4838, 1996.
[20] H. R. Griem, "Principles of Plasma Spectroscopy", *Cambridge University Press*, 2001.
[21] M.A. Heald and C.B Wharton, "Plasma diagnostic with microwaves", *John Wiley & Sons, 1965*
[22] C.A. Borghi, M.R. Carraro, and A. Cristofolini, "Analysis of Magnetoplasmadynamic Interaction in the Boundary Layer of a Hypersonic Vehicle", *Journal of Spacecraft and Rockets*, Vol. 42, n. 1, pp. 45-50, January 2005.
[23] M. Mitchner, C.H. Kruger, "Partially Ionized Gases", *Wiley-Interscience*, New York, 1973

CEAS 2011

Turbines, Fans and Compressors

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Improvements in performance estimation of wind turbine blade airfoils at low Reynolds number

Elia Daniele, Fabrizio Nicolosi and Domenico P. Coiro
Department of Aerospace Engineering, Università degli Studi di Napoli "Federico II"
Via Claudio, 21 – 80125 Napoli, Italy

Abstract

In this paper we propose a comparison between two Reynolds–Averaged–Navier–Stokes (RANS) turbulence models as Spalart–Allmaras (S–A) and the more recent transitional k–ω SST in descibing low Reynolds number laminar bubble bursting stall for three different airfoils: a NACA 0012 used as reference airfoil with a well known behaviour and a NACA 0018, widely used on many vertical–axis–wind–turbine (VAWT), of which TU Delft recent experimental data al low Reynolds number are available. Finally an extension to higher Reynold number is proposed with an analysis on S809 airfoil. Emphasys is posed on different turbulence models desciption of experimental physical condition such as turbulence intensity and flow field description of laminar separation inside bubble.

1 Introduction

In this paper we propose a comparison between two Reynolds–Averaged–Navier–Stokes (RANS) turbulence models as Spalart–Allmaras (S–A [1]) and the more recent transitional k–ω SST (in the following indicated as TR k–ω SST [2, 3]) in trying to predict low Reynolds number laminar bubble bursting stall for three different airfoil.

- NACA 0012 airfoil with a well known behaviour ([4, 6, 5, 7]);
- NACA 0018, widely used on many vertical–axis–wind–turbine (VAWT), of which recent experimental data at low Reynolds number are available [8] by TU Delft;
- NREL S809 airfoil ([9, 10, 11, 12]), typically used on stall controlled horizontal–axis–wind–turbine (HAWT) and characterized by a "flat" stall.

Such calculations would provide a survey of advantages and limitations of low turbulence and low Reynolds numerical simulations in describing both low angle of attack bubble dominated drag coefficient regime and laminar bubble bursting stall for which several investigations based on interactive boundary layer panel methods ([13, 14, 15, 16]) have been carried out in past years. The flow conditions under investigation represent the main working conditions for small–size wind turbines, especially for those which are stall controlled.

Emphasys is posed on different turbulence models capability to descipt the real condition of experimental test, such as turbulence intensity, and flow field description of laminar separation inside bubble.

Concerning the NACA 0012, many numerical analysis have been performed among which: Marongiu and Tognaccini [4] have conducted 2–D steady RANS simulations using the Menter SST turbulence model to investigate the flow at a Reynolds number equal to 135000; Shur *et alii* [7], have made a Detached–Eddy simulation at a Reynolds number of 100000; Alrefai [6] have reported some static pressure measurements at a Reynolds number of 110000 prior to suction control investigation made in the Andrew Fejer

wind tunnel at the Illinois Institute of Technology's Fluid Dynamics Research Center; Lee and Gerontakos [5] have made static pressure measurements at a Reynolds number of 135000 as a starting point of their oscillating airfoil investigations in the low–speed, suction–type wind tunnel in the Aerodynamics Laboratory at McGill University.

Regarding the NACA 0018 airfoil, Timmer [8] has shown experimental results for lift and drag coefficients with and without forced transition in the low–turbulence TU Delft wind tunnel for a Reynolds number between 150000 and 100000.

Concerning the NREL S809 airfoil, Wolfe and Ochs [11], have performed 2D steady CFD simulations using a commercial code that solves the Favre-averaged Navier Stokes equations with finite volume approach. Their approach consists in splitting the computational domain in two regions: a fully turbulent region with a standard k/ε turbulence model and a fully laminar one confined only in a limited portion. In this work an analysis at a Reynolds number of 2.0×10^6 for a range of angle of attack comprises between 0 and 20 degrees has been made. Furthemore, Guerri *et alii* [12] perform numerical computations using an unsteady incompressible RANS solver comparing two different turbulence models as the SST k–ω of Menter and the RNG k/ε assuming in both cases a fully turbulent flow. In this work the flow field is also analyzed at several angles of attack from 0 to 20 degrees considering the influence of the free stream turbulence intensity on results for a Reynolds number of 2.0×10^6. Experimental investigation on this airfoil has been conducted by Somers [10] in the low–turbulence wind tunnel facility of Delft TU for a Reynolds number equal to 1.0×10^6, and is used as a comparison for our numerical results.

The numerical methods on which is based our work is detailed in Section 2. A first validation using NACA 0012 airfoil experiments is introduced in Section 3.1NACA 0018 follows in Section 3.2 and extension to different Reynolds number is made with S809 analysis in Section 3.3 and a final discussion is provided in the concluding Section 4.

2 Numerical Methods

The needs of investigating low Reynolds number flow, such that the flowfield is mostly laminar or transitional, has lead us to take in consideration the use of a transition model.

Even if with the actual S–A and k–ω SST models it is not possible to predict the transition point to agree with the real physical transition, some recommendations for freestream turbulence levels have been made ([17], [18]). Such suggestions were not based upon matching freestream turbulence levels from wind tunnel or flight, but rather upon considerations related to preserving potential cores and maintaining the integrity of the turbulence quantities throughout the boundary layers .

Our study takes as starting point the possibility to achieve a fair agreement between the S–A and k–ω SST turbulence methods acting on *Turbulence Intensity* for k–ω SST for the freestream boundary condition.

The Spalart Allmaras method ([1]) in its standard form has been used for the calculations shown in this work, modifying the *Turbulence Intensity* to very low values, such that a bubble formation, separation and transition to turbulence flow condition could be observed; these value are typically lower than $10e^{-10}$.

The *Turbulence Intensity* values used in the approach of Menter (Menter-2006,Langtry-2006) are related strictly to the one forseeable in wind tunnel testing. This method of predicting transition is based upon experimental correlations that usually relate the free-stream turbulence intensity and the local pressure gradient to the transition momentum thickness Reynolds number, defined as $\mathrm{Re}_\theta = \rho\theta \mathrm{U}_0/\mu$, where θ is the momentum thickness and U_0 the local freestream velocity. Such task is even more complicated in modern CFD methods based on non structured grid (used to approximate complex geometries) and parallel calculation, in wich the information must be shared on different machine, since the integration of boundary layer parameters would result very difficult.

As detailed in [19, 20], the transition model is based on two transport equations: a transport

Table 1 Mesh details

Parameter	Value
Mesh type	C–shape
Mesh size (–)	180493
Airfoil nodes (–)	510
Far–Field nodes (–)	908
Outlet nodes (–)	399
Chord length (m)	0.058
Far–Field — Airfoil distance (unit chord)	30
Outlet — Airfoil distance (unit chord)	30

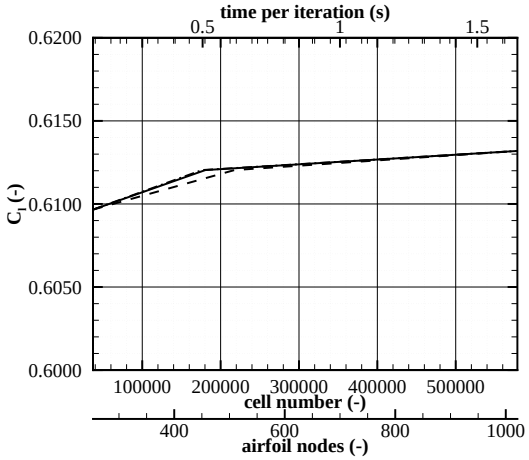
equation for intermittency, used to trigger transition locally; and a second transport equation that computes the transition onset momentum–thickness Reynolds number. The essential part of the model lies in the second transport equation that links the empirical correlation to the onset criteria in the intermittency equation, coupled with SST k–ω model of Menter ([21]).

The analysis on all treated airfoils has been performed generating a C–shape structured mesh whose characteristics are given in table 1. To ensure a full developed wake behind the airfoil and the respect of a far–field condition on the upper and lower outer boundary we have imposed a distance of 30 c (i.e. chords) from leading edge and trailing edge in horizontal direction, and again 30 c from chord line in vertical direction.

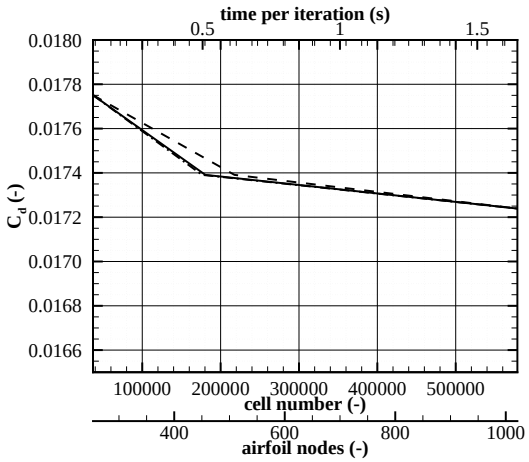
The height for the first cell layer is prescribed imposing $y^+ \simeq 1$ near wall for a correct treatment of boundary layer with the different turbulence models used.

The actual size of the mesh has been choosed after a sensitivity analysis of aerodynamic coefficient (such as the lift coefficient C_l) respect to mesh nodes number on the airfoil, and equivalently the mesh cells number.

In figure 1 can be observed the trend of force coefficients for NACA 0012 airfoil with respect to nodes/cell number, and then the choosen mesh size appears motivated as a compromise between field solution accuracy and computation time. The calculation has been performed on a PC with Intel®Core^{TM}i7 CPU 920 2.67 GHz and 12 GB of RAM. In addition, the specific value for chord length is such to get a value of Reynolds number



(a) C_l sensitivity to mesh size



(b) C_d sensitivity to mesh size

Fig. 1 NACA 0012 Mesh sensitivity analysis:
— dependence on cell number,
−− dependence on airfoil nodes,
· − · dependence on time per iteration.

Re = 135000 with value of air density (ρ) and molecular viscosity (μ) computed at a temperature of $T = 288\,\mathrm{K}$.

The slope of lift coefficient C_l with respect to mesh cell number when passing from the medium to big size mesh is $\frac{\partial C_l}{\partial cell\ number} \simeq 2.5e^{-9}$.

3 Results

3.1 NACA 0012

The turbulence models adopted for our analysis are Spalart–Allmaras (S–A) [1] and the more recent transitional k–ω SST ([2, 3]).

In both cases boundary condition for farfield

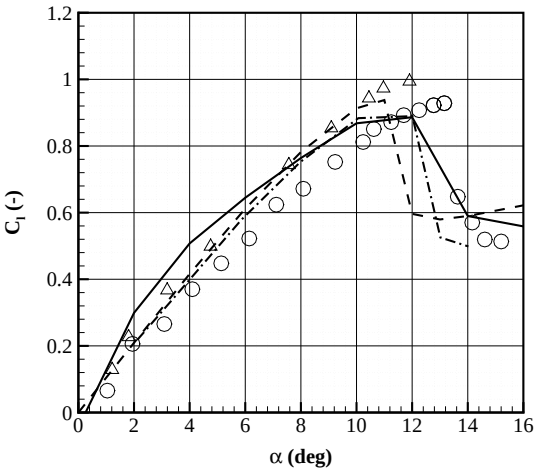
and outlet is set to Pressure-Far–Field and Pressure-Outlet, the Mach and Reynolds number are set to $M = 0.1$ and $\mathrm{Re} = 135000$, the solver type is of coupled–implicit type with gradient evaluated with least square cell method and each equation is discretizated at second order. The convergence criteria is based on equations residual of $10e^{-7}$, and in case of residual oscillation a mean value on 40000 iterations has been considered.

The main difference between the two turbulence methods consists in the definition of turbulence parameters: in TR k–ω SST we have intermittency set to a constant value equal to 1 while in S–A this quantity isn't present; moreover, in TR k–ω SST a more physical value for turbulent intensity, set to 1% (closer to wind tunnel value), has been considered, while in S–A the same value has been lowered to $10e^{-14}$; finally the turbulent length scale has been set equal to chord length, i.e. 0.058m for both methods.

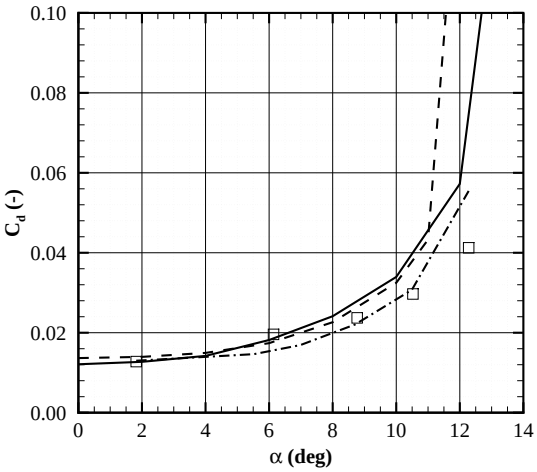
The results obtained can be compared to those of a two–dimensional $k--\omega$ *SST RANS* simulation code at $\mathrm{Re} = 135000$ ([4]), to experimental test data at $\mathrm{Re} = 110000$ ([6]), to experimental data at $\mathrm{Re} = 135000$ ([5]) and to a *Detached Eddy Simulation* (DES) case at $\mathrm{Re} = 100000$ ([7]). We specify that experiments reported in [5] seems to be almost 3D showing a possible separation close to the end parts of the model tested. The comparisons are shown in figure 2.

In figure 2(a) a greater surplus of lift at the low angle of attack range can be noted in the TR k–ω SST curve respect to both S–A and the Marongiu [4] results, while the stall part of the curve is more similar for both methods. The value of maximum C_l appears to be the same for both the TR k–ω SST and the k–ω SST of Marongiu [4], while the S–A predicts a C_l value greater than ~ 0.05 with respect of the previous one and an increase even greater than ~ 0.05 is obtained in the experiment of Alrefai [6]. The stall angle value is again similar between the TR k–ω SST and the k–ω SST of Marongiu [4], i.e. 12 deg, while the S–A method returns a value ~ 1 deg less; in all cases the post–stall decrease of C_l starts about $1 \div 2$ deg after the maximum C_l angle.

In figure 2(b) one could observe a value for drag coefficient (C_d) of TR k–ω SST and S–A respect to Marongiu [4], while stall behavior clearly reflects the lift coefficient trend; at low angle of attack regime very little differences could be noted respect to DNS of Shur [7], but up to even a 20% difference would results at stall angle, where clearly the different Reynold number plays a major role.



(a) C_l VS α



(b) C_d VS α

Fig. 2 NACA 0012 characteristics at $\mathrm{Re} = 135000$**.(a): −− S–A , · − · Marongiu [4], — TR k–ω SST, △ Alrefai [6], ○ Lee [5].(b): −− S–A , · − · Marongiu [4], — TR k–ω SST, □ Shur [7].**

3.2 NACA 0018

Thus, for NACA 0018, in S–A turbulence model simulation, a *Turbulence Intensity* equal to $1e^{-14}$ and $1e^{-10}$ has been choosed for Reynolds number Re = 150000 and 300000. For TR k–ω SST turbulence model the values of *Turbulent Length Scale* is now 0.25 for both Reynolds number Re = 150000 and 300000.

The numerical results can be compared to the experiments presented in [8], and are shown in figures 3 and 4: the experimental data appearing in the mentioned figures heve been captured directly from the graphics of reference [8].

For a Reynolds number of Re = 150000 the results of two turbulence method show differences in stall prediction.

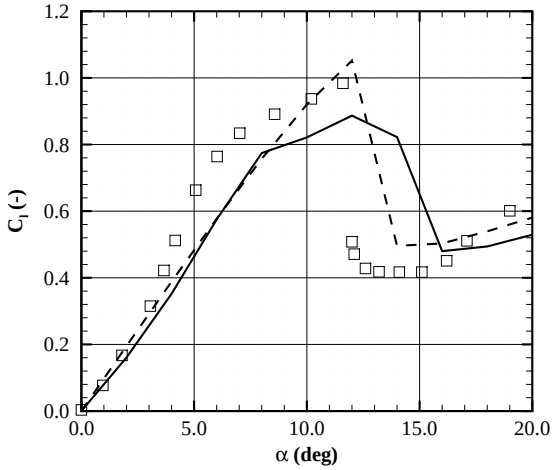
In figure 3(a) one could observe that the S–A turbulence model describes a strong decrease of lift coefficient C_l at an angle of attack $\alpha \in [13, 14]$ deg, very similar to the behavior obtained in the experiment, with a difference in angle of attack of stall onset of about $1 \div 2$ deg. The TR k–ω SST turbulence method fails to describe the abrupt reduction of lift coefficient, even if it predicts a maximum lift coefficient slightly lower (about 0.1) than the experimental one, while, for both low and high values of angle of attack, both turbulence method match the experimental data closely, except the interval $\alpha \in [4, 8]$ deg where seems that a large bubble is present in wind tunnel test, resulting in a local lift coefficient increase.

The difference underlined for lift coefficient are reflected clearly in drag coefficient, C_d (see figure 3(b)), thus a trend closer to experimental data could be observed for S–A model than for TR k–ω SST, that seems to give lower values of C_d in post–stall region.

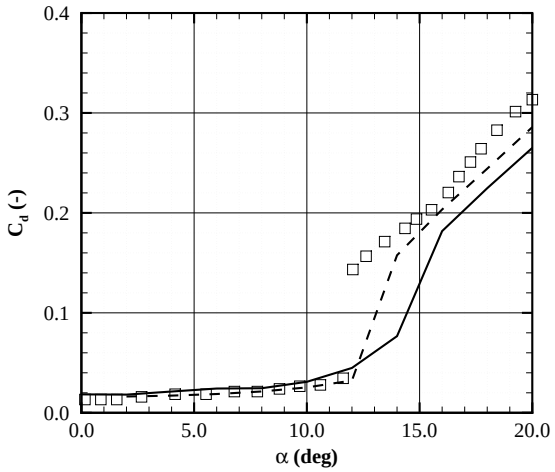
The behavior of moment coefficient respect to the quarter of chord, $C_{m,c/4}$ (see figure 3(c)), is better described by S–A turbulence model, even if both methods capturing well the overall trend.

For a Reynolds number of Re = 300000 again the results of two turbulence method show differences in stall behavior prediction.

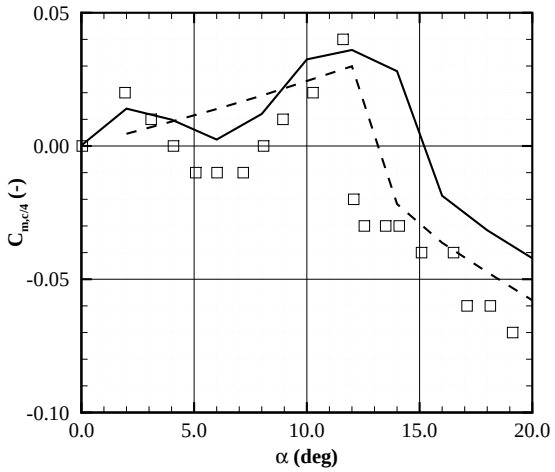
In figure 4(a) one could observe that the S–A turbulence model describes a stronger decrease of lift coefficient C_l at an angle of attack $\alpha \in$



(a) C_l VS α

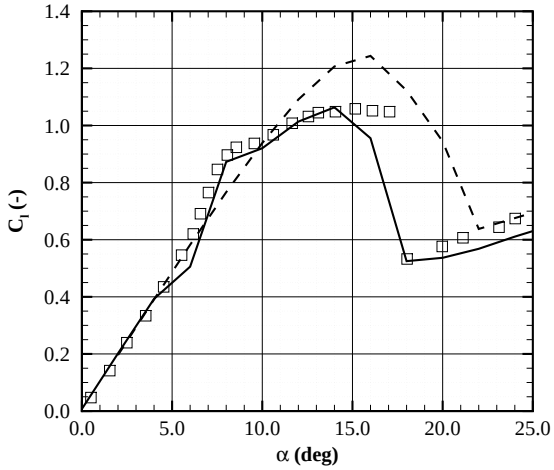


(b) C_d VS α

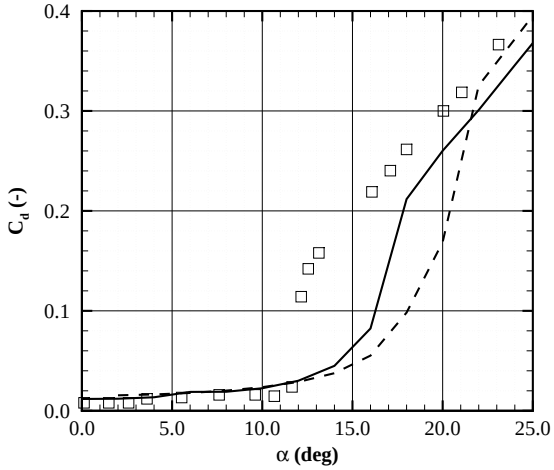


(c) $C_{m,c/4}$ VS α

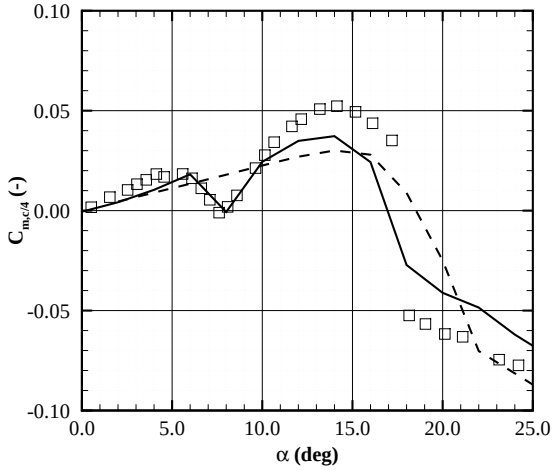
Fig. 3 NACA 0018 characteristics at Re = 150000**.(a), (b), (c): −− S–A , — TR k–ω SST, □ Timmer [8].**



(a) C_l VS α



(b) C_d VS α



(c) $C_{m,c/4}$ VS α

Fig. 4 NACA 0018 characteristics at Re $= 300000$**.(a), (b), (c): −− S–A , — TR** k**–**ω **SST, □ Timmer [8].**

$[16,18]$ deg respect to the behavior obtained with the experiment, with a difference in stall onset angle of attack of about 1 deg and maximum lift coefficient overestimated of $\sim 20\%$. The TR k–ω SST turbulence method describes a less abrupt reduction of lift coefficient, even if it predicts a maximum lift coefficient almost coincident (about 1.05) with the experimental one but with a difference in the angle of attack for maximum lift coefficient of about 1 deg. Both turbulence method are close to the experimental data in post stall region, while in the interval $\alpha \in [6,10]$ deg again seems that a large bubble is present in wind tunnel test, resulting in a lift coefficient increase respect to both numerical results.

For drag coefficient, C_d (see figure 4(b)), there is a good agreement with experimental data in low angle of attack range ($\alpha \in [0,10]$ deg), followed by a great difference in range $\alpha \in [11,18]$ deg, although both turbulence methods are here in good agreement.

The behavior of moment coefficient respect to the quarter of chord, $C_{m,c/4}$ (see figure 4(c)), is described in similar manner by both turbulence methods in the range $\alpha \in [0,16]$ deg, except for the region $\alpha \in [6,10]$ deg where a laminar bubble appears as noticed in figure 4(a). In post–stall region the S–A turbulence model is closer to experimental results than TR k–ω SST that showes a higher (positive) value for $C_{m,c/4}$.

3.3 S809

For the S809, in S–A turbulence model simulation, a *Turbulent Intensity* equal to $1e^{-40}$ has been choosen for Reynolds number Re $=$ 1e6. For TR k–ω SST turbulence model the values of *Turbulent Intensity* is *Turbulent Length Scale* are now, respectively, 0.03 and 0.6 for Reynolds number Re $=$ 1e6.

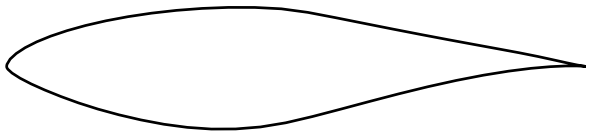
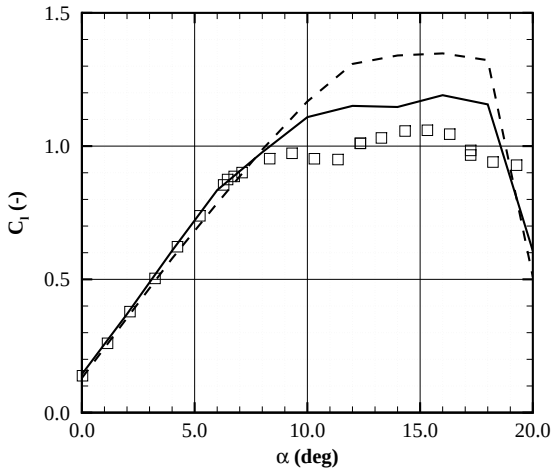
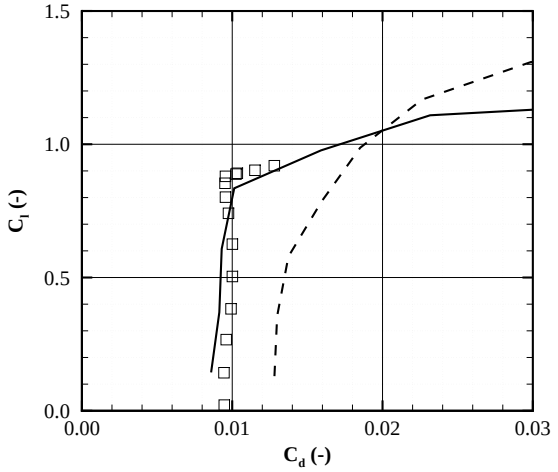


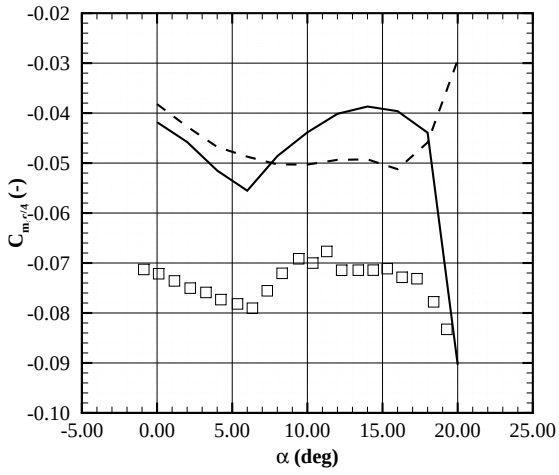
Fig. 5 NREL S809 airfoil.



(a) C_l VS α



(b) C_d VS α



(c) $C_{m,c/4}$ VS α

Fig. 6 S809 polars at Re = 1000000.
(a), (b), (c): $- - - -$ $S- -A$, **— TR** k–ω **SST, □ Somers [10].**

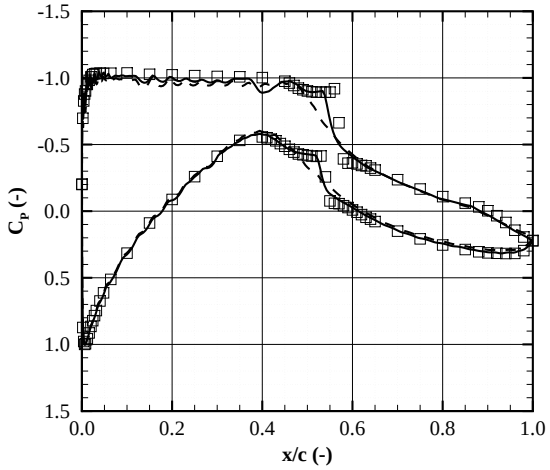
The numerical results are compared to the experiments presented in [10], as shown in figure 6: the experimental data appearing in the figures have been extrapolated through graphical procedure from the reference [10].

For a Reynolds number of Re = 1e6 the results of the two turbulence models show significant differences in stall prediction.

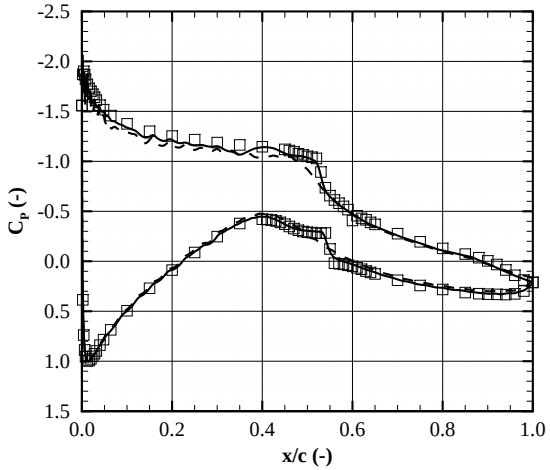
In figure 6(a) it can be observed that both turbulence models estimate correctly the linear part of lift coefficient curve departing from experimental values at an angle of attack of about 8 deg. Neither of them is capable of capturing well the beginning os separation indicated by the *kink* appearing in the $C_l(\alpha)$ curve at 8 deg and both overestimate the value of $C_{l,max}$ of 0.1 for the S–A and 0.25 for TR k–ω SST. The difference in stall angle of attack is for both models equal to about 1 deg greater than that measured in experiments and equal to $\alpha = 15$ deg. The C_l trend in post–linear regime is better captured by the TR k–ω SST model that forsees a two–stage stall behavior as in the experimental curve. On the contrary, the S–A doesn't show a two–region stall curve, but a single bell–shaped curve always greater than that of TR k–ω SST model.

The differences underlined for lift coefficient are reflected clearly in drag coefficient, C_d (see figure 6(b)), thus a trend closer to experimental data can be observed for TR k–ω SST model respect to S–A, while seems to predict greater values of C_d in pre-stall region. However, the C_d experimental data available from [10] describe only the pre-stall region without giving information about what happens at higher angle of attack. Anyway, even in this low angle of attack range it can be observed that the TR k–ω SST tends to approximate the experimental data with a difference that lies in about 20 counts, while the S–A model fails completely the description of the drag coefficient since it predict C_d values while are almost greater of 25 counts than experimental data throughout the range.

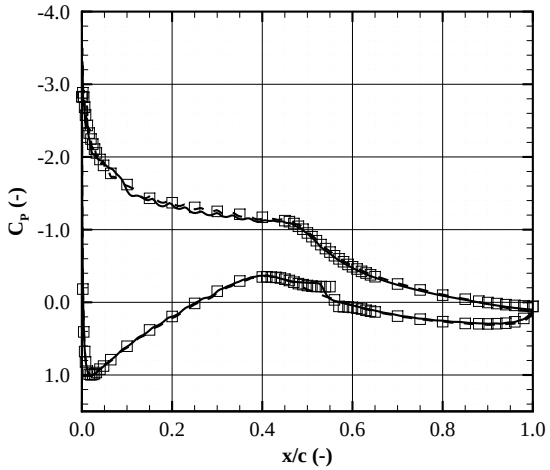
The moment coefficient values are not captured by neither methods with differences respect to experiments, up to 40%. However TR k–ω SST method is capable of predicting the general behavior of $C_{m,c/4}$ curve indicating the capability



(a) $\alpha = 4\,\mathrm{deg}$



(b) $\alpha = 6\,\mathrm{deg}$



(c) $\alpha = 8\,\mathrm{deg}$

Fig. 7 **S809** C_P **at** $\mathrm{Re} = 1000000$**.**
(a), (b), (c): $- - - -$ $S- -A$**, — TR** k**–**ω **SST, □ Somers [10].**

of predicting the main physical properties.

The pressure coefficient, C_P, distribution could be compared to experimental results of [10] only into the range of angle of attack relative to a pre-stall behavior, whit an almost linear lift coefficient vs angle of attack. Asit can noticed in figure 7 some comparisons have been made for $\alpha = 4, 6, 8\,\mathrm{deg}$ for the two aforementioned turbulence methods, of which the TR k–ω SST showes a closer agreements respect to S–A model especially in the area where bubbles are present.

In particular, in figure 7(a) and (b) it can be observed a quite accettable description of TR k–ω SST on C_P sudden variation at the reattachment after laminar separation bubble, while the S–A fails to capture this local variation, failing to describe the presence of laminar separation bubble. The two different turbulence models give very similar results in the 8 deg case where the bubble has moved on leading edge, as depicted in figure 7(c).

4 Conclusions

For the cases investigated in the present work, it can be asserted that both turbulence methods are able to describe fairly well the deep stall regime, where the pressure contribution to aerodynamic forces dominates the skin friction one, i.e. the one that can be more reasonably affected by laminar bubble separation and turbulent reattachement. In the range of angle of attack where more complex phenomen appear, i.e. the presence of a laminar separation bubble and a turbulence reattachement, the discrepancies between the two different turbulence models and the experimetal results appear to be interested by a wider spread than those forseen in the deep stall regime (see e.g. figures 3(a), 4(a), 6(c)).

However, as it has been shown for NREL S809 airfoil, the TR k–ω SST is able to predict the distribution of pressure coefficient C_P for a wide range of angle of attack in a more proper manner than the S–A model. This circumstance justifies the greater adherence of $C_{m,c/4}$ computed by TR k–ω SST to experimental results, since the loss of capability in describing C_P affects more the moment coefficient (due to the arm of lift coef-

Table 2 Summary of qualitative prediction capability for both turbulence methods for all test case

Airfoil Condition	Method	linear range			stall range			post stall range		
		C_l	C_d	$C_{m,c/4}$	C_l	C_d	$C_{m,c/4}$	C_l	C_d	$C_{m,c/4}$
NACA 0012	S–A	good	good	–	poor	poor	–	good	–	–
Re=1.35e5	TR k–ω SST	poor	good	–	poor	poor	–	good	–	–
NACA 0018	S–A	good	poor	poor	poor	poor	poor	good	good	good
Re=1.5e5	TR k–ω SST	good	poor	good	poor	poor	good	good	good	good
NACA 0018	S–A	poor	good	poor	poor	poor	good	good	good	good
Re=3e5	TR k–ω SST	good	good	good	good	poor	good	good	good	good
S908	S–A	good	poor	poor	poor	poor	poor	poor	–	poor
Re=1e6	TR k–ω SST	good	good	poor	poor	good	poor	good	–	poor

ficient) than C_l. Despite the confident predictions made for lift and moment coefficient among all airfoil tested, there is a non negligible difference in drag coefficient C_d prediction, most of all in low angle of attack regime. As it has been shown for NREL S809 airfoil, the drag coefficient predicted by both turbulence methods do not present the laminar bucket observed in experimental results, so very large differences are to be expected for airfoil that are interested by a laminar separation bubble, while a better description is obtained for both NACA airfoil 0012 and 0018.

Finally, it should be taken in account that for the two different turbulence methods are presented different turbulence parameter settings. In order to accomplish an accetable prediction of such low Reynolds number flow conditions, the TR k–ω SST has a more meaningful relation to actual turbulence intensity than S–A, that instead has been used throughout our analysis in a very low turbulence meaning by reducing the *Turbulent Intensity* or the *Turbulence Viscosity Ratio* to very little value. In Table 2 a summary of the qualitative prediction capability of both turbulence models for all test cases analyzed is presented. The great advantage of TR k–ω SST seems to be the capability of stall mechanism prediction, while only nebulous considerations can be made on the actual ability of both turbulence methods to predict the maximum lift coefficient value and the behavior of drag coefficient in laminar bucket.

References

[1] Spalart, P. R. and Allmaras, S. R., "A One-Equation Turbulence Model for Aerodynamic Flows", *Recherche Aerospatiale*, No. 1, pp. 5–21, 1994.

[2] Menter, F. R., Langtry, R. B., Likki, S. R., Suzen, Y. B., Huang, P. G. and Völker, S., "A Correlation–Based Transition Model Using Local Variables – Part I: Model Formulation", *Journal of Turbomachinery*, Vol. 128, pp. 413–422, 2006.

[3] Langtry, R. B., Menter, F. R., Likki, S. R., Suzen, Y. B., Huang, P. G. and Völker, S., "A Correlation–Based Transition Model Using Local Variables – Part II: Test Cases and Industrial Applications", *Journal of Turbomachinery*, Vol. 128, pp. 423–434, 2006.

[4] Marongiu, C. and Tognaccini, R., "Simulation of Dynamic Stall at Low Reynolds Number", *48th AIAA Aerospace Science Meeting Including the New Horizons Forum and Aerospace Exposition*, Orlando, Florida, 2010.

[5] Lee, T. and Gerontakos, P., "Investingation of flow over an oscillating airfoil", *Journal of Fluid Mechanics*, Vol. 512, pp. 313–341, 2004.

[6] Alrefai, M. and Acharya, M., "Controlled Leading–Edge Suction for Management of Unsteady Separation over Pitching Airfoils", *AIAA Journal*, Vol. 34, No. 11, pp. 2327-2336, 1996.

[7] Shur, M., Spalart, P. R., Strelets, M., Travin, A., "Detached–Eddy Simulation of an airfoil at

high angle of attack", *Engineering turbulence modelling and experiments*, Vol. 4, pp. 669–678, 1999.

[8] Timmer, W.A., "Two–dimensional low–Reynolds number wind tunnel results for airfoil NACA 0018", *Wind Engineering Volume*, Vol. 32, No. 6, pp. 525–537, 2008.

[9] Tangler, J., Somers, D. M., "NREL Airfoil Families for HAWTs", *AWEA*, pp. 1-12, 1995.

[10] Somers, D. M., Tangler, J., "Design and Experimental Results for the S809 Airfoil", *NREL/SR–440–6918 UC Category:1213 DE97000206*, pp. 1-104, 1997.

[11] Wolfe, P. W., Ochs, S. S., "CFD Calculations of S809 Aerodynamic Characteristics", *AIAA–97–0973*, pp. 1–8, 1997.

[12] Guerri, O., Bouhadef, K., Harhad, A., "Turbulent Flow Simulation of the NREL S809 Airfoil", *Wind Engineering*, Volume 30, No. 4, pp. 287–302, 2006.

[13] Drela, M. and Giles, M. B., "Viscous–Inviscid Analysis of Transonic and Low–Reynolds Number Airfoils", *AIAA Journal*, Volume 25, No. 10, pp. 1347-1355, 1987.

[14] Coiro, D.P., de Nicola, C., "Prediction of Aerodynamic Performance of Airfoils in Low Reynolds Number Flows", *Low Reynolds Number Aerodynamics*, Proceedings Conference, Notre Dame, Indiana, U.S.A., June 1989, Springer Verlaag.

[15] Dini, P., Coiro, D.P. and Tangler, J., "Effects of Laminar Separation Bubbles and turbulent Separation on Airfoils Stall: Work in Progress", *Windpower '92 Proceedings*, pp. 339–346, 1992.

[16] Coiro, D.P., Dini, P., "The Prediction of Aerodynamic Stall Over Wind–Turbine Airfoils", *International Conference on Fluid and Thermal Energy Conversion '97*, Yogyakarta (Indonesia).

[17] Rumsey, C.L. and Spalart, P. R., "Turbulence Model Behavior in Low Reynolds Number Regions of Aerodynamic Flowfields", 38th AIAA Fluid Dynamics Conference and Exhibit, Seattle, WA, June 23 – 26, 2008.

[18] Spalart, P. R. and Rumsey, C.L., "Effective Inflow Conditions for Turbulence Models in Aerodynamic Calculations", *AIAA Journal*, Vol. 45, No. 10, pp. 2544–2553, 2007.

[19] Menter, F. R., Langtry, R. B. and Völker, S., "Transition Modelling for General Purpose CFD Codes", *Flow, Turbulence and Combustion*, Springer, Vol. 77, pp. 277–303, 2006.

[20] Langtry, R. B., Menter, F. R., "Transition Modeling for General CFD Applications in Aeronautics", *AIAA*, Vol. 100, 2005.

[21] Menter, F. R., "Two–equation Eddy–viscosity turbulence models for engineering applications", *AIAA J.*, 32(8), pp. 1598–1605, 1994.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Thermal Cycle Analysis of Constant Volume Combustion Gas Turbine Engine

Roberto Andriani and Umberto Ghezzi
Politecnico di Milano, Department of Energy, Milan, Italy

Fausto Gamma and Antonio Agresta
Università "La Sapienza", Dep. of Mechanics and Aerospace Engineering, Rome, Italy

Keywords: *Constant Volume Combustion, Thermodynamic, Propulsion.*

Abstract

New interest rises in gas turbine systems where the combustion process takes place not at constant pressure, like in conventional gas turbine engines, but at constant volume. The main feature of such system is that the gas pressure increase is not entirely provided by the compressor, but also by the combustion process, that taking place in a closed volume increases the gas pressure. In this way the power adsorbed by the compressor is less than that adsorbed in a conventional gas turbine engine, and as consequence the available power at the turbine shaft is greater. By means of a thermodynamic numeric program that simulates the constant volume combustion turbine behavior, a parametric study is performed, computing at different operating conditions the main system performances and efficiencies.

1 Introduction

The aircraft cost reduction has led in the last decades to the development of new high-power and low fuel consumption propulsion systems. Together with these characteristics, another aspect, the pollutant emissions, has become of primary importance in the development of new aero engines. In this research, the constant volume combustion turbine plays an interesting role. The idea is not new, because at the beginning of the last century some prototypes were built for electrical energy production [1-2]. However, the constant pressure turbine became more powerful and efficient, and all the engineer efforts were addressed to that direction, neglecting the constant volume turbine. Today new interest rises on this subject because, together with the possibility to reach high thermal efficiency, the new technology and in particular the electronic engine control makes the engine use more flexible, and also suitable for propulsion systems [3-5]. Some applications consist in using it as the high pressure section in a conventional gas turbine engine, in particular the valveless configuration [6-10]. However, if high frequencies can be attained, also the valved configuration can be considered as aero engine [11-13]. The constant volume combustion gas turbine is a particular engine that joins together some features of gas turbine and piston engine. In fact, instead of the constant pressure combustion chamber used in gas turbine engine, the combustion process takes place in a closed volume (constant volume), similar to that of a piston engine.

In this way the working fluid pressure rise occurs in two times: first by means of the low pressure compressor, second as consequence of the closed volume combustion. So, once fixed the maximum cycle pressure, the power adsorbed for compression is less than that of a gas turbine engine. Consequently the available output power and the efficiency are greater.

The aim of this work is to study the thermodynamic behavior of the system, computing the thermal cycle and hence the main efficiencies and performance.

2 Thermodynamic analysis

Fig. 1 shows the working scheme of the constant volume combustion turbine.

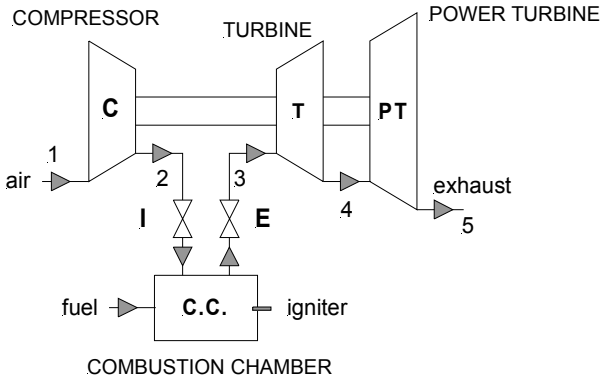


Fig. 1 Working scheme of constant volume combustion gas turbine

Air, at ambient conditions 1, enters the low pressure compressor C raising its temperature and pressure to condition defined at point 2.

Then it gets in the combustion chamber C.C. passing through intake valve I, while fuel is introduced by a nozzle. Once the chamber is completely filled, intake valve I is closed and, started by igniter, combustion begins. Since it takes place in a closed volume both temperature and pressure raise, reaching their maximum value. When combustion process is over, exhaust valve E opens, allowing gas to pass through turbine T, to drive compressor C, and turbine PT that provides the output power. When combustion chamber has been evacuated, valve E is closed, valve I opens and the cycle starts again.

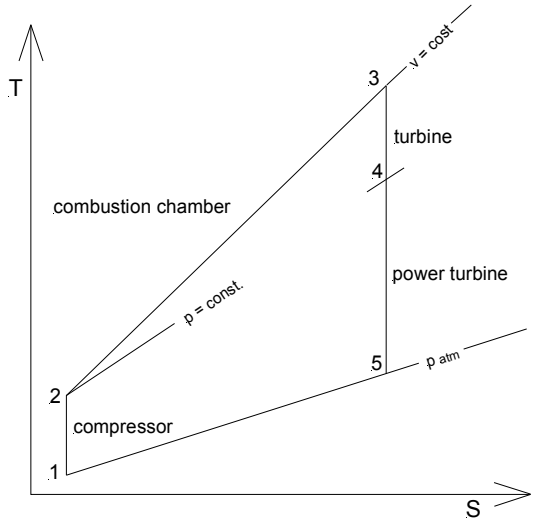


Fig. 2 Thermodynamic cycle in a Temperature/Entropy plane of constant volume combustion gas turbine

In Fig. 2 it is reported in a Temperature/Entropy plane the thermal cycle. For sake of simplicity the transformations are considered ideal, and the ambient pressure and constant volume lines are marked. The thermal efficiency of the system is the ratio of output power and heat introduced; referring to the scheme and the cycle of Figs. 1 and 2 we can say that the output power is represented by the power turbine work, because the power developed by high pressure turbine is fully absorbed by compressor.

The heat for unity of mass, Q_{in}, introduced at constant volume in the combustion chamber, is equal to the gas internal energy variation between the points 2 and 3 in the graphic:

$$Q_{in} = U_3 - U_2 = c_v(T_3 - T_2) \qquad (1)$$

The heat for unity of mass released is the internal energy variation between the points 5 and 1:

$$Q_{out} = H_5 - H_1 = c_p(T_5 - T_1) \qquad (2)$$

Since the flow is intermittent, the turbine does not operate in steady conditions, as in a conventional gas turbine engine. In fact, the gas pressure and temperature at the turbine entry vary with time, following the combustion chamber conditions. When the exhaust valve opens, gas pressure and temperature are at their maximum values, corresponding with the cycle maxima. Then temperature and pressure at the turbines entry reduce their value until reaching their minimum, represented, in the ideal case, by the external pressure. In this way, at any cycle, the turbines are fed by gas at different pressure and temperature.

The *HPT* and the power turbine output work must therefore be computed in a different way respect to the constant pressure turbine.

The temperature drop T_3-T_5 is divided into N parts, each part equal to:

$$\Delta T = (T_3 - T_5)/N \qquad (3)$$

After a certain time, the temperature inside the combustor has reduced to the value $T - \Delta T$. If we consider the expansion inside the combustion chamber, due to the mass reduction, as adiabatic and, as first approximation isentropic, then the pressure inside the chamber is:

$$p_{T-\Delta T} = p_T \left(\frac{T - \Delta T}{T}\right)^{\frac{\gamma}{\gamma-1}} \qquad (4)$$

where γ is the specific heat ratio respectively at constant pressure and constant volume, and the point 3 in the graphic of fig.1 moves downward at constant entropy. Although the initial temperature and pressure are different for different gas parts, since the expansion in both turbines is adiabatic-isentropic, the final temperature and pressure of the whole flow is constant and equal to the condition defined by point 5. The mass inside the combustion chamber at the temperature T is:

$$m_T = \varrho_T V \qquad (5)$$

where ϱ_T is the gas density inside the combustion chamber. Considering the perfect gas state equation:

$$\varrho_T = \frac{p_T}{RT} \qquad (6)$$

During the temperature variation ΔT the mass through the turbines is equal to $m_T - m_{T-\Delta T}$, and the works provided by this mass through the *HPT* and the power turbine is:

$$L_T = (m_T - m_{T-\Delta T}) c_p (\bar{T} - T_5) \qquad (7)$$

where $\bar{T}$ is a mean value between T and T-ΔT.

The whole turbine work L is the sum of the N works L_T, where N is the number of the ΔT intervals.

$$L = \sum_{i=1}^{N} L_{T_i} \qquad (8)$$

As previously reminded, the work L is the sum of the *HPT* and of the power turbine work. Since the *HPT* work balances the compressor work, the power turbine work L_{pt} is the difference between L and the compressor work L_c:

$$L_{pt} = L - L_c \qquad (9)$$

the compressor work is given by:

$$L_c = m c_p T_1 \left(\beta_c^{\frac{\gamma-1}{\gamma}} - 1\right) \qquad (10)$$

where m is the mass passing through the compressor and the turbines (except the fuel mass), and β_c the compressor pressure ratio.

The heat for unity of mass introduced during combustion is given by the internal energy difference (Equation 1). If we consider also the mass we can write (neglecting the fuel mass):

$$Q = mc_v(T_3 - T_2) \qquad (11)$$

The thermal efficiency is then given by:

$$\eta_{th} = \frac{L_{pt}}{Q} \qquad (12)$$

3 Results

A thermodynamic numerical code has been developed to simulate the behavior of a constant volume combustion gas turbine. The code solves the ideal cycle, computing the air/gas thermodynamic characteristics (pressure, temperature, density, etc.) at any point. The operating conditions, as compressor pressure ratio and maximum cycle temperature, can be varied, obtaining a parametric analysis of the system. In Fig. 3 is reported the behavior of the specific work (work for unit of mass) as function of the compressor pressure ratio and the maximum cycle temperature. The latter coincides with the high pressure turbine entry temperature *TET*. For this first analysis it has been considered the ideal cycle, meaning ideal behavior components. The gas expansion in turbine has been divided into 5 parts, following the procedure above described. The air conditions at the compressor entry are those at sea level (288K and 1 bar), and no flight speed is considered.

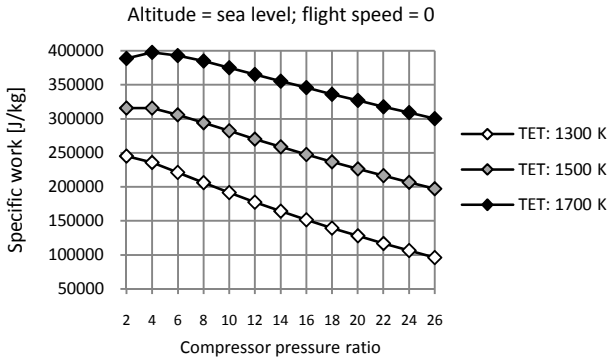


Fig. 3 Power turbine specific work (output work for unit of mass) as function of compressor pressure ratio and for three *TET* values

The graphic shows a similar behavior for the three *TET* considered. For higher temperatures the curves are shifted towards higher work values. The curves, except for the case at 1300K, whose maximum is not present inside the graphic, show a maximum at low pressure ratios, and then decrease as pressure rises. The position of the curve maximum values is different for the three cases; as the *TET* rises, the maximum moves towards higher pressure ratios. The work difference with temperature is great; if we consider the maximum value, it is about 250 kJ/kg at 1300K, about 320 kJ/kg at 1500K, and about 400 kJ/kg at 1700K of *TET*.

In Fig. 4 is shown the amount of introduced heat during the constant volume combustion process, for unity of mass, as function of the compressor pressure ratio and for the three *TET* analyzed.

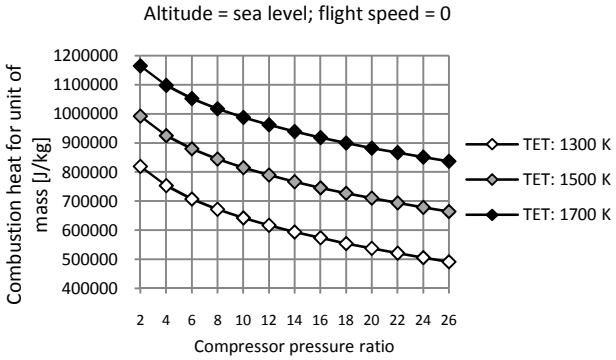


Fig. 4 Heat amount introduced during the constant volume combustion process as function of compressor pressure ratio and for three *TET* values

The introduced heat is continuously reduced as the compression ratio rises, and the behavior is common to the three *TET*. Similarly to the output work, also in this case the heat amount is greatly influenced by *TET*, greatly growing with it.

The thermal efficiency behavior, shown in Fig. 5, comes directly from the previous two graphics. It is strongly influenced by maximum cycle temperature, especially at higher pressure ratios, rising with it. The curves present a maximum, whose position depends on *TET*: at low temperature the maximum is at low

pressure, while as temperature rises, the maximum is present at higher pressure ratios. However also at the higher temperatures, the maximum is practically reached at not very high values: about 5 at 1300K, 7-8 at 1500K and about 10 at 1700K. At low temperature the maximum is more marked, while at higher *TET* it becomes smoother. If we compare the graphics of the output work with those of the thermal efficiency, we notice that, at the same *TET*, the maximum power turbine work is at pressure ratios lower than those at which the thermal efficiency has its maximum. For this reason, also the constant volume combustion turbine, similarly to the conventional constant pressure combustion turbine, has two typical working conditions: the maximum power condition, where the output work is maximum, at lower pressure ratio, and the minimum fuel consumption condition, at higher pressure ratio, where the thermal efficiency reaches the maximum.

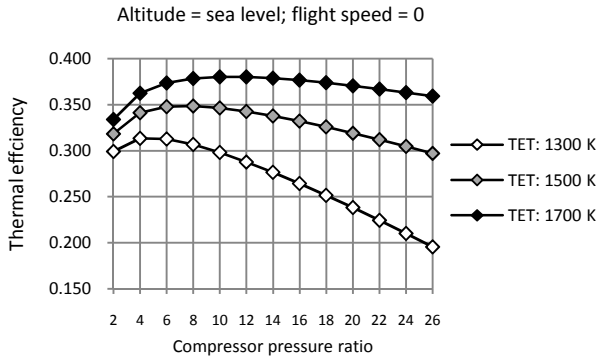


Fig. 5 Thermal efficiency as function of compressor pressure ratio and for three *TET* values

In the graphics above described, in x axis is reported the compressor pressure ratio, to specify that it is the pressure ratio provided by the compressor. As described in the introduction, this is not the actual pressure ratio, because the pressure is increased in the combustion chamber as consequence of the combustion, that is the main feature of the system. For this reason the pressure inside the combustion chamber after the combustion, just before the exhaust valve opening, is much more than that provided by the compressor. It depends on the compressor pressure ratio and on the *TET*, obviously rising with both of them. In Fig. 6 are reported the graphics that show this dependence, showing as function of the compressor pressure ratio and *TET* the maximum pressure inside the combustion chamber. In the graphics, the external pressure before the compressor is equal to 1 bar. It is interesting to note the combustion amplification effect on the gas final pressure. This explains because the output work maximum is at very low compressor pressure ratio. The actual gas pressure is much higher. For instance, at *TET* equal to 1300K the output work maximum is about 2 of pressure ratio, while the actual pressure ratio in the combustion chamber is about 7-8. At 1700K the maximum is at about 5 of pressure ratio, while the actual pressure ratio is about 17-18.

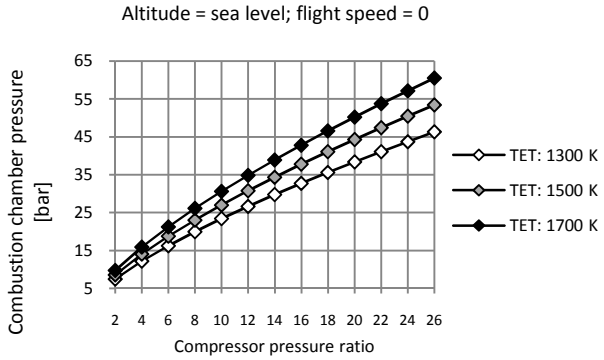


Fig. 6 Pressure values inside the combustion chamber as function of compressor pressure ratio and for three *TET* values

4 Conclusion

A parametric study of a constant volume combustion turbine has been performed by means of a thermodynamic numeric program able to compute the ideal output work and the ideal thermal cycle of the system. The simulated combustion chamber is with intake and exhaust valves, so that combustion takes place in a closed volume. The program allows to simulate the engine behavior at different operating conditions, as compressor pressure ratio and maximum *HPT* entry temperature. The codes solves the unsteady turbine expansion by

dividing it into a certain number of parts. Once computed the working cycle the codes allows to obtain the main performance and efficiency of system, as specific output work, introduced heat, thermal efficiency.

The main results can be summarized as follows:

- The turbine expansion division in finite intervals leads to good approximation in computing the turbine work and the thermal efficiency.
- Although the unsteady flow reduces the turbine work, the power turbine output work remains high thanks to the low adsorbed work by the low pressure compressor.
- The output work and the thermal efficiency are high also at low pressure ratio. This means that a low pressure compressor is sufficient to have good performance and efficiency.

References

[1] Stodola, A, "Steam and Gas Turbines, Volume 2", McGraw-Hill Book Company, Inc.
[2] E. Garuffa, "Motori a Combustione Interna", Ulrico Hoepli Editore, Milano, 1925.
[3] Kailasnath, K. "Review of Propulsion Applications of Detonation Waves," AIAA J. Vol. 38, No. 9, pp. 1698-1708, 2000.
[4] Mawid, M., Park, T., Sekar, B., Arana, C., "Turbofan Engine Thrust Augmentation With Pulse Detonation Afterburners - Nozzle Influence," AIAA 2002-4073, 2002.
[5] M.A. Mawid, T.W. Park, B. Sekar, and C. Arana, Air Force Research Laboratory, "Application of Pulse Detonation Combustion to Turbofan Engines", 2001- GT-0448, 2001.
[6] V. Tangirala, A. Dean, D. Chapin, "Performance Estimates of a Pulse Detonation Engine", ISABE 2003-1215, July 2003.
[7] Schauer, F., Bradley, R., Hoke, J., "Interaction of a pulsed detonation engine with a turbine," AIAA 2003-0891, 2003.
[8] A. Rasheed, V.E. Tangirala, C.L. Vandervort, and A.J. Dean, "Interactions of a Pulsed Detonation Engine with a 2D Turbine Blade Cascade", 42nd Aerospace Sciences Meeting and Exhibit, 5-9 January 2004, Reno, Nevada, AIAA 2004-1207.
[9] A. Rasheed, A. H. Furman, A. J. Dean, "Experimental Investigations of the Performance of a Multitube Pulse Detonation Turbine System", Journal of Propulsion and Power, Vol. 27, No. 3, May–June 2011.
[10] Caldwell, N., Glaser, A., and Gutmark, E., "Performance Measurements of a Pulse Detonation Engine Array with a Turbine," 42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Sacramento, CA, AIAA Paper 2006-4307, 9–12 July 2006.
[11] Snyder, P.H., Alparslan, B., Nalim, M.R., "Gas Dynamic Analysis of the CVC, A Novel Detonation Cycle," AIAA-2002-4096, 2002.
[12] R. Andriani, U. Ghezzi, E. Infante, "Thermal Analysis of Constant Volume Combustion Gas Turbine", 43rd AIAA Aerospace Science Meeting and Exhibit, AIAA 2005-1010, January 10-13, 2005, Reno, Nevada.
[13] Glaser, A., Rasheed, A., Dunton, R., and Tangirala, V. E., "Investigations of Thrust Generated by a Valved, Multi-Tube PDE with Exit Nozzles," 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Hartford, CT, AIAA Paper 2008-4692, 21–23 July 2008.

CEAS 2011

Human Factors and Safety

Further Investigation into Landing Flare Guidance Using the Optical Parameter Tau

Garnet R. Ridgway and Michael Jump
The University of Liverpool, UK

Keywords: *Landing, Flare Manoeuvre, Visual Perception, Displays, Tau*

Abstract

The landing flare is widely regarded as being one of the most difficult aspects of piloting a fixed wing aircraft, and this phase of flight accounts for a high proportion of fatal accidents for civil transport aircraft. Previous studies have sought to identify the piloting strategy used for the flare manoeuvre, and thence develop pilot aids to improve performance and safety. A previous study at the University of Liverpool designed and evaluated a flare guidance algorithm, based on the Tau theory of visual perception, with some success. This study sought to address the key limitations of the previous work; firstly by refining the display symbology used for the piloted evaluation; secondly by implementing a variation on the guidance strategy that was shown to be more successful. The new algorithm was evaluated in a piloted simulation experiment, and was shown to provide performance comparable to an in-service example display.

1 Introduction

For the first few years of the 21st century, the number of fatal accidents involving passenger airliners had been in steady decline [1]. However, a significant number still occur annually - in 2010 there were 8 fatal accidents, including 3 cases with over 100 fatalities [1]. The statistics for the period 1999 – 2010 are shown in Fig. 1. These show that during the last 10 years the greatest proportion of fatal accidents have occurred in the approach and landing phase of flight. In the last four years, the number of occurrences of fatal accidents in this phase of flight has remained approximately constant. This suggests a slowing of the downward trend of fatal accidents in the approach and landing flight phase which has been in evidence since the 1970s [1]. Reasons for the high proportion of accidents in this phase of flight include proximity to the ground and high perceived task difficulty [2]. Indeed, Ref. [2] showed that the majority of pilots surveyed indicated that the landing flare (typically an aft longitudinal stick input resulting in a pitch up to achieve the required reduction in vertical velocity in the last few seconds before

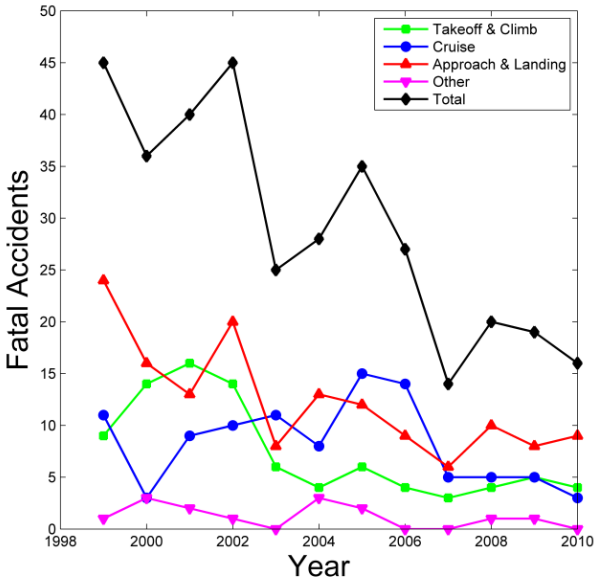


Fig. 1 Fatal airliner accidents worldwide 1999-2010

touchdown) is the most difficult manoeuvre performed during a typical flight, and that the most significant factor for improving performance is practice. These factors suggest that current methods may not be fully effective in training pilots to consistently perform the flare manoeuvre in varying conditions, and that the learning processes involved are not well understood. In fact this is alluded to in flight training literature. For example the training manual, Ref. [3], states *"Learning where to start the flare is a matter of trial and error"*. In addition to being a difficult manoeuvre, correct implementation of the flare is critical to the safety of the aircraft. Flaring too early can lead to the aircraft "ballooning", with the subsequent risk of a runway overshoot. Flaring too late can result in a hard landing, i.e. making contact with the runway surface at a higher than recommended vertical velocity, potentially causing damage to the aircraft structure and injury to passengers.

Although Automatic Landing Systems (ALS) are capable of performing fully automatic landings (depending on the aircraft / airfield configuration) [4], regulations mandate that a proportion of landings are carried out manually in order to maintain pilot proficiency [4]. For this reason, systems that provide flare cues to the pilot but which do not automatically control the aircraft are of particular relevance to this manoeuvre. An example of such a system is the Visual Guidance System (VGS) Head Up Display (HUD) developed by BAE Systems [5], which provides the pilot with guidance cues prior to and during the flare manoeuvre. However, despite such systems being in use for at least 10 years [6], the statistics presented in Fig. 1 illustrate that there is more work to be done. A question therefore arises as to how this problem might be approached from a different perspective. This paper presents the latest results from research undertaken at the University of Liverpool with respect to display algorithms based upon the psycho-physics based Tau Theory of visual perception.

Previous research at the University of Liverpool has identified piloting strategies based on the Tau theory of visual perception. Ref. [7] demonstrated that pilots flying nap-of-earth manoeuvres in helicopters make use of the parameter *time-to-contact*, or *tau* (τ), to define their control inputs. A subsequent study demonstrated that pilots of both simulated and actual helicopters make use of τ-related parameters to perform runway traverse manoeuvres [8]. A key conclusion of Ref. [7] was that such strategies were eminently suitable for application to other phases of flight, and that Tau might form a suitable basis for the development of pilot aids. The work cited here related primarily to helicopter flight. This is to be expected given the nature of many rotary-wing tasks (i.e. close to the ground) and hence its relevance to guidance theories that require some form of visual sensory input. However, the development of a pilot guidance aid was actually first undertaken for a fixed-wing application in Ref. [4], that of the flare manoeuvre. Here, coherent tau-based descent strategies were observed during this phase of flight in both simulated and actual flight test data. One of these strategies was utilized to develop a flare guidance algorithm which was implemented in a simple head up display format. This concept was evaluated against an in-service example [5] with some success. However, a number of limitations of this novel design were identified in Ref. [4]. The guidance strategy selected for evaluation in Ref. [4] was shown to provide the most consistent performance of the two. However, the alternative strategy was shown to be the more successful, albeit with less consistent performance. Secondly, the primitive symbology used for the evaluation exercise caused some confusion amongst the pilots whilst attempting to follow the guidance cue. The objectives of this study were therefore twofold. Firstly to assess the second coherent tau guidance strategy observed in Ref. [4], and secondly to address the identified limitations. To this end, a new flare guidance algorithm has been designed, implemented and evaluated in a piloted flight simulation experiment.

The remainder of the report is structured as follows. Section 2 outlines the fundamental concepts of Tau theory, Section 3 details the

objectives of the study. The design of the new guidance algorithm is outlined in Section 4, followed by details of the evaluation exercise in Sections 5 and 6.

2 Tau Theory

2.1 Tau Fundamentals

The concept of performing a manoeuvre in order to decelerate, of which the flare manoeuvre is an example, is common in the natural world. For example birds and insects perform manoeuvres analogous to the flare when landing [10]. Gibson, the pioneer of the Ecological approach to visual perception, hypothesized that animals control their movement based on the relative motion of the elements in their visual scene, known as the *optic flow* [11]. This approach to visual perception was further developed by Lee, who introduced the optical parameter *tau* (τ), the time to contact with an object at a distance of x at the current closure rate $\dot{x}$, as shown in Eq. (1).

$$\tau_x = \frac{x}{\dot{x}} \qquad (1)$$

A number of subsequent studies demonstrated the use of τ-based guidance strategies in the animal kingdom. For example, Ref. [10] showed that pigeons landing on perches regulate their deceleration by keeping the rate of change of time to contact, $\dot{\tau}_x$, constant. Such a strategy has also been observed in hummingbirds landing on feeders [12] and car drivers braking to a halt [13].

2.2 Aerospace Applications

Prior research at The University of Liverpool has sought to identify the control strategies used by pilots to both initiate [14] and execute [9] the flare manoeuvre. Ref. [9] proposed that pilots make use of the parameter "time-to-contact with runway", τ_h, to control the trajectory of the flare (Fig. 2).

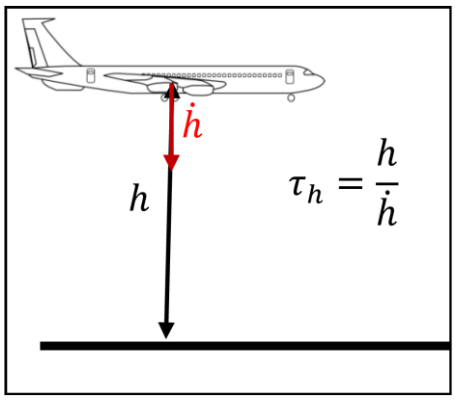


Fig. 2. Gap definition for the flare manoeuvre

The selection of a constant value of rate of change of time to contact with runway, $\dot{\tau}_h$, was shown to determine the characteristics of the flare. Two distinct strategies were found to exist when closing the motion gap in Fig. 2: (1) a so-called '2-stage' landing.

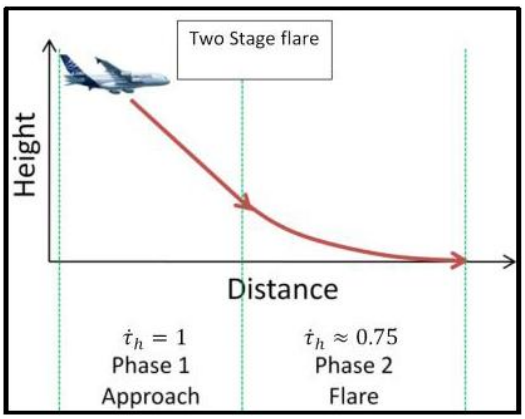


Fig. 3 Schematic of 2-stage flare strategy

Here a constant vertical approach velocity ($\dot{\tau}_h$=1) was reduced to the touchdown velocity by selecting some value of $\dot{\tau}_h$<1 which results in a continuous but not necessarily constant deceleration to touchdown (Fig. 3); (2) a '3-stage' landing. Here the constant vertical velocity of the approach is reduced approximately exponentially ($\dot{\tau}_h \approx 0$). In theory, of course, this would result in the aircraft never reaching the runway surface. As such, a third phase in which $\dot{\tau}_h$ is some value greater than zero is observed (Fig. 4).

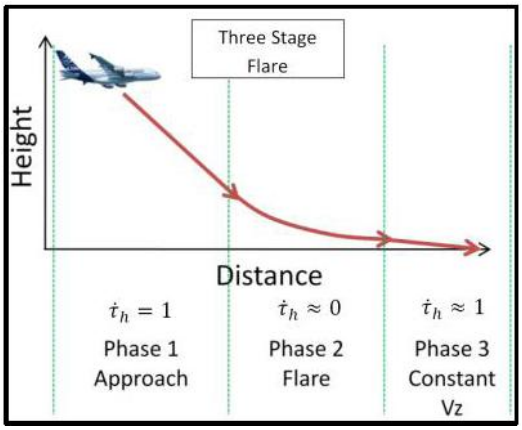


Fig. 4 Schematic of 3-stage flare strategy

Analysis of flight test and simulated flight flare manoeuvres indicated that, in terms of touchdown velocity, the former technique was predictable whilst the latter gave consistently lower but less predictable touchdown velocities.

2.3 Previous Tau-Based HUD Algorithm

The predictability of the 2-stage landing technique led to the development of a novel flare guidance algorithm based on a constant $\dot{\tau}_h$ strategy [4]. A flight path angle cue was generated to command a $\dot{\tau}_h$=1 (approach at constant vertical velocity) and $\dot{\tau}_h$=0.75 (flare), as defined in Eq. (2):

$$\gamma_{CMD} = \theta - \frac{1}{cos\theta . cos\varphi}\left[\frac{1}{u}\sqrt{\frac{\Delta h \ddot{h}}{1-\dot{\tau}_h}} + sin\theta\right] \qquad (2)$$

where γ is the vertical flight path angle, θ is the pitch angle, φ is the roll angle and u is the forward body velocity. This flare algorithm was used to drive a basic symbology set on a Head Up Display (HUD) for the purposes of piloted evaluation (Fig. 5).

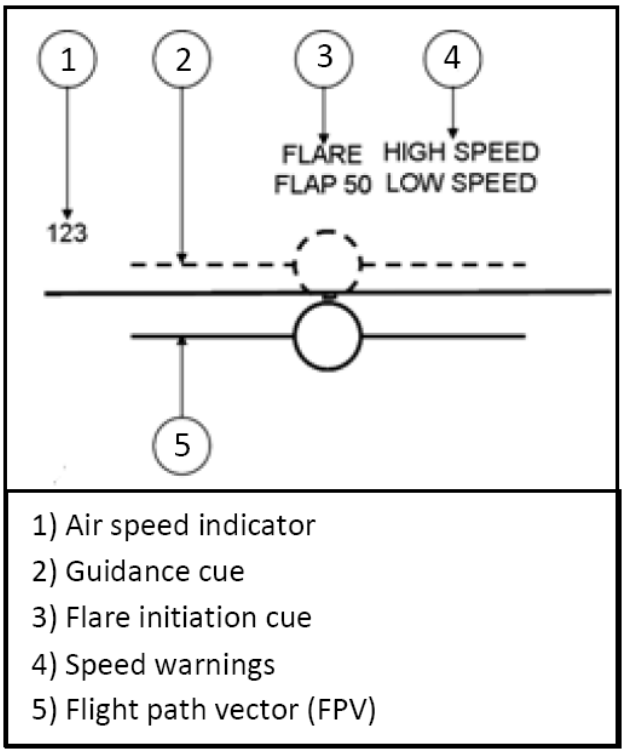


Fig. 5 HUD symbology for piloted evaluation in Ref. [4]

The results of Ref. [4] demonstrated that, in good visual conditions, the novel tau-based algorithm produced results comparable to an in-service example HUD. There was also an improvement in performance in a degraded visual environment compared with a baseline case using only a Head Down Display (HDD); however this was not as consistent as for the in-service example. Two key limitations of this novel tau-based flare algorithm were identified by [4]:

- When flown manually, the 3-stage approach to flaring resulted in consistently lower, but less predictable, touchdown velocities than for flares conducted using the two-stage approach.
- The display symbology was not initially intuitive, resulting in some confusion as to which symbol the pilot should be following.

3 Objectives

The objectives of this study were twofold. Firstly, to implement a tau-based flare guidance algorithm to command the 3-stage strategy as identified in [4]. Such a strategy has been shown to be more successful than the 2-stage strategy that formed the basis of the design of the previous tau-based flare algorithm [4]. When flown manually, the 3-stage approach to flaring resulted in consistently lower, but less predictable, touchdown velocities than for flares conducted using the two-stage approach.

Secondly, the display symbology used for piloted evaluation of the flare algorithm was to emulate the ergonomics of an in-service example as closely as possible, in order to allow a more objective comparison between the guidance of the novel and in-service algorithms.

4 Algorithm Design

A novel flare guidance algorithm was implemented based on the 3-stage flare strategy outlined in Section 2.1. In order to achieve this, the value of the $\dot{\tau}_h$ command had to be varied as the manoeuvre progressed. The behavior of the $\dot{\tau}_h$ parameter is shown in Fig. 6.

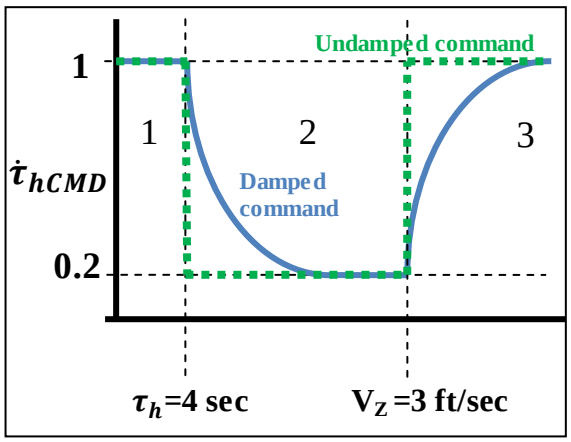


Fig. 6 Behavior of $\dot{\tau}_h$ during the flare manoeuvre

During the approach stage (1), the value of $\dot{\tau}_h$ was set to 1, and the HUD commanded a constant flight path angle of 3°. The parameter τ_h was used to trigger an anticipation cue ("STAND BY") at τ_h=6 sec and the flare initiation cue ("FLARE") at τ_h=4 sec. The use of τ_h as a cue for flare initiation is consistent with both the design of the previous tau-based flare algorithm [4] and the findings of [14], which showed the value of τ_h=4 sec to be appropriate for the evaluation aircraft (Section 5.1).

Upon passing the threshold value of τ_h (2), the flight path angle command switched from the constant approach value to being a function of $\dot{\tau}_h$ as shown in Eq. (3):

$$\gamma_{CMD} = \gamma + k(\dot{\tau}_h - \dot{\tau}_{h\,CMD}) \qquad (3)$$

The piloting task was, therefore, to minimize the error between $\dot{\tau}_h$ and $\dot{\tau}_{h\,CMD}$ by controlling the pitch angle (and hence vertical flight path angle) of the aircraft. Shortly after flare initiation, the reduction in the value of $\dot{\tau}_{h\,CMD}$ resulted in a positive value of γ_{CMD}, corresponding to a nose up command which was sustained until the value of $\dot{\tau}_h$ was sufficiently reduced. The amplitude of the nose up command was determined by the value of the proportional gain k*. This deceleration was sustained until the vertical velocity (V_Z) of the aircraft passed below the target touchdown value of 3 ft/sec, a value which was defined by Ref. [15] as desired performance for this task. At this point (3), the $\dot{\tau}_{h\,CMD}$ value returned to 1, corresponding to a constant vertical velocity command which was sustained until touchdown.

During initial testing of the algorithm, the transition between values of $\dot{\tau}_{h\,CMD}$ generated a rapid shift in the position of the guidance cue, significantly adding to the task difficulty. For this reason, the transition between $\dot{\tau}_{h\,CMD}$ values was damped by the short period natural frequency of the aircraft in order to improve the controllability of the guidance cue as shown in Fig. 6 (solid blue line).

* For the purposes of this study, the value of k was set such that the difference in γ_{CMD} produced immediately after flare initiation was equal to the approach angle i.e. 3°.

5 Experimental Setup

5.1 Equipment

Piloted evaluation was carried out using the HELIFLIGHT Simulator in the Flight Simulation Laboratory, now part of the Centre for Engineering Dynamics at the University of Liverpool (Fig. 7). HELIFLIGHT features a 6 degree of freedom motion base, 6 visual channels and an integrated HUD capability [16]. The aircraft model used in this experiment was a variant of the *Generic Large Transport Aircraft* (GLTA) developed for Ref. [4]. The simulation software used for the aircraft dynamics was *FLIGHTLAB* [17], and the visual scene was generated by *Landscape* [18]. Display symbology was defined using the *VAPS* HMI software package [19] and then integrated into *Landscape* for real time simulation.



Fig. 7 HELIFLIGHT Simulator at the University of Liverpool

5.2 Pilots

Three pilots were selected for the evaluation exercise:

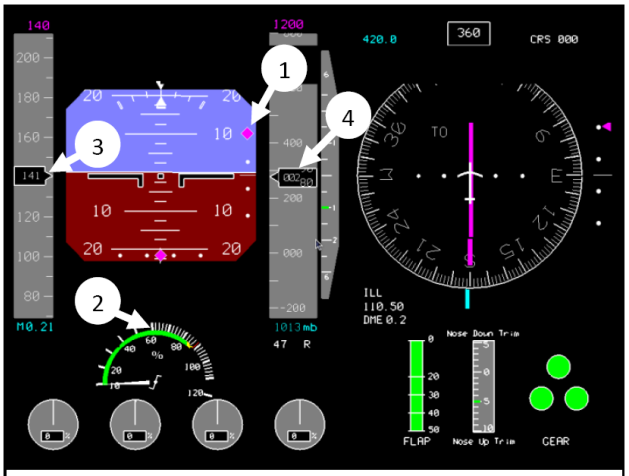
- P1: Professional airline pilot and former military test pilot, currently flying short haul Airbus aircraft for a major carrier.
- P2: Current civil and former military test pilot, with operational experience of military large fixed-wing transport aircraft.
- P3: Retired professional airline pilot with operational experience on a number of large transport aircraft.

Both P1 and P2 had extensive previous experience of the HELIFLIGHT simulator. All pilots completed a number of familiarization flights before commencing the display evaluation task.

5.3 Displays

In addition to the novel HUD algorithm, a standard head down display, the Primary Flight Display (PFD) and an in-service example of a HUD, the Visual Guidance System (VGS) were evaluated for the purposes of comparison. The key features of these displays are outlined in this section.

PFD: This display was used as a baseline case against which to compare the relative performance of the novel HUD, as in Ref. [4]. The functions of particular relevance to the approach and landing task are shown in Fig. 8. It contains symbology that is representative of a typical flight deck but is not a facsimile of any particular display in current service.

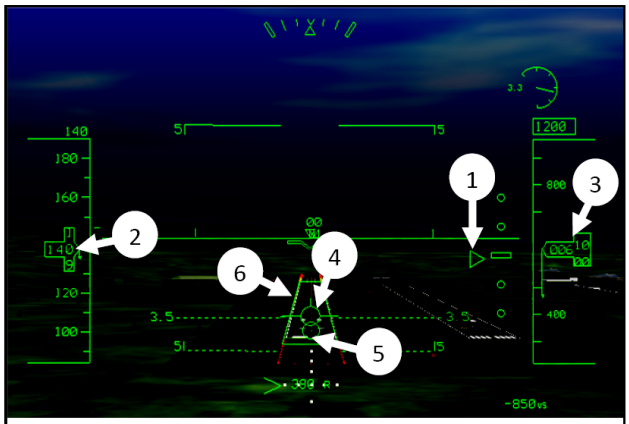


1) ILS glide slope indicator
2) Throttle position indicator
3) Air speed indicator
4) Pressure altimeter

Fig. 8 Primary Flight Display

VGS: The BAE VGS was used as a baseline in-service example HUD. The VGS provides

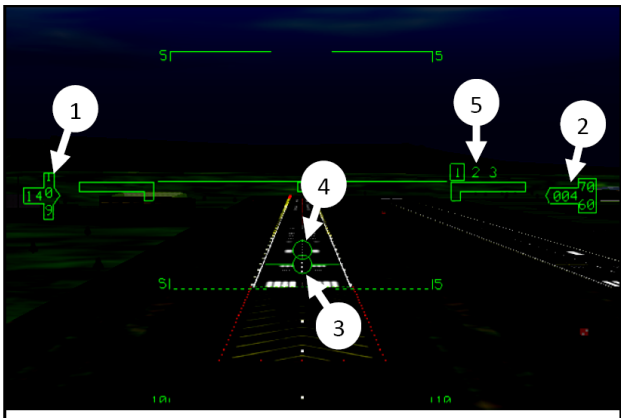
flight information for all phases of normal operational flight [5], and for this reason not all of the symbology was relevant to the approach and landing task (Fig. 9)†. During the approach, the ILS glide slope indicator on the HUD displays the same information as that on the PFD. A conformal outline of the runway is also displayed during the approach phase, but this is automatically disabled once the aircraft passed below 100ft. As the aircraft descended towards the flare initiation point, a pre-flare cue ("+" symbol) appears from the bottom of the screen, moving towards the flare guidance cue. Once the pre-flare cue overlays the flare guidance symbol, the pilot is required to follow the cue with the FPV symbol, hence applying and aft stick input and performing the flare manoeuvre. Once main gear touchdown is achieved, the flare guidance symbology is disabled.



1) ILS glide slope indicator
2) Air speed indicator
3) Pressure altimeter
4) Flight path vector (FPV)
5) Flare guidance cue
6) Runway outline

Fig. 9 Visual Guidance System (VGS) HUD

TDE (Tau Dot Error): Since the objective of the study was to evaluate a novel HUD algorithm rather than a complete system, relevant aspects of the VGS symbology were emulated for use with the TDE (Fig. 10). During the approach stage, the guidance cue (4) is used to indicate the appropriate approach angle based on information from the ILS glide slope indicator. The "STANDBY" and "FLARE" messages appear in the centre of the screen as detailed in Section 4. During the flare, the task is the same as for the VGS; i.e. to overlay the flare guidance cue with the FPV. This is achieved by the application of an aft stick input. Following touchdown the guidance cue is disabled.



1) Air speed indicator
2) Pressure altimeter
3) Flight path vector (FPV)
4) Flare guidance cue
5) Flare phase indicator

Fig. 10 Tau Dot Error (TDE) HUD

5.4 Task Definition

A simplified approach and landing task was used for the purpose of evaluating the novel HUD algorithm. The task started with the aircraft trimmed in a standard approach at -3° vertical flight path angle at 145kts IAS. The pilots were required to maintain this approach condition through use of the airspeed indicator and ILS glide slope indicator, both of which were located on the PFD HDD (Fig. 8). The flare manoeuvre was initiated either at the discretion of the pilot for the PFD-only flights (control case), or as directed by the HUD for the VGS and TDE flights. In order to focus on the

† Note that in order to clearly demonstrate the HUD symbology, the background scene in Figures 8 and 9 has been darkened. For the piloted evaluation, a daytime scene was used.

longitudinal axis (the primary axis of motion for the flare), the lateral states of the aircraft model were disabled for this study. Therefore any lateral stick or pedal movements made by the pilot had no effect on the motion of the aircraft. Two visual conditions were used[‡]:

- GVE: 60000ft visibility, day
- DVE: 150ft visibility (fog), day, equivalent to Cat IIIb RVR

After each flight, the pilots rated both the task workload through use of the Bedford Workload Scale [20], and the performance of the display with the Display Controllability Scale [21]. In addition to the pilot ratings, the vertical velocity at touchdown, V_{ZTD}, was recorded to provide a measure of task performance. Adequate and desired performance criteria were defined as 3ft/sec and 5ft/sec respectively as specified in Ref. [15] Each pilot / display / visual condition combination was repeated 5 times.

6 Results

6.1 GVE

Fig. 11 shows the vertical velocity at touchdown for each of the flights in the GVE condition (for key to notation see Fig. 12). The red and green lines represent the adequate and desired performance criteria respectively, as defined in [15]. As stated in Section 1, the flare manoeuvre is often considered to be one of the most difficult aspects of piloting a fixed wing

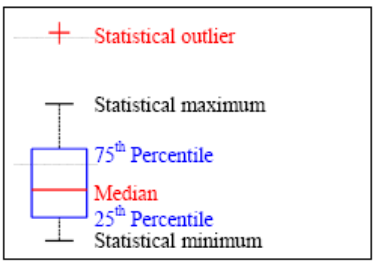


Fig. 12 Key to box plot notation

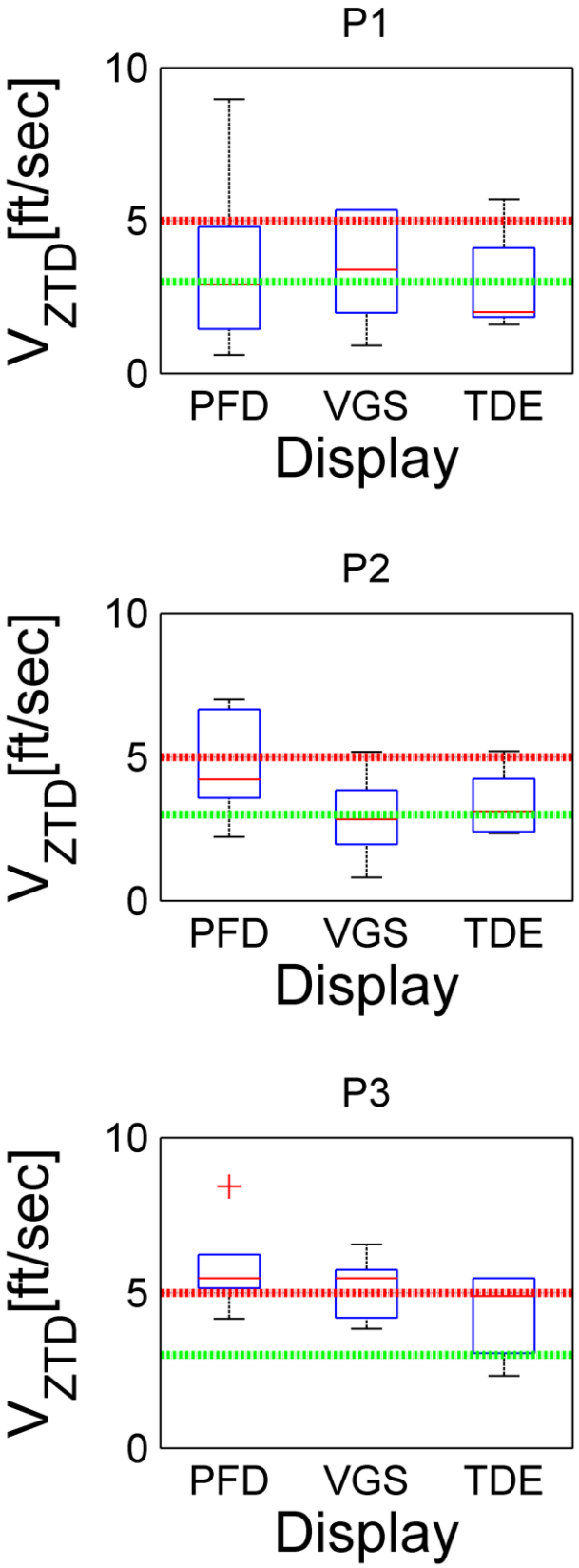


Fig. 11 V_{ZTD} performance for 3 displays and 3 pilots in GVE conditions

[‡] Due to time constraints, only P1 and P2 completed flights in the DVE condition.

aircraft [2]. This supposition is supported by the results of the PFD flights, in which each pilot performed at least one "hard landing" (i.e. with a V_{ZTD}> 5ft/sec) in spite of high levels of experience on large transport aircraft and practice in the simulated environment. This was especially true for P3, who performed the greatest number of landings outside of the adequate performance criterion. Investigation into the performance trend with the number of flights suggested that this was not due to the pilot requiring more practice, as the V_{ZTD}values converged over time. This was supported by the pilot comments, which suggested that the pilot believed he was achieving adequate performance. A potential explanation for this could be a difference in appropriate piloting strategy between P3's most recent large transport aircraft type and the aircraft model used in this study.

The VGS HUD gave a significant improvement in performance for P2, and a moderate improvement for P3. For P1, the VGS HUD improved the consistency of the V_{ZTD} values, but moved the median value from desired to adequate performance.

The TDE HUD gave an improvement in V_{ZTD} performance for all three pilots. This improvement was particularly marked for P3, with the median V_{ZTD}value moving inside the adequate performance criterion.

Pilot workload and display controllability ratings are shown in Fig. 13. The PFD was consistently awarded low ratings (for both scales a low rating indicates desirable characteristics). Pilot comments suggested that this was due to the familiarity with and simplicity of the display symbology, the only objectionable characteristic being a slight oversensitivity of the ILS glide slope indicator. For this study, this instrument was driven by a simple function of the aircraft position and the desired touchdown position, which may have been less well damped than the in-service implementation.

P1 awarded significantly higher controllability values to the HUDs than to the baseline HDD. One reason given for this was the difficulty in distinguishing between the

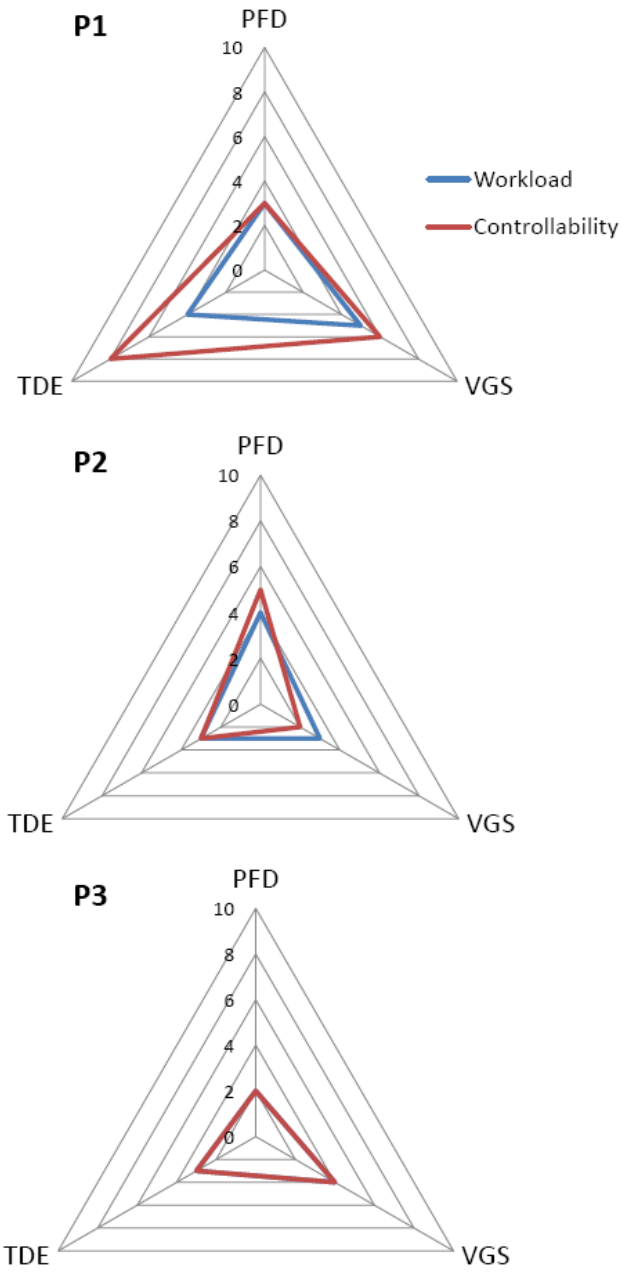


Fig. 13 – Workload and display controllability for 3 displays and 3 pilots in GVE conditions

guidance cue and FPV symbols. At the time of the test the display hardware appeared to be functioning correctly, and P1 fulfils the eyesight requirements for the JAR class 1 pilot medical examination [22]. P2 and P3 were subsequently asked to comment on the clarity of the display, with positive responses. As such, P1's perceived difficulty with the clarity of the display is unexplained. The TDE was awarded a higher controllability rating than the VGS by P1 due to the higher sensitivity of the flare

guidance cue. Eq. (2) shows that the command is a function of the current flight path angle, which resulted in a slight inconsistency in the behavior of the guidance cue. However it should be noted that all pilots reported that this became easier to predict once they were more familiar with the display.

Both P2 and P3 awarded similar workload and controllability ratings for the VGS and TDE (note that P3's workload and controllability scores directly correlate, hence only one line appears in Fig. 13), suggesting that the TDE algorithm provides guidance cue behavior comparable with the in-service example. The awarded controllability ratings of 2 and 3 correspond to display dynamics which are "*Satisfactory without improvement"* [21]. For a development-standard display to provide comparable performance to an in-service example is a significant result.

6.2 DVE

Results for the flights in the DVE condition are shown in Fig. 14. Both pilots performed a number of hard landings for the PFD cases, a result which is unsurprising given the lack of visual cues. Indeed, both pilots commented that the manoeuvre would not be attempted in such conditions during normal operations. P1 was able to perform a single landing within the adequate performance criterion by initiating an "open-loop flare" at 50ft above ground level (AGL); i.e. estimating the required aft stick input based on his experience of the GVE flights rather than on the available visual cues.

The VGS HUD gave a significant improvement in performance for both pilots in this degraded visual condition. A Wilcoxon signed-rank test, performed on the PFD and VGS data for both pilots yielded a p value of 0.002, suggesting that this is a statistically significant result. This was mainly attributed to the runway outline symbology (Fig. 9) restoring appropriate visual cues during the approach, and the pre-flare cue giving sufficient anticipation of the flare (the findings of Ref. [14] previously highlighted the importance of initiation point to the success of the subsequent manoeuvre). One

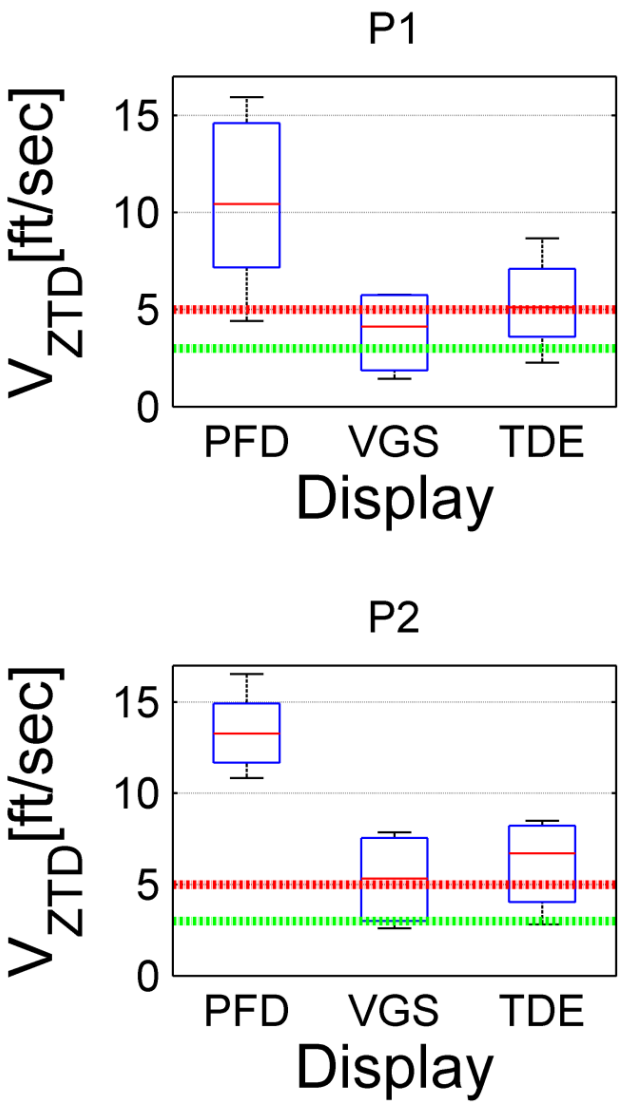


Fig. 14 V_{ZTD} performance for 3 displays and 2 pilots in DVE conditions

reported negative aspect of the runway outline symbology was that it's disappearance below 100ft caused a brief disorientation as the pilot's attention shifted from it to the guidance cue. Clearly such disorientation is undesirable given the aircraft's proximity to the ground at that time.

The TDE HUD also gave a significant improvement in performance vs. the PFD cases, albeit marginally less significant than the VGS (Wilcoxon signed-rank test p = 0.005). The lack of a runway outline cue was highlighted as a deficiency of the TDE, in addition to the relatively basic flare anticipation and initiation cues. The workload and controllability ratings (Fig. 15) reflect the pilots' preference for the

VGS HUD in this visual condition. P1 again awarded a high controllability rating to the TDE, partly due to the aforementioned display clarity issue and partly due to the less predictable nature of the guidance cue. It should, however, be noted that the pilot accompanied the rating with the comment "*but it did get me down safely every time".*

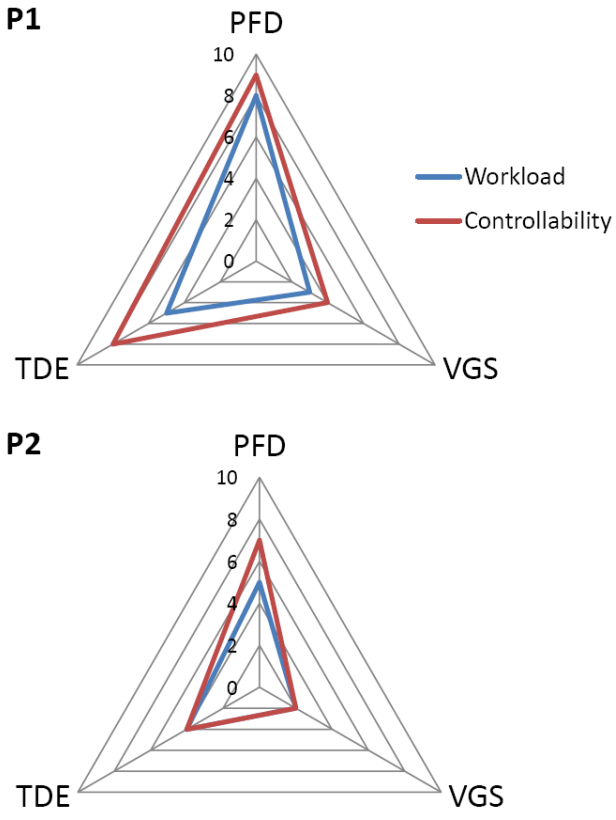


Fig. 15 Workload and display controllability for 3 displays and 2 pilots in DVE conditions

7 Conclusions and Recommendations

A study was undertaken to design and implement a flare guidance algorithm based on a 3-stage constant $\dot{\tau}_h$ strategy ("Tau Dot Error" – TDE). In order to perform piloted evaluation of this concept, the algorithm was used to drive a set of HUD symbology which was intended to emulate the "look and feel" of a conventional in-service example. This study built on the findings of Ref. [4], which implemented a flare algorithm based on a 2-stage constant $\dot{\tau}_h$ strategy with some success.

The TDE was evaluated against a baseline head down display (PFD) and an in-service example HUD (VGS). Vertical velocity at touchdown, V_{ZTD}, was used as the main performance metric for the flare manoeuvre. The results showed that the TDE HUD improved flare performance for all of the pilots tested in both good (GVE) and degraded visual (DVE) conditions c.f. the PFD flights. For 2 of the 3 pilots tested in the good visual conditions the TDE gave improved performance c.f. the VGS HUD. Pilot rating scales were also used to assess the workload and controllability associated with each display. For the GVE, 2 of the 3 pilots tested gave ratings which indicated the TDE to be "*Satisfactory without improvement"*. In the DVE condition, the TDE HUD improved performance for both of the pilots tested, albeit not to the same extent as the VGS. Pilot P1 reported difficulty in distinguishing between the symbols on the HUD, which contributed to the higher ratings awarded. As this issue was not reported by the other pilots and no problems were found with the experimental setup, it remains unexplained.

A secondary objective of the study was to address the limitations of the flare command algorithm developed in [4]. Specifically, the symbology used for the evaluation of the previous concept did not compare favorably to the in-service HUD example, leading to some confusion for the pilots when following the guidance cue. Two of the three pilots tested in this study reported no issues with the symbology; the third reported clarity issues with both HUDs. For this reason it can be concluded that this study made significant progress in fulfilling this objective.

In terms of recommendations, the pilot comments suggest that although the TDE was able to provide adequate performance, there is still scope for improvement. In particular it was reported that the simple text-based flare anticipation cue "standby" was not as intuitive as the moving anticipation symbol on the VGS. Secondly, the dynamics of the flare guidance cue were reported to be slightly unpredictable before the pilot became familiar with the TDE. This suggests that further tuning may be

required to improve the dynamics. Both of these issues will be addressed in the next design iteration.

Acknowledgements

This study was conducted at the Flight Science & Technology Research Group, a part of the Centre for Engineering Dynamics, University of Liverpool, UK. The research is funded as part of a Doctoral Training Award from the Engineering and Physical Sciences Research Council (UK). The authors would like to thank the Pilots and Simulation Laboratory Manager Dr Mark White. Finally, thanks to Professor Gareth Padfield for his guidance in the role of supervisor and mentor.

References

[1] Anon, *Aviation Safety Network*, [cited; Available from: **http://www.aviation-safety.net/about/**.]
[2] Benbassat, Abramson, ―Landing Flare Accident Reports and Pilot Perception Analysis‖, *International Journal of Aviation Psychology*, 2002, 12:2,137 — 152
[3] Pratt, J.M., *The Private Pilot's License Course*, AFE, 2007, ISBN 1 874783 73 X
[4] Jump, M, ―Prospective Sky Guides: Developing Guidelines for Pilot Vision Aids‖, Ph.D Dissertation, Department of Engineering, University of Liverpool UK, 2007
[5] Wisely, P.L., Bartlett, C.T., ―Gate to Gate Operation of a Civil Transport Head Up Display‖, *SPIE, Cockpit Displays VII: Displays for Defense Applications*, Vol.4022, 2000, pp. 392-398.
[6] Anon, *American Airlines Historical Information* [cited; Available from: **http://www.aa.com/**.]
[7] Padfield, G.D., D.N. Lee, and R. Bradley, "How Do Helicopter Pilots Know When to Stop, Turn or Pull Up?" *Journal of the American Helicopter Society*, 2003. 48(2).
[8] Lu, Padfield, Jump, "Optical Tau in Boundary-Avoidance Tracking – A new Perspective on Pilot-Induced Oscillations*"*, *European Rotorcraft Forum*, 2010.
[9] Jump M and Padfield GD, ―Development of a nature-inspired flare command algorithm‖, *Guidance, Navigation and Control Proceedings, AIAA Guidance, Navigation and Control Conference and Exhibit*, Marriott Hilton Head Beach and Golf Resort, Hilton Head Island, South Carolina: 2007
[10] Lee, D.N., et al., ―Visual Control of Velocity of Approach by Pigeons When Landing‖. *Journal of Experimental Biology*, 1993. 180: p. 85-104.
[11] Gibson, J.J., *The Ecological Approach to Visual Perception*. 1986, Hillsdale, New Jersey: Lawrence Erlbaum Associates.
[12] Lee, D.N., ―Aerial docking by hummingbirds‖. *Naturwissenschaften*, 1991. 78: p. 526-527.
[13] Lee, D.N., ―A theory of visual control of braking based on information about time-to-collision‖. *Perception*, 1976. 5(4): p. 437-459.
[14] Ridgway, G.R., Jump, M, ―Arresting a Descent: What Visual Cues do Pilots Use to Initiate a Flare Manoeuvre‖, *Guidance, Navigation and Control Proceedings, American Institute of Aeronautics & Astronautics*, Portland, Oregon, AIAA-2011-6687, 2011
[15] Heffley, R.K., T.M. Schulman, and W.F. Clement, ―NASA-CR-166404: An Analysis of Airline Landing Flare Data Based on Flight and Training Simulator Measurements‖ NASA, 1982
[16] White, Padfield, ―The Use of Flight Simulation for Research and Teaching‖, *Collection of Technical Papers - 2006 Atmospheric Flight Mechanics Conference*, v 2, p 954-967, 2006.
[17] Anon, FLIGHTLAB Theory Manual (Vol. One)‖. March 2004, Advanced Rotorcraft Technology: Mountain View, California.
[18] Bickerstaffe, I.H. ―Portrait of Landscape‖. *Proceedings of IMAGE Conference*. 1998. Scottsdale, Arizona.
[19] Anon, ―Presagis VAPS QCC Data Sheet‖, available from *http://www.presagis.com/*, 2008
[20] Roscoe, A.H. and G.A. Ellis, ―A Subjective Rating Scale for Assessing Pilot Workload in Flight: A Decade of Practical Use‖. Royal Aircraft Establishment TR 90019. 1990
[21] Haworth, L.A. and R.L. Newman, ―Test Techniques for Evaluating Flight Displays‖, NASA TM 103947, 1993
[22] Anon, "JAR-FCL 3; Flight Crew Licensing (Medical), Civil Aviation Authority, UK, 2006

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

User-Centered-Design of a Human-Machine-Interface for the Tower Controller

Jörg Bergner and Christoph Schmand
DFS Deutsche Flugsicherung GmbH, Germany

Thomas Hofmann and Christina König
Technische Universität Darmstadt, Germany

Hendrik Ebert
delair ATS GmbH, Germany

Keywords:
Air Traffic Control, Tower Controller, Human-Machine-Interface, Ground Radar, User-Centered-Design

Abstract

The current paper reports on the design of a human-machine-interface (HMI) for the local controller at a hub-airport. The aim of the investigation was to integrate related information that originates from separate sources in a concise and sensible way. The HMI consolidates arrival and departure planning-information, ground radar and stop-bar switchboard functionality in a single 30" high-resolution display with optical-touch overlay.

The development followed a multidisciplinary approach combining operational, software, ergonomic, and product design expertise. To assure acceptability by the future users, operational controllers were involved right from the start. During the iterative and incremental design process the real-time simulation environment of DFS Deutsche Flugsicherung GmbH, Langen was used to evaluate the design-steps of the HMI.

1 Introduction

At the large hub-airports, the continuous increase in traffic count especially when combined with a complex runway layout results in a tremendous workload for air traffic controllers [1], [2]. The primary means by which a tower controller keeps track of the current position of an aircraft is by looking out the window of the tower cab. But additionally, he also has to monitor a lot of information distributed over an assembly of individual displays, each for every system of a specific domain, like weather, radar, departure- and arrival information. The sequence in which the information is needed by the controller while he is coping with his tasks is commonly not reflected by the segregated arrangement of the information.

The recent trend towards enhanced controller assistance- and decision-support tools (e.g. Advanced Surface Movement Guidance and Control System (A-SMGCS), Airport Collaborative Decision Making (A-CDM), Arrival-and

Departure Management (AMAN, DMAN)) leads to even more information that has to be presented to the controller. Typically, this has been achieved by incorporating additional displays into the controller working position but this is difficult during continuous 24h operation and often simply constrained by the lack of additional space.

The aim of this investigation therefore was to develop an HMI that integrates information from various different sources, improves legibility and readability, harmonises interaction and is tailored to the specific needs of a tower controller under heavy workload. Consequently, the local controller at the major hub-airport Frankfurt/Main was used as an example.

2 Background

Eurocontrol has investigated the integration of the various airport airside system components into an Integrated Tower Working Position (ITWP) [3]. The study included a survey of the tower environment at many different-sized airports, the definition of high-level technical, operational, human-factors and safety requirements and an exemplary prototype (Fig. 1).

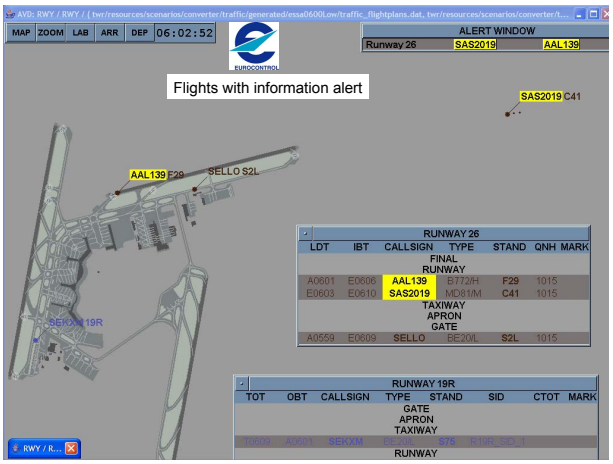


Fig. 1 Eurocontrol prototype: Integrated tower working position (ITWP)

The approach in the present investigation was slightly different as it aims specifically at controller positions under heavy workload and a successive implementation.

In a first step the focus was on the integration of arrival and departure flight-plan data since the coordination of the arrivals and departures on the main runway is a central part of the local controller‘s task and until today usually is provided by two separate systems that use different representations for the data (see Fig. 2 and Fig. 3).



Fig. 2 Monochromatic and alpha-numeric display of the Departure Coordination System (DEPCOS)

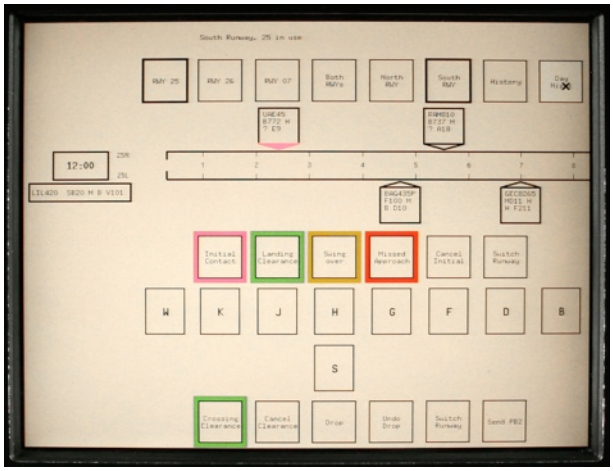


Fig. 3 Timeline representation of the arriving aircraft on the Tower-Touch-Input-Device (Tower-TID)

Considering the presently used systems for the arrival and departure information an integrated HMI for the planning information (TWR-HMI) was designed and evaluated. The central element of the HMI is a timeline representation of the estimated arrival time and accordingly the expected time of departure (Fig. 4). The present status or ATC-clearance of an aircraft is colour-coded.

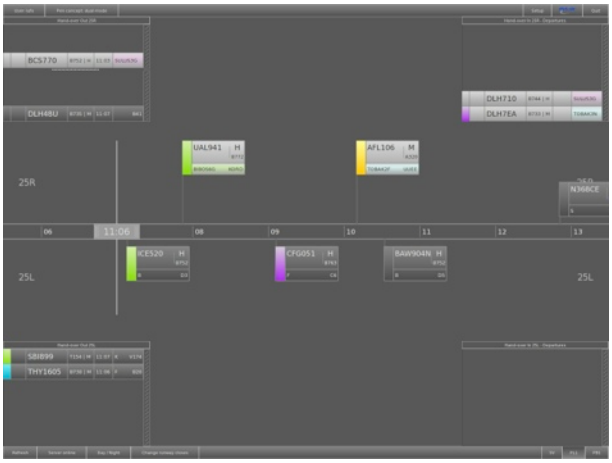


Fig. 4 Screenshot of the integrated planning HMI

The interface was designed for the WACOM Cintiq 21UX touch-screens. Since these displays are to be used with a grip-pen, the HMI was designed so that all necessary interactions can be conducted solely by using this pen. No additional input-device like a keyboard or mouse is required. Details on the development of this integrated planning HMI are given in [4].

The design was honoured with the iF design award 2009, the red dot award 2010 and nominated for the Designpreis der Bundesrepublik Deutschland 2011. A refined version of the prototype went into operation in the new control tower at Frankfurt International Airport in 2011.

Following the idea of an incremental implementation, this paper describes both the second step which is the integration of flight-status information and several other data and functionalities into the ground radar display (GND-HMI) as well as the third step which is the integration of both planning and ground HMI into a single, high-resolution, finger-touch enabled 30“ display.

3 Method

To assure acceptability by the future users, operational controllers were involved right from the start. The HMI was developed using a multi-disciplinary approach combining operational, software, ergonomic, and product design expertise.

3.1 Test environment

During the iterative and incremental design process the real-time simulation facilities of DFS Deutsche Flugsicherung GmbH, Langen were used to evaluate the design-steps of the HMI (Fig. 5). To mimic the complex operational procedures at the Frankfurt/Main International Airport the tower simulator was coupled with an approach and apron simulator as well.

Pseudo-pilots were involved to handle radio- and voice communication and to pilot the aircraft according to the controllers instructions. All simulations were run by operational controllers and supported and supervised by industrial psychologists.



Fig. 5 Real-time simulation facility

3.2 Implementation of the HMI prototype

The HMIs are realised using the QT 4 graphic library from Trolltech/Nokia in SVG. This allows for a flexible graphic design of the HMI in a standardised graphics format.

A client-server architecture was chosen with one server process dealing with the business logic of the flight strips and one with the graphic generation. The actual rendering of the SVG is performed for every single display by a dedicated thin-client. For the communication between the clients and the server the SOAP protocol is used, which is supported by QT through special libraries.

3.3 Evaluation

Each design-step was evaluated during three simulation runs of 90 min each. For each simulation run the local controller position was typically taken by a different controller. After a brief introduction to the HMI and the possible interactions, the controllers were supported and supervised by industrial psychologists who logged all their difficulties with the HMI as well as any suggestions for improvement.

Additionally the controllers were asked to verbalise their thoughts and considerations while working with the HMI. This special ergonomic method is described in [5]. Furthermore the industrial psychologists used a technique called the observer-interview. Findings made during the observation are promptly discussed with the controller. Thus they can immediately be verified in the context of the situation where they occurred. To complement this data every run closed with a general debriefing and a guided interview of the controller with the focus on the design of and the interaction with the HMI.

After the careful review of the evaluation results, modifications of the HMI where derived and implemented, thus leading to the next design-step. Over the course of this study a total of four design-steps, each with a typical duration of about a month, have been conducted.

4 Results

4.1 Ground radar display

The ground radar HMI (GND-HMI) was designed to complement the TWR-HMI and its look-and-feel. The first design iteration introduced the display of the colour-coded status information (e.g. initial contact, cleared to land, line-up, etc.) in the label of an target analogue to the TWR-HMI. Since the simulator only provided track data, targets are shown as aircraft symbols.

The presentation of colours takes into account the fact that they had to be easily recognisable, differentiated and non-dazzling for both day and night displays and different ambient light conditions. Since only about five colours can be easily distinguished at a glance, the application of colours was carefully chosen.

The differentiation between arrivals and departures is achieved not only by the information that is displayed but also by the background colour of the label. Departures are light grey, while arrivals are dark grey.

To reduce the risk of label overlap a cropped enlargement of the area where a departure-queue might build up supplements the overall view (Fig. 8). Whereas the label in the overall view only contain the call sign and in case of an arriving aircraft also the taxiway (Fig. 6), labels in the enlargement are bigger and additionally contain info on the standard departure route (SID) as shown in Fig. 7.

During the course of the study different label sizes have been investigated. It was found that their size has to be adapted to the available screen size and the expected traffic density, carefully balancing readability versus the risk of label overlap.



Fig. 6 Different label sizes for arriving aircraft

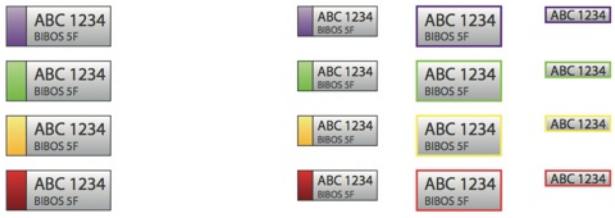


Fig. 7 Different label sizes for departing aircraft

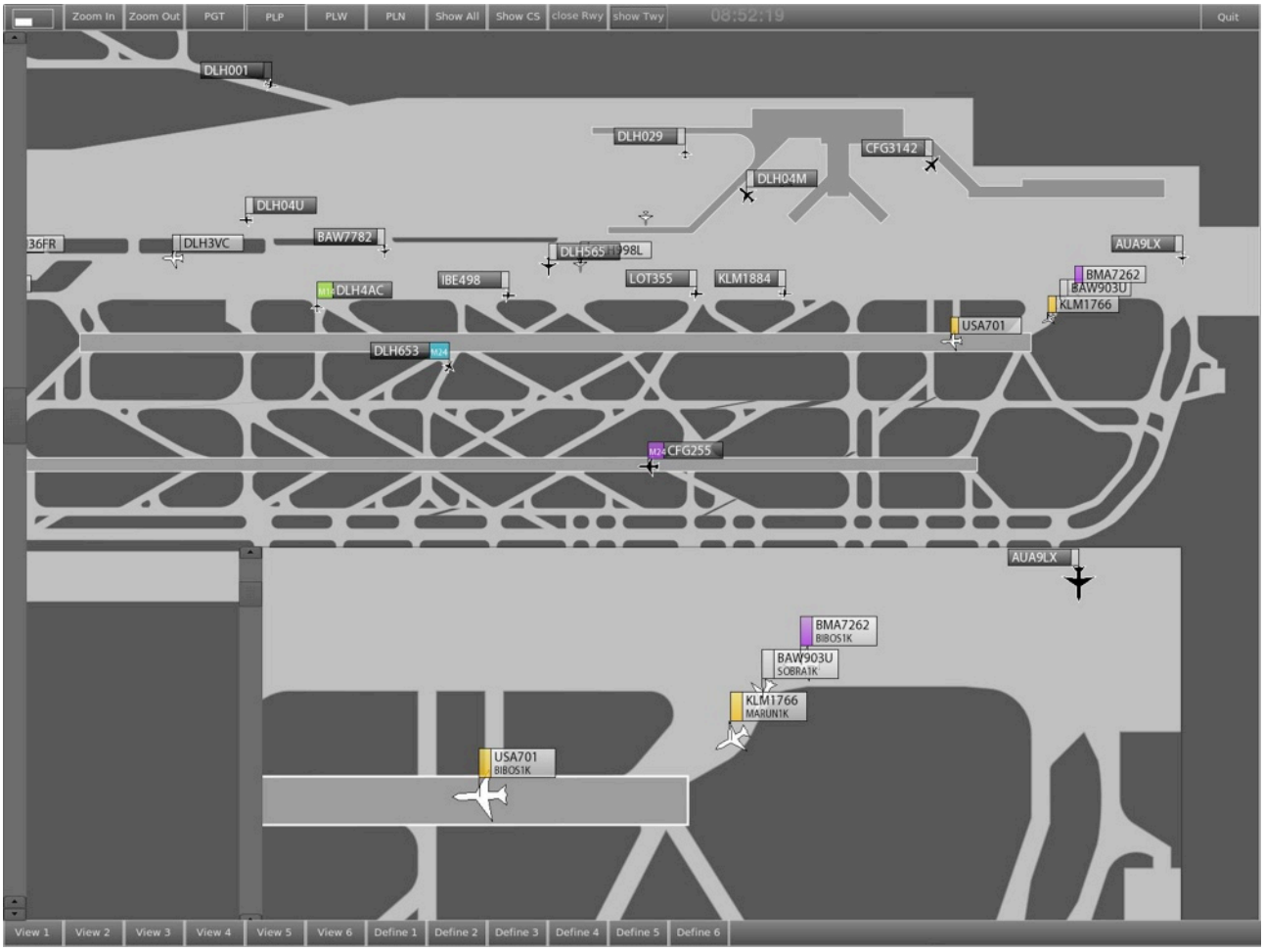


Fig. 8 Screenshot of GND-HMI

Further information and functionalities are integrated into the GND-HMI using the layer metaphor. Besides the optional display of the taxiway nomenclature, stop bars can be switched and runways can be marked as closed.

Stop bars

In the first implementation of the stop-bar functionality both interaction and display area were the same size (Fig. 9). Since the large red coloured areas might be confused with a closed taxiway, the display area of an stop bars status was reduced whereas the interaction area was kept unchanged.

The interaction was designed so that with one click on the interaction area the stop bar turns green and a timer appears that gives the remaining time in seconds after which the stop bar automatically turns back to red (Fig. 10).



Fig. 9 Cropped enlargement of the GND-HMI with active stop-bar and taxiway nomenclature layer

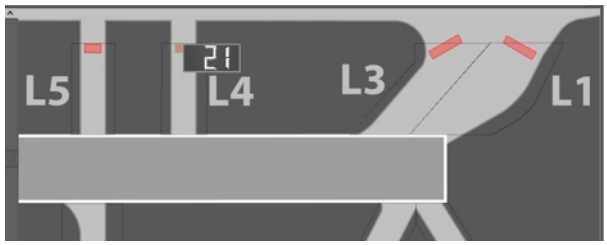


Fig. 10 Cropped enlargement of the GND-HMI with stop-bar timer

Runway closure

To visually indicate the temporary closure of a runway it can be highlighted in red (Fig. 11). This is achieved by opening a small window showing a simplified representation of the airports runway layout and clicking the accordant runway (Fig. 12).

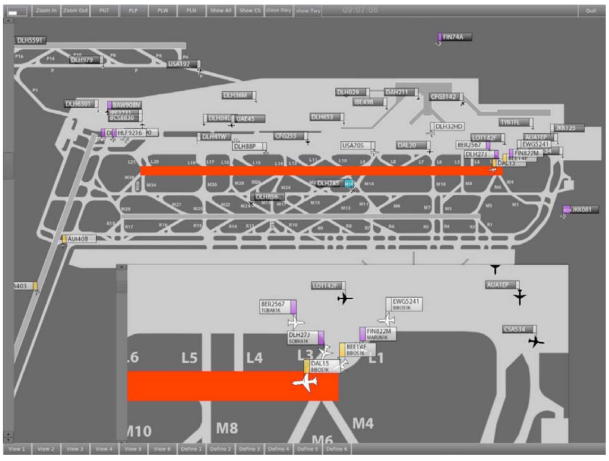


Fig. 11 GND-HMI with temporarily closed runway

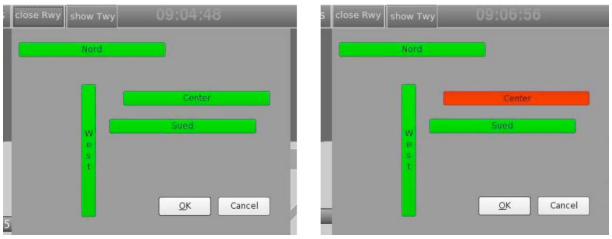


Fig. 12 Menu window with simplified runway layout for selection of closed runway

4.2 Integrated planning and ground radar display

After integrating arrival and departure information into one 21“ WACOM 21UX display and accompanying that with a second WACOM that integrates ground radar and the lighting panel, the next logical step was to join both of them on a large touchscreen.

When choosing a suitable display size, first of all the view of field and the boundaries of hand reach have to be considered. As can be seen in Fig. 13, the optimum size for such a display appears to be approx. 27-30“.

The additional display area of larger displays not only lies outside of reach but also outside of the viewing area. In addition, as top to bottom viewing angles may still be restricted by design since manufactures commonly focus on the horizontal viewing angle, the usable display area might be even smaller. Thus great care has to be taken when arranging information on the available display area of large screens.

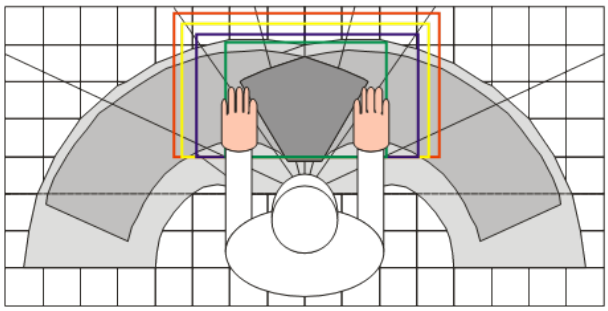


Fig. 13 Field of view and reachable area compared to 21“, 27“, 30“ and 32“ displays

An additional goal was to evaluate the use of finger-based touch interactions as an alternative to the previously used pen-based WACOM touchscreens. Despite a wide availability of large touchscreens in general, their display resolution usually is constraint to 1080p. With sizes of about 32“ this leads to a pixel density of only approximately 80 dpi.

As a consequence, it was decided to equip a high-resolution 30“ computer display (NEC LCD3090WQXI) with a native resolution of 2560x1600 with an glossy, anti-reflex coated, optical touch overlay. This allowed for finger based touch interactions and at the same time lead to a pixel density of 101 dpi which is even greater than the 95 dpi of the WACOM.

Optical imaging is a relatively new touchscreen technology, in which two or more image sensors typically are placed at the corners of the screen. Opposite to them infrared back-lights are positioned in the camera's field of view. A touch on the surface of the screen shows up as a shadow and each pair of cameras can then be used to locate the coordinates of the touch.

During the investigation three different installation angles of the display (15°, 30 & 60°) have been evaluated. Figure 16 shows the angle of 30° which was preferred by the operators. The steep angle of 60° provided a more perpendicular view on the screen and thus a increased legibility. However, the prolonged use of the arm held in an unsupported horizontal position is well known to cause fatigue (gorilla arm). On the contrary, the shallow angle reduced legibility but more important it significantly increased the risk of unintended interactions.

The great flexibility of this prototype enables the analysis of a variety of different layouts of the time-line, the ground radar and additional representations of different kinds of information like e.g. the meteorological conditions and aerodrome configuration.

Display of additional information

Since meteorological conditions at the airport are of great importance for the safe operation, today a dedicated display provides the controller with all relevant weather related parameters. Figure 14 shows a concise layout of this information making a separate display obsolete.

In a similar way the controls for the instrument landing system (ILS) of the airport and its status information have been arranged to be integrated into the 30“ display (Fig. 15).

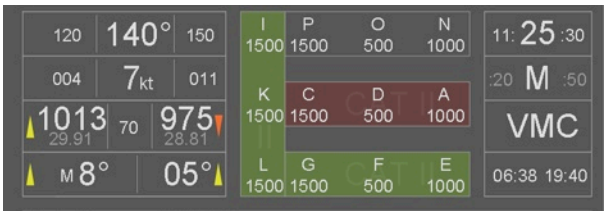


Fig. 14 Arrangement of weather related information

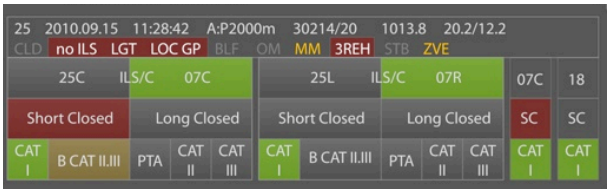


Fig. 15 Arrangement of ILS controls and status information

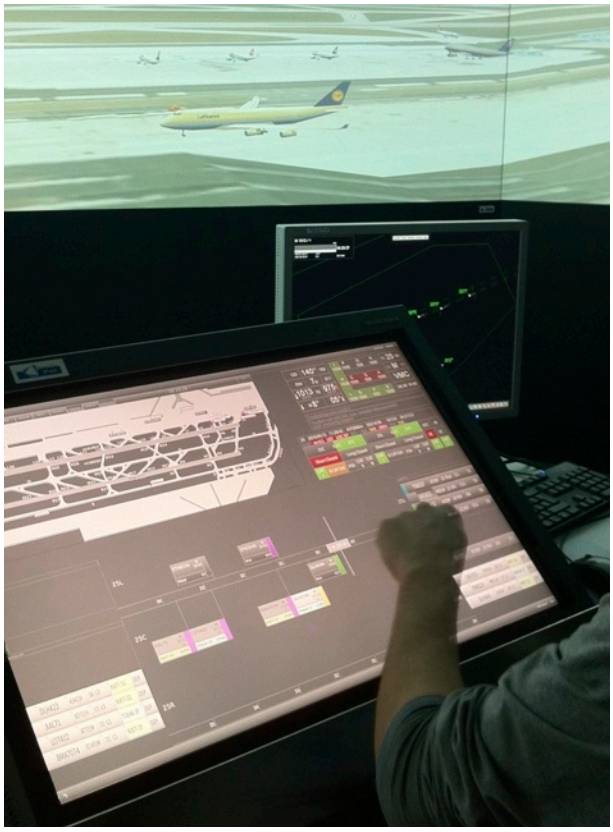


Fig. 16 Evaluation of the integrated HMI with weather related information and ILS control panel

Figure 17 shows a screenshot of a layout for the local controller who is responsible for the parallel runway at Frankfurt International Airport. This particular configuration excludes the display of additional weather or ILS information in favour of providing as much space as possible for the ground radar. This allows for the use of the stop bar functionality without having to zoom the ground radar view or the use of an additional enlarged crop. Thus, this 30“ display fully integrates the functionality previously split across four separate displays which are the arrival & departure information, the ground radar and the lighting panel.

5 Conclusion

The evaluation of the HMI confirmed the general approach of consolidating related information that originates from separate sources into a large, high-resolution, wide aspect-ratio touchscreen. The concise and sensible integration helps to reduce the number

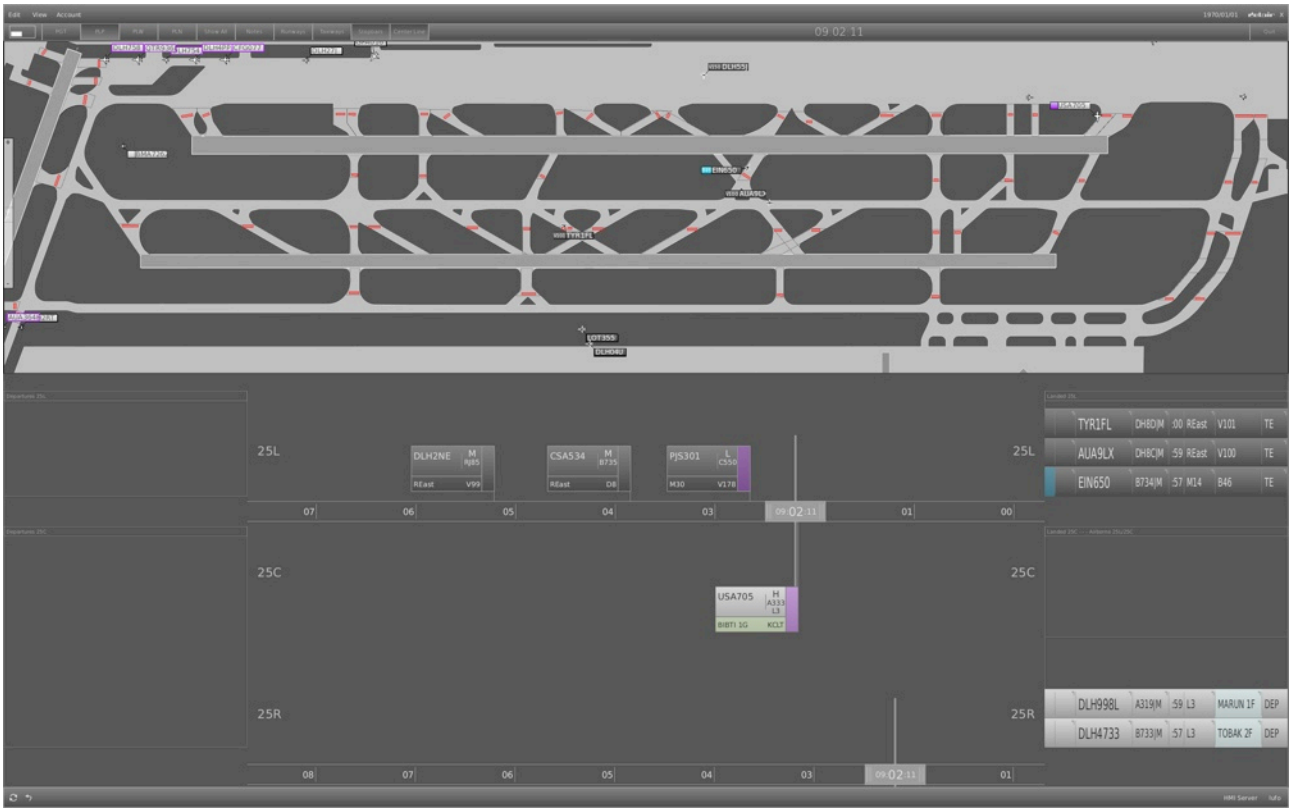


Fig. 17 Screenshot of the integrated HMI with planning information, ground radar and stop bar functionality

of necessary displays at the controller working position and supports an efficient visual scan pattern.

The use of a finger appears to be the most natural and intuitive way of interacting with a touchscreen. Nevertheless, the presently available touch technology that does not degrade the optical display quality and is sufficiently precise when used with large screen sizes basically is based on the photosensor principle. This in turn introduces the risk of unintended interactions when e.g. the hell of the hand or a shirt cuff reaches the trigger distance of the photosensor or the threshold of the camera.

The great flexibility of the prototype enables the analysis of a variety of different layouts of the time-line, the ground radar with stop bar switchboard functionality and additional representations of different kinds of information like e.g. the meteorological conditions and aerodrome configuration.

However, to utilise the full potential of this approach, the layout of a integrated HMI has to be carefully adapted to the individual area of responsibility and the information needs at a specific working position.

Acknowledgments

Being sub work-packages of the research projects „Wettbewerbsfähiger Flughafen“ and „Innovativer Airport“ this work was in part funded by the German Federal Ministry of Economics and Technology (BMWi) within the forth national aerospace research program (LuFo IV).

Our special thanks are due to Wolfgang Witteborg (DFS) who served as subject matter expert and all participating air traffic controllers from Tower Frankfurt for their invaluable feedback. We also would like to thank all colleagues that have been involved in the preparation and realisation of the extensive real-time simulations.

References

[1] J. Leonhardt, „Implementation of Critical Incident Stress Management at the German Air Navigation Services," Human Factors Impacts in Air Traffic Management, edited by Kirwan, B.; Rodgers, M. and Schäfer, D., Ashgate Publishing Limited, Hampshire, 2005, pp. 207-225.

[2] C. Manning and E. Stein, „Measuring Air Traffic Controller Performance in the 21st Century", Human Factors Impacts in Air Traffic Management, edited by Kirwan, B.; Rodgers, M. and Schäfer, D., Ashgate Publishing Limited, Hampshire, 2005, pp. 283-31.

[3] Eurocontrol „Integrated Tower Working Position Initial Study Report, version 1.0", Eurocontrol Experimental Center, Bretigny, 2006.

[4] J. Bergner, C. König, T. Hofmann, H. Ebert, „An Integrated Arrival and Departure Display for the Tower Controller", in Proc. 9th AIAA ATIO Conf., 2009, submitted for publication.

[5] A. Genov, „Iterative Usability Testing as Continuous Feedback: A Control Systems perspective," Journal of Usability Studies, Vol. 1, No. 1, 2005, pp. 18-27.

Towards a Wake Encounter Avoidance and Advisory System (WEAA) – Conceptual Design and First Implementations

T. Bauer* **and D. Vechtel****
German Aerospace Centre (DLR), Institute of Flight Systems,
Lilienthalplatz 7, D-38108 Braunschweig, Germany

(*) *tobias.bauer@dlr.de*
(**) *dennis.vechtel@dlr.de*

Keywords: *wake vortex separation, wake encounter avoidance, flight safety, trajectory planning, situational awareness*

Abstract

The encounter of wake vortices of preceding aircraft can present a substantial hazard to the follower. Separations as procedural means are applied to avoid such encounters; however, the multitude of parameters governing vortex evolution makes them overly conservative (and thus capacity limiting) in the majority of cases without preventing incidents altogether.

German Aerospace Centre DLR is working on a Wake Encounter Avoidance and Advisory (WEAA) system to allow pilots to avoid potentially dangerous wake vortex encounters by a small scale tactical evasion manoeuvre. Such manoeuvre tries to adhere to the planned flight track as closely as possible and is not intended to require ATC permission.

Work on the system is on-going with a larger time horizon but initial solutions will be presented: the system conceptual design and selected components such as wake vortex prediction and hazard evaluation, avoidance trajectory generation, and a human machine interface for increased situational awareness.

1 Motivation

The encounter of a sufficiently strong wake vortex of a preceding or crossing aircraft can pose a threat to the follower as a consequence of

- significant induced (linear and angular) accelerations leading to injuries of passengers / crew or structural damages,
- induced forces and moments (often predominantly rolling moments) leading to a poorly controllable flight condition or even (temporary) loss of control,
- hazardous flight path deviations resulting from the above effects, including altitude loss in the downdraft of the wake.

Wake vortices are often encountered in departure and approach operations – even close to the threshold, which makes them particularly dangerous due to ground proximity – but encounters also occur during cruise. Separations are defined under IFR to prevent hazardous encounters [17]; these have generally proven acceptably safe but the level of safety still could be increased

- under VFR, or
- in extreme meteorological situations when vortices behave abnormally, or
- in situations where separations are inadvertently violated, cf. Learjet 45 accident in Mexico City in 2008 [24].

In these cases an encounter avoidance system could improve flight safety by providing a safety net function. As current wake vortex separations limit (especially approach and departure) capacities unnecessarily in most cases, such a system could alternatively assist in increasing airspace capacity because of its mitigating effect on the encounter probability.

The aim of WEAA (Wake Encounter Avoidance and Advisory) system development as undertaken by DLR in the framework of its internal project "Wetter und Fliegen" (Weather and Flying) is to support pilots in avoiding hazardous small scale atmospheric phenomena such as wake vortices without necessity to change the flight plan or violate the active clearance. To that effect, potentially hazardous conflicts are detected and where possible evasive manoeuvres are defined, guided and monitored. Increasing the pilots' situational awareness in case of a predicted encounter is one of the main design objectives.

Where circumstances do not allow encounter avoidance, timely alerts can help the pilots to prepare themselves. Here the potential benefit of dedicated wake encounter alleviation functions of the flight controller has already been successfully demonstrated in flight experiments [1][2][3]. Such functions can be seamlessly integrated with WEAA even if they are not part of the avoidance system.

WEAA is initially intended as a safety net i.e. to improve the existing level of safety. It is conceivable to use it alternatively as a means to increase capacity in as much as separations may be reduced for future aircraft so equipped; however, the system itself is not designed to suggest appropriate separations.

The approach is in principle extendable to other phenomena such as small fields of clear air turbulence (CAT) but, due to the lack of suitable instruments for their detection, investigations currently focus on wake vortices. DLR work in this respect is complementary to earlier (I-WAKE [4], FLYSAFE) and on-going (SESAR [5]) efforts.

2 Initial Considerations for System Concept

Due to the stochastic nature of wake evolution and transport, zones of potential danger become very large over time especially if meteorological parameters and those of the generating aircraft (weight, airspeed, configuration) are unknown. The complete avoidance of these zones of uncertainty is not compatible with current operations and traffic densities. Thus a viable avoidance system requires a more exact knowledge of the actual wake position, which is only attainable with improved prediction based on more accurate generator and meteorological parameters and/or remote detection of the wake position.

Remote detection of wake vortices could be performed using forward-looking LIDAR sensors. Unfortunately most of these sensors are only able to measure the line-of-sight (LoS) component of the flow field velocity. As wake encounters often occur when the involved aircraft have the same nominal flight path (as during approach and landing) the main vorticity components of the wake are not measurable with LIDAR. Additionally the measurement resolution and update rate of current instruments at distances relevant for avoidance manoeuvres are below requirements by several orders of magnitude [21]. Conceptually, however, it appears desirable to combine the measurement with a prediction based on data of the leading aircraft, which might be received via Automatic Dependent Surveillance ADS-B. A combination of ADS-B based prediction and LIDAR measurement appears promising for this purpose [21].

Once position and spatial orientation of the wake are known an evasion trajectory can be calculated. Aim of system development is to initiate the manoeuvre early enough so that it requires only small accelerations and ideally is not perceived by the passengers. With the expected small deviations from the original flight path 30 to 40 s appear a sufficient tentative time horizon; this value will be further analysed in simulation and ultimately in flight experiments.

3 Avoidance Manoeuvre Design Restraints

The WEAA system enables avoidance of small scale phenomena with immediate return to the original flight path so by definition a change in the flight plan is not envisaged, nor does the small time horizon allow to request changes of the current clearance from ATC. This requires manoeuvres to be designed such that informing ATC about their execution suffices.

As a consequence, the airspace available for wake evasion is limited. In the absence of dedicated ATC procedures the idea is to restrain manoeuvring to the tolerable navigation deviations. In principle evasive manoeuvres can be performed either vertically or laterally, or by adapting airspeed only. All three options may be combined if compatible with the capabilities of the flight guidance and control functions.

3.1 Airspace and Navigational Restraints

Under normal conditions each aircraft has to acquire the desired flight path as accurately as possible given its capabilities. ICAO's Required Navigation Performance (RNP) concept [17][18][19] specifies a minimum navigational accuracy (Navigation System Error, NSE) to be met by the on-board equipment for 95% of the time, cf. Fig. 1. The concept includes on-board monitoring meaning an estimation of current NSE is available. Numerical RNP requirements are set by ICAO member states dependent on type of airspace and flight phase. Typical values are ±2 NM for cruise, ± 1 NM for arrival and departure and ± 0.3 NM for approach [19].

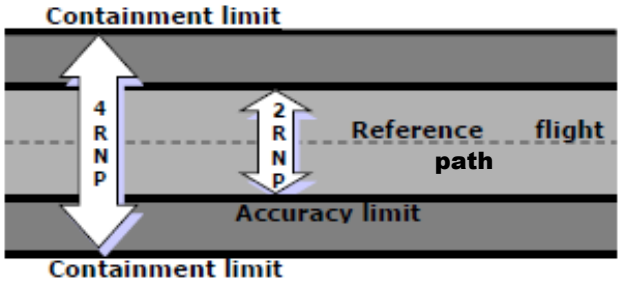


Fig. 1: RNP containment and accuracy limits [25]

Twice the width of the accuracy limit defines the containment limit where, assuming Gaussian distribution of NSE, the indicated position is located with a probability of 99.999%.

Taking into account horizontal navigation performance achieved by modern aircraft for most flight phases [25] and confidence intervals of wake vortex habitation volume predictions, see section 5.1, it appears that often sufficient lateral margin exists for manoeuvring, cf. Fig. 2.

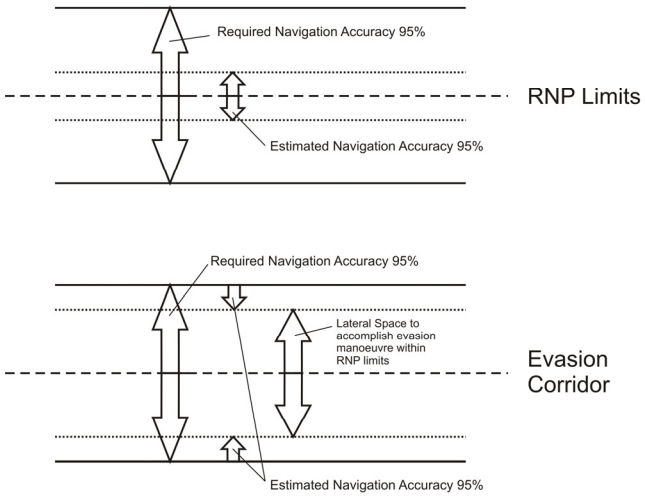


Fig. 2: Definition of lateral evasion corridor limits

The tolerances on vertical position deviation, listed in Table 1, are significantly smaller than those in lateral direction. This is due to the high accuracy of barometric height measurement and the much lower vertical separations. However, as the vertical motion of the vortices is limited this option is not precluded per se. Considering that ATC should be informed when the altitude deviates ±100 ft from the cleared flight level (FL), it is conceivable though that for this type of manoeuvre a quick, possibly automated go/no-go decision from ATC would be required.

Table 1: Tolerable vertical flight path deviations due to pilotage errors ([19] II-A-6).

	Level flight segments and climb/descent intercept altitude region of specified altitudes	*Climb/descent along specified vertical profile (angle)*
At or below 1 500 m (5 000 ft)	45 m (150 ft)	60 m (200 ft)
1 500 m to 3 000 m (5 000 ft to 10 000 ft)	73 m (240 ft)	91 m (300 ft)
Above 3 000 m (10 000 ft)	73 m (240 ft)	91 m (300 ft)

As to speed adaptations, true airspeed has to be maintained within a tolerance margin of ±5% [20]. On its own this option does not appear effective during the envisaged time horizon. Moreover, configuration and airspace dependent speed restrictions (such as the limit of 250 kts below FL100) need to be taken into account.

3.2 Interoperability Restraints

In addition to navigational and ATC requirements interoperability restraints must be observed during an evasive manoeuvre such that it is not impaired by triggering alerts from other systems. It is the intention of the WEAA system design that the independent safety net functions of terrain and traffic collision avoidance are not affected by the avoidance manoeuvre. Incidentally, these functions obviously need to have higher priority than wake avoidance.

Navigating in areas close to ground, mountains or other obstacles it must be assured that the performance of an evasive manoeuvre does not result in an (Enhanced) Ground Proximity Warning System ((E)GPWS) alert. Therefore at low altitudes or in mountainous areas information about the proximity of ground must be considered for the generation of the evasion trajectory, with regard to the fact that (E)GPWS not only takes into account height above ground but also terrain closure rates.

As to surrounding traffic, the WEAA system is not intended to handle existing TCAS conflicts, which take priority over wake avoidance. However, it must be assured that the wake evasive manoeuvre does not trigger new TCAS alerts during its execution. This does not necessarily mean that the wake evasion system must comprise a built-in TCAS functionality but the TCAS principle of operation needs to be taken into account for trajectory generation. In analogy to (E)GPWS, not only actual proximity of the aircraft is evaluated but also closure rates are observed.

3.3 Performance Restraints

When determining an evasion trajectory, the aircraft performance, manoeuvrability and passenger comfort (small accelerations and load factors) need to be taken into account. These restraints depend on the flight phase (e.g. in final approach passengers can be assumed to be seated with seatbelts fastened such that higher accelerations can be tolerable). The general aim is to achieve sufficient lead times for the manoeuvre such that, unlike TCAS, hard manoeuvring is not required. Beneficial for reaching this aim is that usually only small flight path deviations will be required.

4 WEAA System Architecture

With the above restraints in mind, a system architecture for WEAA has been developed, a high-level overview of which is presented in Fig. 3.

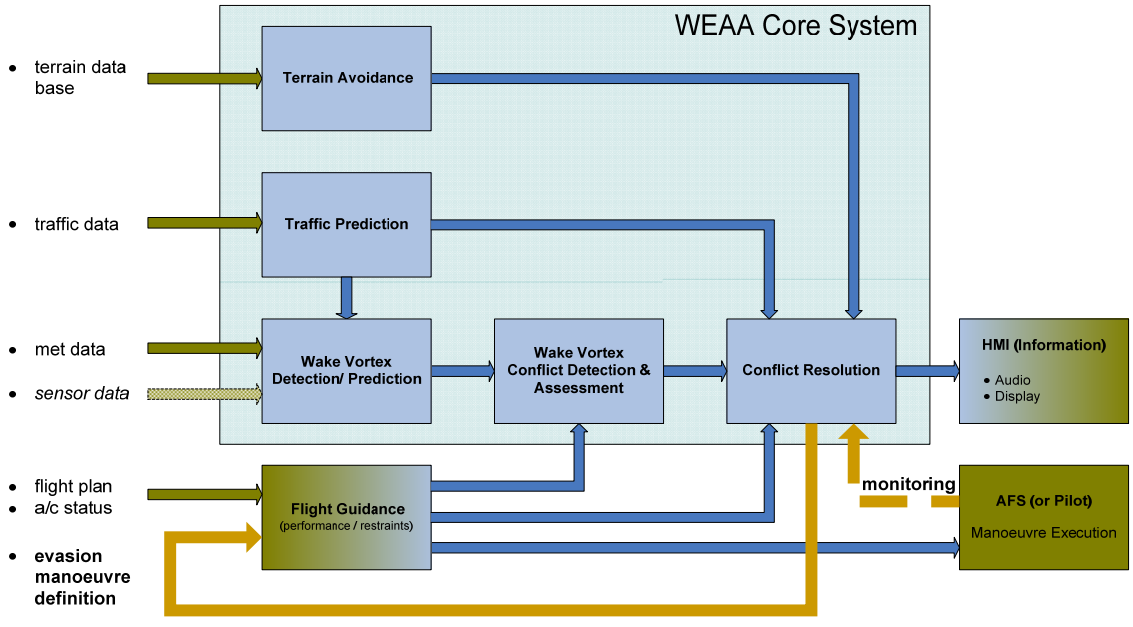


Fig. 3: Simplified top-level functional architecture of WEAA system

According to requirements, the following tasks have to be performed by the system:

- predict wake vortex transport and decay for surrounding traffic using meteorological data and information on the traffic (aircraft type, track, velocity; performance data and planned trajectories where available), possibly enhanced by remote wake vortex detection employing a suitable forward looking (LIDAR) sensor. Such prediction can include own vortices as they might be encountered under special circumstances;
- detect wake conflicts, using prediction of own trajectory and that of other traffic (assuming in absence of better information that its flight track is properly described by extrapolating the current flight path vector), followed by hazard assessment where required. These two functions are closely interconnected, see section 5.2;
- in case a conflict is detected, alert the flight crew and generate an evasion trajectory, taking into account terrain and traffic collision avoidance requirements as well as own ship flight performance, flyability restraints and passenger comfort;
- provide guidance for the necessary evasive manoeuvres to the pilots, e.g. on Primary Flight Display (PFD) and Vertical Speed Indicator (VSI);
- monitor manoeuvre execution (by pilot via flight director in manual flight or by AFS when autopilot is in use); and
- display an overview of the situation to increase the pilots' situational awareness, e.g. on the Navigation Display (ND).

The WEAA system constantly updates the trajectory upon reception of new input data. The stability of the prediction remains a major point for further investigations.

Two design options have been defined for the determination of necessary wake parameters:

- The basic design is based on prediction of the wake vortex location using information and sensors that are in principle available and onboard a transport aircraft today.
- The enhanced system design employs wake characterisation from prediction and vortex measurement by means of a suitable sensor such as LIDAR [12]. These have not yet achieved sufficient technological maturity [21] and their installation on transport aircraft is not foreseeable in the near future. Hence, this option is seen as a long term one but among the aims of further WEAA system assessment is to evaluate benefits and necessity of a LIDAR sensor.

Figure 5 shows a more detailed functional system break-down of the WEAA system with the two design options. The system architecture is as generic as possible but where necessary adapted to DLR's ATRA research aircraft (Fig. 4) which is the envisioned test bed for future in-flight system integration and evaluation.



Fig. 4: DLR's ATRA research aircraft

5 System Components and Initial Results

The WEAA system realisation, which is a work-in-progress, uses a modular architecture to allow a step-by-step implementation of components and to allow trade-off studies for different configurations. For instance, the benefit of using a LIDAR sensor in addition to a pure wake prediction will be evaluated.

Various existing and on-going component implementations are discussed hereafter.

5.1 Wake Vortex Prediction

Wake vortex prediction is effected with HOLZÄPFEL's P2P tool [6][7], a real-time capable probabilistic two-phase wake vortex prediction model. P2P takes into account effects of aircraft configuration, wind, wind shear, turbulence, stratification and ground proximity. It has been extensively validated on LIDAR measurements and numerical (LES) simulations [7][8], cf. Fig. 6. Alternatives to the P2P model could be NASA's APA [10] or PVM [11].

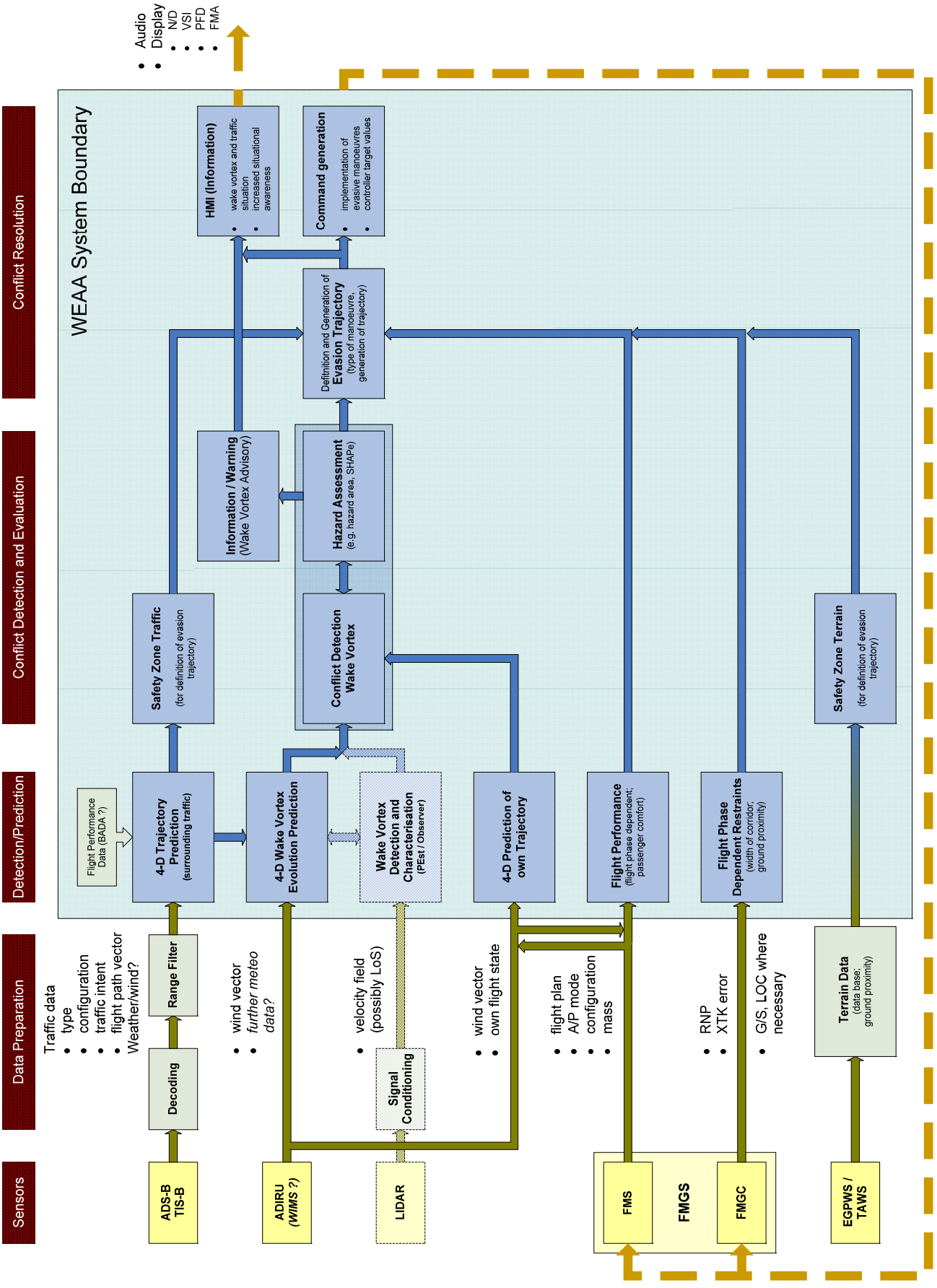


Fig. 5: Functional architecture of WEAA system (with optional LIDAR detection)
(refer to appendix for abbreviations)

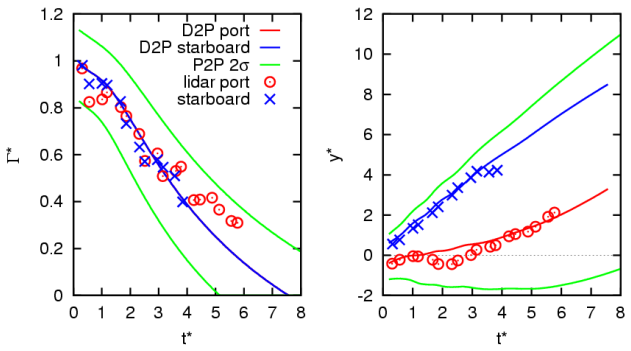


Fig. 6: Sample wake vortex prediction from P2P showing calculated nominal (D2P, red and blue lines) and measured (symbols) values for vortex decay Γ^* and lateral positions y^*. Green lines are associated 2σ confidence intervals for prediction (P2P) [9]

5.2 Hazard Assessment

WEAA's hazard assessment function uses the Simplified Hazard Area Prediction (SHAPe) method by HAHN and SCHWARZ [13][14][15] to quantify the severity of a potential wake encounter. In relation to the nominal wake vortex positions, areas are determined outside of which no hazard due to the wake exists (Fig. 7) i.e. where flight operations can be considered as safe and undisturbed.

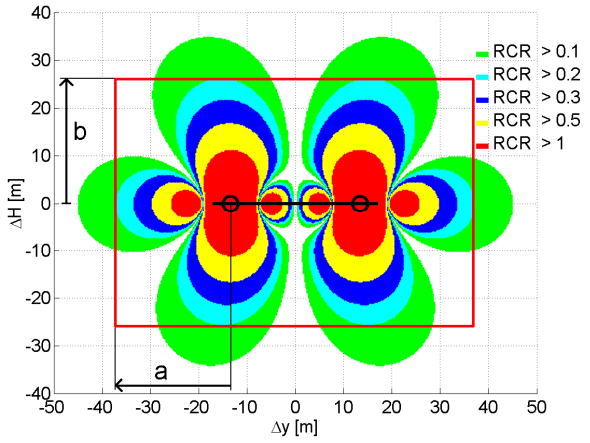


Fig. 7: Sample result from SHAPe hazard assessment showing a simplified hazard area (red rectangle) based on roll control ratio *RCR* (cf. Eq. 1) around nominal wake vortex positions (black circles) (modified from [16])

The hazard assessment criterion used by SHAPe is the roll control ratio *RCR* defined as

$$RCR = \left| \frac{C_{l,WV}}{C_{l\max,\delta_l}} \right| \qquad (1)$$

where $C_{l,WV}$ is the coefficient of the rolling moment induced by the vortex wake and $C_{l\,\max,\,\delta l}$ the coefficient for the maximum rolling moment obtainable by roll motivator deflection (roll control power). Numerical values for acceptable *RCR* have been derived experimentally from piloted trials [15].

The extent *a*, *b* of the hazard areas given by SHAPe (Fig. 7) is added to the wake vortex habitation volumes predicted by P2P. Hazard assessment thus is closely connected to conflict detection, as the permissible distances *a*, *b* to the assumed position of the vortex cores determine the size of the volumes for which conflict detection is necessary (with $a = b = 0$ meaning the wake does not represent a relevant hazard any longer). In addition, the level of hazard and hence *a*, *b* may depend on the encounter geometry.

5.3 Evasive Trajectory Generation using a Potential Field Based Method

The conflict resolution function, cf. Figs. 3 and 5, has the task to modify the planned flight path as little as possible while avoiding the hazardous zones (comprising wakes, terrain, and the safety zones around other traffic). Due to the limited manoeuvrability of a passenger aircraft many algorithms successfully used in unmanned aerial vehicle (UAV) technology cannot be applied. The method chosen uses a potential field approach. Danger zones are modelled by penalty barrier functions which are added to the nominal potential field function characterising the intended flight path. This approach assures avoidance trajectories close to the originally planned flight path and, by contrast to TCAS, automatic reacquisition of the nominal flight path after clearance of the danger zone.

For first studies the evasion options have been restrained to pure lateral or vertical motion with a decision algorithm based on encounter geometry, cf. Fig. 8. Motion in all three spatial dimensions plus adjustment of flight speed, i.e. 4-D evasion, is conceivable though. The trajectory generated by the algorithm is smoothed to ensure flyability.

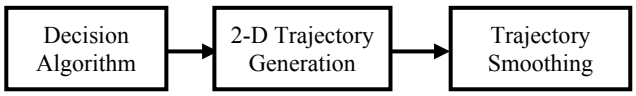


Fig. 8: Sequence of avoidance trajectory generation (modified from [22])

The algorithm has been successfully tested in flight simulation in cruise conditions [22] and will be further applied to other flight phases. An example is depicted in Fig. 9.

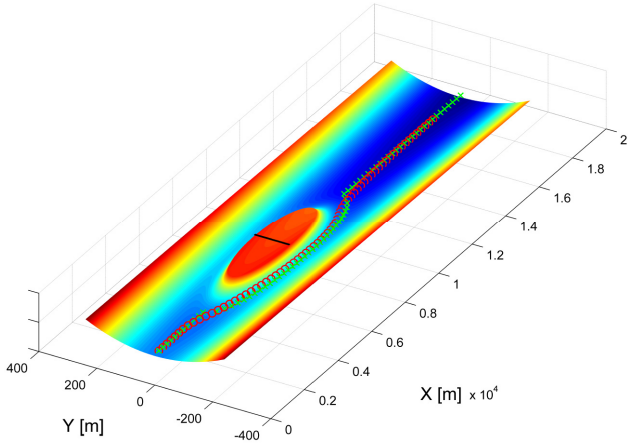


Fig. 9: Example of generated avoidance trajectory on a potential field represented by the surface [22]. Black line indicates wake segment to be avoided. Initial (green ×) and smoothed (red ○) trajectories are shown

5.4 Human-Machine-Interface

A human machine interface (HMI) for WEAA has been further developed from existing third party design work in the FLYSAFE project. Guidance symbology established for TCAS manoeuvres is reutilised as far as applicable. Flight Director and Primary Flight Display are used to guide the avoidance manoeuvre. Hazard zones, the track for the (lateral) evasive manoeuvre and headings currently to be avoided are depicted in the Navigation Display. The latter has been favourably evaluated for increasing situational awareness of the pilots [23], which is crucial even if the manoeuvre is executed by the autopilot. If available, the Vertical Situation Display (VSD) is also used.

An alerting logic and callouts supporting the manoeuvre have been defined adopting TCAS principles where possible. Positively evaluated in a fixed-based simulator study [23], this initial HMI design will be further refined.

6 Conclusion

DLR is developing a Wake Encounter Avoidance and Advisory System (WEAA) that allows tactical small-scale evasion to avoid possibly hazardous wake encounters. WEAA initially has a pure safety net function, assuring interoperability with existing safety functions TCAS and (E)GPWS / TAWS. The pilots' situational awareness is a key driver in system design and development.

DLR's objectives in WEAA development are a system proof-of-concept and in-depth investigation of selected components. This includes trade-off studies to assess the necessary level of detail in modelling and hence complexity of the system functions.

Pre-existing DLR knowledge has been exploited in the form of the P2P wake vortex prediction model, wake parameter identification from forward-looking sensor measurements (as a long term option), and severity assessment for a predicted encounter (SHAPe method). This means that several key system components are already available and well tested.

Initial development work has focused on the approach and cruise phases, encompassing a method for evasion trajectory generation using a potential field based approach, an investigation into forward looking sensor requirements for wake detection, and the enhancement of a pilot display concept for manoeuvre guidance and increased situational awareness. Current stand-alone implementations of these functions have been successfully tested in flight simulation. Further testing of HMI and wake prediction in a flight experiment with DLR's research aircraft ATTAS is imminent.

Further Work

On-going work includes further development of component functions such as conflict detection concept and algorithms as well as a refinement of the conflict resolution algorithms taking into account (E)GPWS & TCAS interoperability, passenger comfort, and aircraft performance. Next steps comprise enhancing

severity assessment (graded hazard assessment, non-roll dominated encounters) and further studies on enhanced wake characterisation using (LiDAR) measurements. The WEAA concept of operation will be assessed for all flight phases (current implementations focus on cruise flight and the approach phase). Different concepts for pilot assistance will be analysed with respect to workload and situational awareness.

System integration is a major task for the future; the proof-of-concept implementation in an engineering flight simulator is already on-going. As a perspective increasingly complete functions are planned to be evaluated in motion based simulation and during flight-testing with ATRA. These experiments will support an eventual benefit analysis under realistic conditions.

Acknowledgements

The authors gratefully acknowledge contributions from the following colleagues: Jobst Diekmann (trajectory generation via force field method), Christian Gall (HMI), Frank Holzäpfel (P2P wake vortex prediction model), Carsten Schwarz (SHAPe severity assessment).

Symbols

Γ^*	normalised circulation (current circulation Γ over theoretical initial value Γ_0)
C_l	rolling moment coefficient
$C_{l,\,WV}$	coefficient of rolling moment induced by wake vortex
ΔH	vertical distance
RCR	roll control ratio (see Eq. (1))
X	longitudinal distance in horizontal plane (along intended flight path)
Y	lateral distance in horizontal plane (perpendicular to intended flight path)
δ_l	deflection of roll moment motivator (usually aileron and/or roll spoiler)
σ	standard deviation
a	horizontal extent of hazard area (relative to nearest vortex centre)
b	vertical extent of hazard area (relative to nearest vortex centre)
t^*	normalised wake vortex age
y^*	normalised lateral wake vortex motion (distance travelled over initial wake vortex separation)
Δy	distance in spanwise direction

Abbreviations

a/c	aircraft
ADIRU	Air Data and Inertial Reference Unit
ADS-B	Automatic Dependent Surveillance – Broadcast
AFS	Auto(matic) Flight System
A/P	Autopilot
APA	Aircraft Vortex Spacing System [AVOSS] Predictor Algorithm
ATC	Air Traffic Control
ATM	Air Traffic Management
ATRA	[DLR's] Advanced Technology Research Aircraft
ATTAS	[DLR's] Advanced Technologies Testing Aircraft System
BADA	Base of Aircraft Data [Eurocontrol Aircraft Performance Data Base]
CAT	Clear Air Turbulence
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Centre)
D2P	Deterministic 2-Phase [Wake Vortex Engineering Model]
EGPWS	Enhanced Ground Proximity Warning System
FL	Flight Level
FMA	Flight Mode Annunciator
FMGC	Flight Management and Guidance Computer
FMGS	Flight Management and Guidance System
FMS	Flight Management System
GPWS	Ground Proximity Warning System
G/S	Glide Slope [of the Instrument Landing System]
HMI	Human Machine Interface
ICAO	International Civil Aviation Organisation
IFR	Instrument Flight Rules
kts	knots
LES	Large-Eddy Simulations
LIDAR	Light Detection and Ranging
LOC	Localiser [of the Instrument Landing System]
LoS	Line-of-Sight
NASA	[U.S.] National Aeronautics and Space Administration
ND, N/D	Navigation Display
NM	Nautical Mile
NSE	Navigation System Error
PEst	Parameter Estimation
PFD	Primary Flight Display
P2P	Probabilistic 2-Phase [Wake Vortex Engineering Model]
PVM	Probabilistic Wake Vortex Model
RCR	Roll Control Ratio
RNP	Required Navigation Performance
SESAR	Single European Sky ATM Research
SHAPe	Simplified Hazard Area Prediction

TAWS	Terrain Awareness and Warning System
TCAS	Traffic Alert and Collision Avoidance System
TIS-B	Traffic Information Service – Broadcast
UAV	Unmanned Aerial Vehicle
VFR	Visual Flight Rules
VSI	Vertical Speed Indicator
VSD	Vertical Situation Display
WEAA	Wake Encounter Avoidance and Advisory [System]
WIMS	Weather Information Management System
XTK	cross-track

References

[1] Hahn, K.-U. and Schwarz, C., "Alleviation of atmospheric flow disturbance effects on aircraft response", *ICAS 2008*, Anchorage, 14th – 19th Sep. 2008. proceedings: Edinburgh, 2008.

[2] Hahn, K.-U., Fischenberg, D., Niedermeier, D. and Horn, C., "Wake Encounter Flight Control Assistance Based on Forward-Looking Measurement Processing", *AIAA ASE 2010*, 2nd – 5th Aug. 2010, Toronto, Canada.

[3] Schwarz, C. and Hahn, K.-U., "Automated pilot assistance for wake vortex encounters", *Aerospace Science and Technology*, Vol. 15, 2011, pp. 416-421.

[4] Rouwhorst, W., Groeneweg, J., Winckelmans, G. and Luckner, R., "Piloted Simulation Results of an Onboard Wake Vortex Detection, Warning and Avoidance System and Cockpit Display Aspects", *ICAS 2008*, Anchorage, 14th – 19th Sep. 2008.

[5] Kauertz, S., Reinke, A., "SESAR projects 9.11 & 9.30 'Aircraft Systems for Wake Encounter Alleviation' ", *3rd Major WakeNet3-Europe Workshop*, Southampton, 10th – 11th May 2011.

[6] Holzäpfel, F., "Probabilistic Two-Phase Wake Vortex Decay and Transport Model," *Journal of Aircraft*, Vol. 40, No. 2, 2003, pp. 323-331.

[7] Holzäpfel, F. , "Probabilistic Two-Phase Aircraft Wake-Vortex Model: Further Development and Assessment," *Journal of Aircraft*, Vol. 43, No. 3, 2006, pp. 700-708.

[8] Holzäpfel, F., and Steen, M., "Aircraft Wake-Vortex Evolution in Ground Proximity: Analysis and Parameterization," *AIAA Journal*, Vol. 45, No. 1, 2007, pp. 218- 227.

[9] Holzäpfel, F. et al., "The Wake Vortex Prediction and Monitoring System WSVBS – Part I: Design", in: Holzäpfel, F, Gerz, T., *The DLR Project Wirbelschleppe – Detecting, Characterizing, Controlling, Attenuating, Understanding, and Predicting Aircraft Wake Vortices*, DLR research report FB 2008-15, Köln, 2008., pp. 73-88.

[10] Sarpkaya, T., Robins, R. E., and Delisi, D. P., "Wake-Vortex Eddy-Dissipation Model Predictions Compared with Observations," *Journal of Aircraft*, Vol. 38, No. 4, 2001, pp. 687–692.

[11] Winckelmans G., The Vortex Forecast System (VFS): prediction of trajectories and decay for aircraft wake vortices in an operational environment, Proc. *Conference on Capacity and Wake Vortices*, Imperial College, London, UK, 11th – 14th Sept., 2001.

[12] Schönhals, S., Steen, M. and Hecker, P., "Wake vortex prediction and detection utilising advanced fusion filter technologies", *The Aeronautical Journal*, Vol. 115, No. 1166, 2011, pp. 221-228.

[13] Hahn, K.-U., Schwarz, C. and Friehmelt, H., "A simplified hazard area prediction (SHAPe) model for wake vortex encounter avoidance", *ICAS 2004*, Yokohama, 29th Aug. – 3rd Sep. 2004. proceedings: Edinburgh, 2004.

[14] Schwarz, C. and Hahn, K.-U., "Full-Flight simulator study for wake vortex hazard area investigation", *Aerospace Science and Technology*, Vol. 10 No. 2, 2006, pp. 136-143.

[15] Hahn, K.-U. and Schwarz, C., "Safe limits for wake vortex penetration", *AIAA GNCC 2007*, Hilton Head, 20th – 23rd Aug. 2007. Paper AIAA-2007-6871.

[16] Schwarz, C., Hahn, K.-U., Holzäpfel, F. and Gerz, Th.: "DLR Wake Solutions in Development", *WakeNet USA Meeting October 2010*, Cambridge, MA, 20th -21st Oct. 2010.

[17] ICAO Doc 4444 ATM/501, "*Procedures for Air Navigation Services – Air Traffic Management*", 15th Edition, Montreal, 2007.

[18] ICAO Doc 8168 OPS/611, "*Procedures for Air Navigation Services – Aircraft Operations, Volume II, Construction of Visual and Instrument Flight Procedures*", 5th Edition, Montreal, 2006. *here*: Part III, Sec. I, Chap. 7.4.

[19] ICAO Doc 9613 AN/937, "Performance-based Navigation (PBN) Manual", Volume II, 3rd Edition, Montreal, 2008.

[20] ICAO Annex 2, "*Rules of the Air*", 10th Edition, Montreal, 2005.

[21] Wolkensinger, C., "*Vergleich messtechnischer Konzepte zur bordgestützten Ermittlung atmosphärischer Störphänomene*", DLR internal report IB 111-2010/35, Braunschweig, 2010.

[22] Diekmann, J. H., "*Entwicklung eines Software-moduls zur Generierung von Trajektorien zum kleinräumigen Ausweichen vor Wirbelschlep-pen*", DLR internal report IB 111-2010/59, Braunschweig, 2010.

[23] Gall, Ch., "*Entwicklung und Systembetrachtung eines Assistenzsystems zur Durchführung von Ausweichmanövern*", DLR internal report IB 111-2011/10, Braunschweig, 2010.

[24] Malkin, E.: "*Mexico Crash Inquiry Points to Pilot Error in Turbulence". The New York Times*, New York, 15th Nov. 2008.

[25] AIRBUS Flight Operations Support & Services, "*Getting to Grips with RNP-AR*", Blagnac, 2009.

CEAS 2011

New Configurations

Recent Developments of the FanWing Aircraft

George R Seyfang
Formerly Principal Concepts Engineer ~ BAE Systems, UK

Keywords: *STOL, Wind Tunnel, High Lift, Cross-flow Fan, Cargo*

Abstract

FanWing is an aircraft configuration that uses a simple cross-flow fan mounted in the wing to provide distributed propulsion and augmented wing lift at very low flying speeds. One of the current flight-test vehicles has been modified to accept both a novel OHS twin-tail arrangement and a new wing section and this has shown both increased flight stability and reduced drag, leading to significantly higher cruise speeds.
The good slow flight capability, inherent safety and relatively quiet propulsion of the FanWing rotorcraft could fit well with cargo operations close to urban areas. A comparison with aircraft and helicopters showed that the recent developments of the FanWing concept could now uniquely offer short-field performance close to that of helicopters and tilt-rotor aircraft, but with operating economies close to that of conventional aircraft.

1. Introduction

This paper briefly reviews the history of the FanWing aircraft concept and then reports some of the author's recent aerodynamic activities to improve flight characteristics and to increase flying speeds. Finally it reports an economic study of short-field cargo operations comparing this configuration with established aircraft configurations.

FanWing is a powered-lift aircraft configuration that uses a simple cross-flow fan mounted in the wing to provide distributed propulsion and augmented wing lift at very low flying speeds. The fan's tip speed is considerably lower than the tip speeds of conventional aircraft propellers or helicopter rotors, so it offers unique opportunities for improved propulsive efficiency and reduced noise footprint. The result is that FanWing may offer many of the advantages of both fixed-wing aircraft and helicopters.

2. Background

Although several earlier, and more complex, horizontal-axis rotor concepts [1] had been explored since the early 1900s, it was only in 1998 that the simple FanWing concept was evolved and first flown by Patrick Peebles.



Fig. 1 Early flight model of FanWing

Early wind tunnel tests confirmed the key performance parameter to be the ratio of fan blade tip speed to aircraft speed.....the Tip Speed Ratio, TSR [2, 3]. Maximum lift coefficient values of over 10 have been measured at low speed on the original FanWing section, and this offers remarkably

low take-off speeds and true STOL performance.

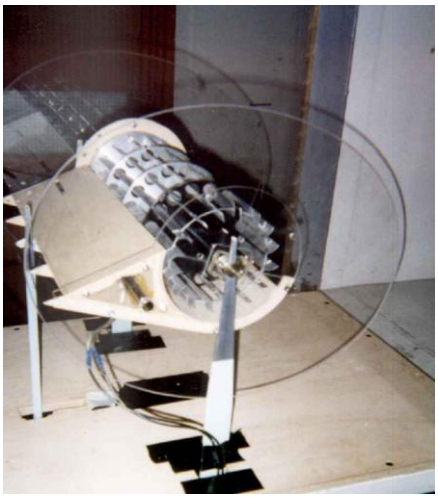


Fig. 2 A 2-D wind tunnel model at ICL

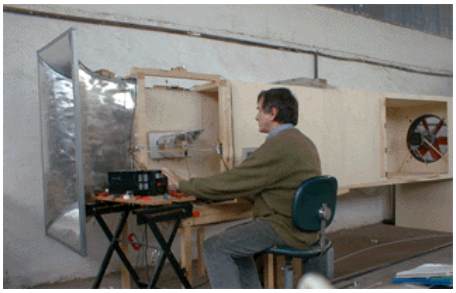


Fig. 3 Early wind tunnel testing

Other models of increasing size and power efficiency have flown and most recently there were daily flight demonstrations of a UAV variant of FanWing during the Farnborough Air Show in 2010.



Fig. 4 FanWing at Farnborough in 2010

Thus far the FanWing concept has demonstrated the ability to fly slowly and safely in the regime of helicopters, and efficiently in the regime of fixed-wing aircraft. This economical loitering capability would be ideal for surveillance UAVs.

In an early attempt to widen the FanWing's flight speed envelope there have been further development flight tests. At the low speed end of the envelope, exploratory tethered flight tests have demonstrated the potential to hover and fly vertically.



Fig. 5 Tethered VTOL hovering tests

3. Recent Aerodynamic Developments

3.1 Objective -- Higher Cruising Speed

Recent developments of FanWing have concentrated on further widening the flight envelope, in particular with three aerodynamic modifications to allow flight at higher speeds. The combination of ultra-short field performance and moderate cruising speeds would greatly widen the field of potential applications of FanWing.

3.2 OHS Tail Configuration

One of the current flight-test vehicles has been modified to accept a novel OHS twin-tail arrangement which moves the tails from the intense fan-stream and wing downwash flow directly behind the wing, to positions where they can exploit the upwash flow from the wingtip vortices.

The OHS tail arrangement was first developed and flown on radio-controlled models over 20 years ago by Professor John Kentfield at Calgary University, [4, 5]. The current author has derived design rules for the optimum use of OHS and estimates of the performance advantages of OHS [6].

Flight testing of FanWing with the OHS twin-tail has shown both increased flight stability and reduced drag, leading to higher flight speeds.
In addition, when the elevators of the OHS tails are used differentially, they can provide roll control as well as pitch control. This allows the wing to be built as a simple lifting surface without ailerons.



Fig. 6 FanWing model with OHS tails



Fig. 7 Twin-tail OHS model 2011

3.3 High-Speed Wing Section Design

The fixed wing section of FanWing has recently been modified following a series of 2-D and 3-D wind tunnel tests of many alternative fan and wing combinations.

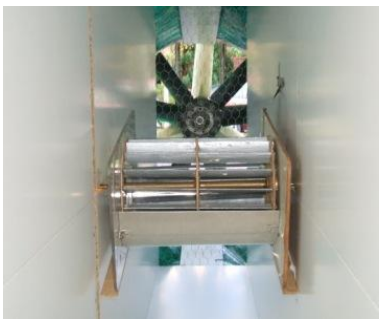


Fig. 8 2-D wind tunnel tests in 2010

Examples are shown below of the wing section with its trailing edge lengthened to reduce profile drag at lower lift coefficients whilst retaining the excellent CLmax values.

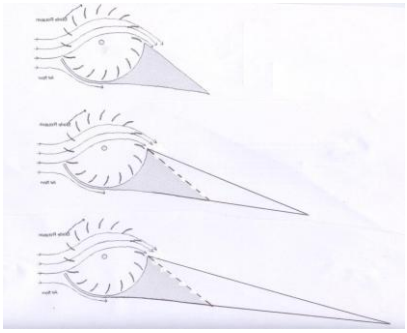


Fig. 9 Baseline (top) and two high-speed sections

Wind tunnel tests of the 2-D model did indeed show that the CLmax values were little changed, but the longer section was found to have less drag and a much improved thrust margin at cruise.

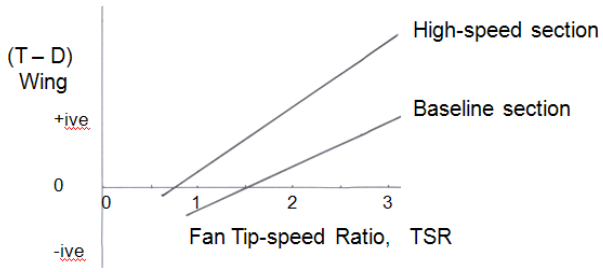


Fig. 10 Cruise (Thrust – Drag) margin

Based on the results of these 2-D model tests and using appropriate 'fan laws', it is estimated that a full-size FanWing aircraft will have a cruise speed near to that of a helicopter.

The opportunity was taken with the 2-D model to measure lift, drag, thrust and power characteristics over a much wider range of TSR values than had been possible previously. This was done to provide data which could be used in take-off and cruise performance calculations, as reported later in section 5.

An example is shown of the CLmax vs TSR and it is seen that very high values were achieved.

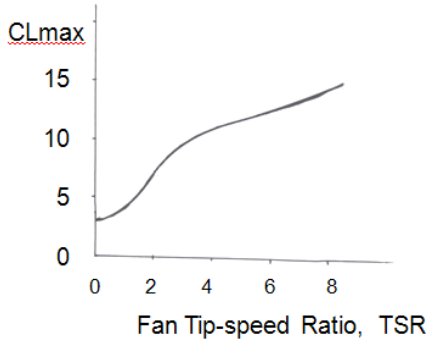


Fig. 11 Maximum lift vs fan tip-speed ratio

In addition to improving the cruise drag characteristics, the long chord section greatly improved the 'gliding' performance in the power-off, auto-rotation mode.

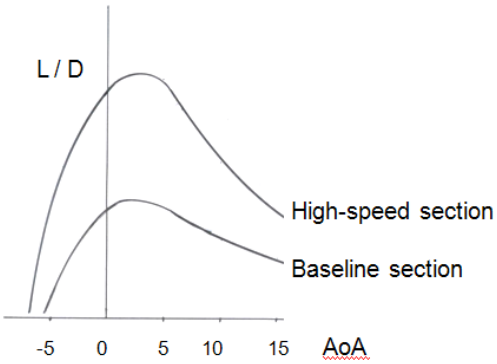


Fig. 12 Glide ratio in auto-rotation mode

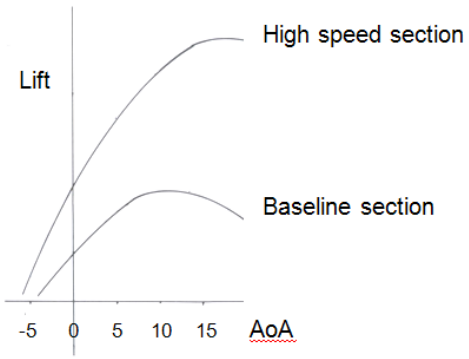


Fig. 13 Lift in auto-rotation mode

Other testing with a short rounded nose extension was found to further improve the performance in the auto-rotation mode.

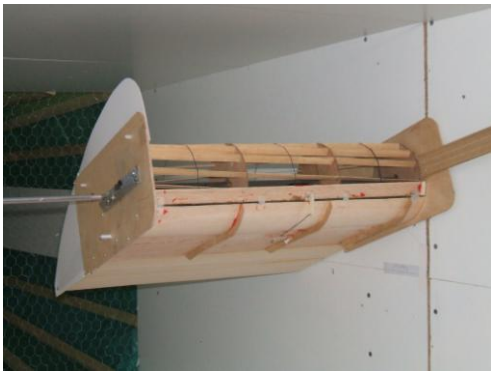


Fig. 14 3-D wind tunnel model 2011

The long chord high-speed section was also tested at 3-D on a half-model, mounted at one side of a larger wind tunnel, and similar improvements performance were obtained.
The 3-D model was also used to examine the horizontal wing-tip extensions fitted to earlier flight models. These were found to add little benefit and have since been removed for simplicity.
Flight testing of the new, longer wing section on a large FanWing model has confirmed the improved flight performance characteristics and has achieved much higher cruise speeds.



Fig. 15 Flight tests of high-speed section 2011



Fig. 16 Example of flight test telemetry

3.4 Fan Blade Design

The fan blade sizes, numbers, cambers and setting angles have been evolved to satisfactory performance levels in an ad hoc way with wind tunnel and flight models. There have also been several partially successful attempts to model the complex fan airflows with CFD. [7, 8, 9].

Recent activities have included testing many fan blade arrangements in a simple water channel at different TSRs.

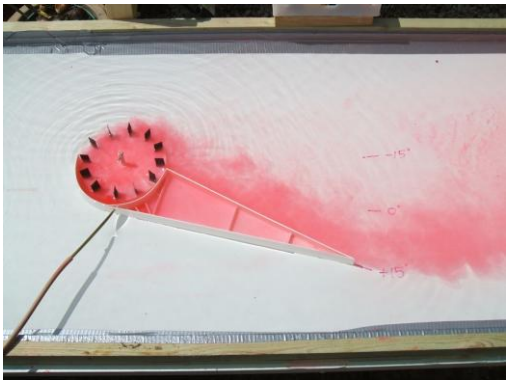


Fig. 17 Water channel tests 2011

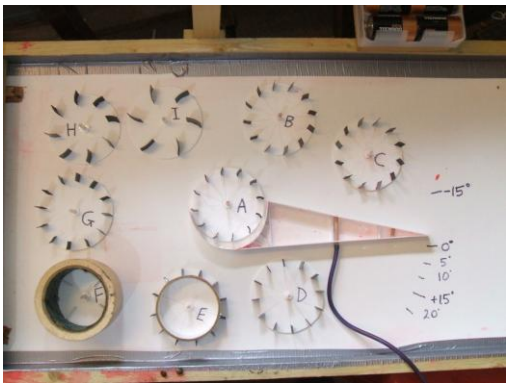


Fig. 18 Examples of the fan-blades tested

Indications from the flow-visualisation tests suggest that the optimum fan blade arrangement is not overly-critical. It is intended to repeat this wide-ranging fan blade study in a wind tunnel to capture the thrust, drag and power requirements at various airspeeds and fan rotation speeds.

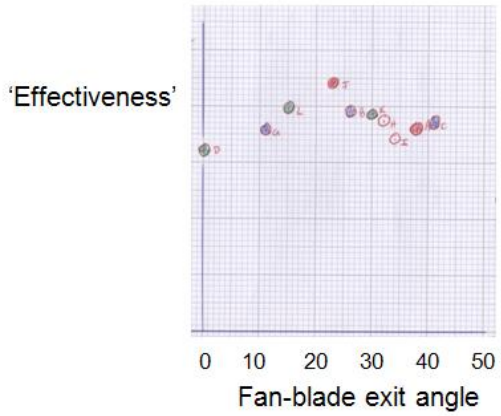


Fig. 19 Fan-blade efficiency measurements

4. Operational Features of FanWing

4.1 Flight Safety

Relative to conventional twin-engine aircraft with propellers, the FanWing concept has three inherent safety features.

Firstly, it does not stall in the way that a normal wing stalls. This is because the wing is blown by the fan and the air does not separate if the rotor is sufficiently powered.

Secondly, it avoids any asymmetric thrust and lift in the event of a single engine failure. FanWing aircraft will have the fans on each wing connected by a cross-shaft so that both fans continue to be powered equally at all power levels. The mounting positions for the engine do not affect this feature and engines may be positioned centrally in the fuselage, at mid-wing span or at the wing tips, depending on operational or structural needs.

Thirdly, even in the event of total power loss then the FanWing is able to auto-rotate and continue to develop lift at low speeds and so provide a safe and controllable glide.

4.2 Flyover Noise

Relative to conventional propeller aircraft and helicopters, the FanWing will cause less fly-over noise nuisance for several reasons. The fan has a larger projected area, it has lightly loaded blades and they are moving at much lower tip speeds. These factors all reduce the noise generated at source. In addition, the fan exhausts above the wing surface and so is partially shielded from observers on the ground.

However, there is an additional noise source on FanWing due to the moving blades passing close to the fixed wing surface. Solutions have included a slight skew to the rotor blades or to the fixed structure to spread the interaction, and minor changes to the separation gap between blades and fixed structure.

5. A STOL Cargo Study

5.1 Opportunities for STOL Operations

Previous economic studies have suggested that there would be a greater demand for air-cargo operations if they could be operated from multiple short airstrips close to industrial areas, and away from crowded regional airports. The demonstrated safe, slow flight capability and relatively quiet propulsion of the FanWing rotorcraft would fit well with such operations close to urban areas.

Other potential applications of a STOL cargo aircraft might include short-range military transportation and civilian operations in countries and terrains which do not have an extensive road, rail or water infrastructure.

This STOL study was carried-out privately with the objective of answering two questions; 'what is the cost of designing for short-field performance and which aircraft technologies are most appropriate?'



Fig. 20 A notional FanWing cargo aircraft

5.2 Types of Aircraft Compared

This study considered four types of aircraft; Conventional STOL, FanWing STOL, Helicopter and Tilt-Rotor. Each type offers different combinations of cruise speed, airfield length and operating cost. The comparisons shown here were done at a design concepts level, with a sizing mission to transport a 5000kg shipping container over a range of 500km.

Conventional STOL aircraft data was taken from an unpublished STOL cargo study. Variants with different field lengths were generated by changes of wing area and engine power, and re-sizing to the design mission.

FanWing STOL aircraft were based on the Conventional aircraft, re-computed for the effects of the FanWing propulsion and lift system, as measured in wind tunnel and flight reported in section 3.

Helicopters have true VTOL capability and data was based on open-source information for the Boeing CH-47 Chinook.

Tilt-Rotor aircraft also have true VTOL capability plus a much higher cruise speed and data was based on open-source information for the Boeing V-22 Osprey.

5.3 Comparison of Aircraft Types

The two key performance and economic parameters chosen for presentation here are the take-off run and D.O.C., the direct operating cost. D.O.C. is a very powerful parameter since it includes flying costs, maintenance costs and capital costs. In turn, these reflect such relevant flight performance features as fuel consumption and cruise speed. The results of the study are shown in Fig. 21 below and it is seen that FanWing variants offer a niche capability between a Conventional STOL aircraft and the two VTOL types having significantly higher costs.

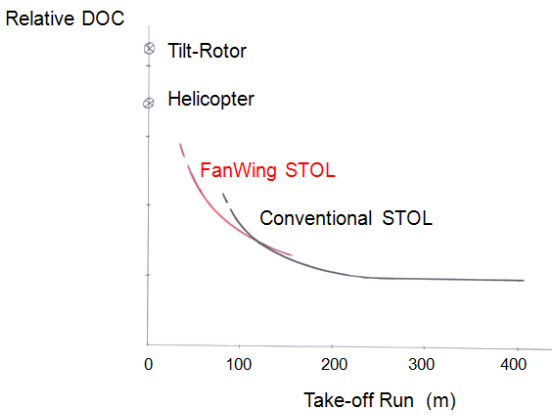


Fig. 21 Results of Cargo Study

6 Conclusion

Recent aerodynamic developments of the FanWing aircraft configuration have increased its cruise speed considerably.
A comparison of four different aircraft and rotorcraft configurations has shown that the FanWing concept now offers an interesting and unique capability, with short-field performance close to that of helicopters and tilt-rotor aircraft, but with operating economies close to that of conventional aircraft.

Acknowledgment

The author has received much help and encouragement with details of the history of FanWing, and the provision of wind tunnel models for testing, from Mr. Patrick Peebles, the inventor of FanWing.

References

[1] Foshag W.F., Boehler G. D., "Review and Preliminary Evaluation of Lifting Horizontal-Axis Rotating-Wing Systems (HARWAS)", Aerophysics Co., Washington DC, 1969

[2] Forshaw S., "Wind-Tunnel Investigation of FanWing" Final Year Project, M.Eng., Imperial College London, 1999.

[3] Kogler K.U., "FANWING, Experimental Evaluation of a Novel Lift & Propulsion Device" M.Eng. Thesis, Imperial College, London 2002

[4] Kentfield J.A.C., "Aircraft configurations with outboard horizontal stabilizers", Journal of Aircraft October 1991

[5] Kentfield J.A.C., "Drag reduction and improved fuel consumption of aircraft employing outboard horizontal stabilizers", AIAA Paper 2005-0811

[6] Seyfang G.R., "Aerodynamics and applications of the OHS configuration", Royal Aeronautical Society Conference on the Aerodynamics of Novel Configurations, London 2008

[7] Duddempudi D., Yao Y., Edmondson D., Yao J., Curley A., "Computational study of flow over generic fan-wing airfoil", Aircraft Engineering and Aerospace Technology, 2007 Vol. 79, Iss: 3, pp.238 – 244

[8] Askari S., Shojaeefard M.H., "Shape optimization of the airfoil comprising a cross flow fan", Aircraft Engineering and Aerospace Technology, 2009 Vol. 81, Iss 5, pp 407 - 415

[9] Chiang A. L., China N., Jie L., Xu-hui, "Research on the Wing Composition of FanWing Flight", ARIC Aeronautical Science and Technology 2009 (3)

[10] Website www.fanwing.com

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

The Conflict of Aerodynamic Efficiency and Static Longitudinal Stability of Box Wing Aircraft

D. Schiktanz, D. Scholz
Aero – Aircraft Design and Systems Group, Hamburg University of Applied Sciences, Hamburg, Germany

Keywords: *box wing, static longitudinal stability, induced drag, aerodynamic efficiency*

Abstract

The induced drag of box wing aircraft is assessed with the help of literature data. The theoretical foundations of static longitudinal stability and controllability are presented and applied to the box wing aircraft. The results are interpreted and put into practice with the help of a medium range box wing aircraft based on the Airbus A320. Stability in cruise is attained by increasing the ratio $C_{L,1}/C_{L,2}$ to a value of 1,74, which is the ratio of lift coefficients of the forward and the aft wing. According to biplane theory this results in a 3,4 % increase of induced drag. Applying aerodynamic theory for closed wing systems no increase would be expected. With the stated ratio of lift coefficients results a relatively small envelope for the center of gravity (CG). Consequently the aircraft is designed to be well balanced with regard to its CG. The individual CGs of the airframe, engines, fuel and payload are all located approximately at the same position. Hence the CG shift is minimized for different payload and fuel quantities.

1 Introduction

As stated in Flightpath 2050 civil aviation transport is facing challenges like globalization, climate change and a cumulative scarcity of resources [1]. To cope with these challenges, aircraft have to become more efficient, especially concerning energy and fuel consumption. Flightpath 2050 states: "In 2050 technologies and procedures available allow a 75% reduction in CO2 emissions per passenger kilometre ... these are relative to the capabilities of typical new aircraft in 2000." Benefits of alternative fuels are seen in addition to this. [1]

It is questionable, if these goals are realistic. Without doubt the latest aircraft emerging on the market have pretty much exhausted the inherent saving potential of conventional configurations without reaching the goal of a 75% CO2 (or fuel) reduction. Therefore it becomes a necessity also to exploit all saving potentials originating from the aircraft configuration. It is not a question anymore if unconventional configurations find acceptance or not. Quite simply, without unconventional configurations stated Flightpath 2050 goals will not be reached. One of these configurations is the box wing aircraft, a biplane with oppositely swept wings whose tips are connected by

extended winglets (see Fig. 1). The most recognized benefits of this configuration are its low induced drag and alleged structural superiority.

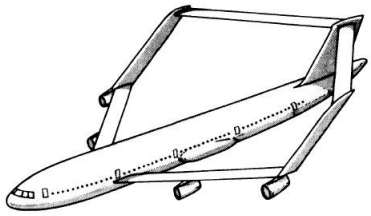


Fig. 1 Example of a box wing aircraft [2]

From literature it can be understood that it is difficult to attain static longitudinal stability and controllability for box wing aircraft. Civil transport aircraft have to be stable in flight [3]. Consequently the design of the aircraft is greatly influenced by the requirements according to static longitudinal stability and controllability. In this paper these requirements are analyzed and possibilities for attaining stability are presented. The easiest to accomplish is having different lift coefficients at both wings which, in the case of equal wing areas, leads to an unequal division of lift between both wings. Applying biplane theory this results in an increase of induced drag [4]. Thus assuring stability is in conflict with aerodynamic efficiency.

In this paper the above approach is applied to the design of a medium range box wing aircraft. Its special design characteristics because of stability requirements are outlined and the decrease of aerodynamic efficiency according to biplane theory is assessed.

2 Aerodynamic Efficiency of Box Wing Aircraft

2.1 Induced Drag

A prominent aerodynamic feature of the box wing configuration is its low induced drag $(D_i)_{box}$. It can be expressed in relation to the induced drag $(D_i)_{ref}$ of a conventional reference wing with the same span. The published equations mostly have the form

$$\frac{(D_i)_{box}}{(D_i)_{ref}} = \frac{k_1 + k_2 \cdot h/b}{k_3 + k_4 \cdot h/b} \qquad (1)$$

where h/b is the height to span ratio, which is the vertical gap between both wings divided by the wing span of the whole configuration. k_1 up to k_4 are free parameters which are adjusted so that the result of Eq. (1) comes close to measured or calculated data. Equation (1) is only applicable when both the box wing and the reference wing configuration have their optimum span loading. In literature there exist several solutions for the parameters k_1 to k_4. A common result is given in [4] reading

$$\frac{(D_i)_{box}}{(D_i)_{ref}} = \frac{1 + 0{,}45 \cdot h/b}{1{,}04 + 2{,}81 \cdot h/b} \qquad (2)$$

which slightly overestimates the actual reduction of induced drag, as confirmed in [5]. For the design of the medium range box wing aircraft a more conservative approach is taken, coming from [6]:

$$\frac{(D_i)_{box}}{(D_i)_{ref}} = \frac{0{,}44 + 0{,}9594 \cdot h/b}{0{,}44 + 2{,}219 \cdot h/b} \qquad (3)$$

Figure 2 shows a graphic comparison of the Eqs. (2) and (3).

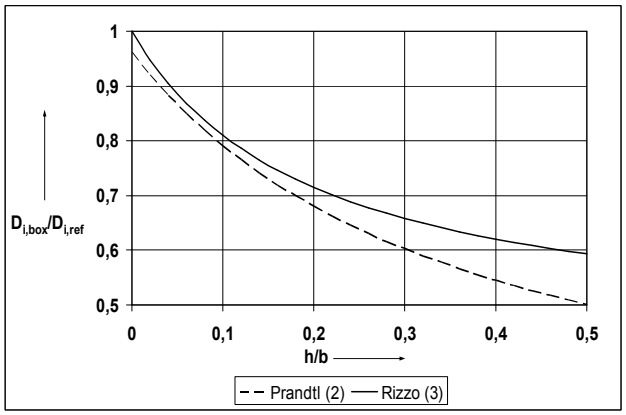


Fig. 2 Reduction of induced drag acc. to Eqs. (2) and (3)

2.2 Conditions for Minimum Induced Drag

For a conventional cantilever wing the lift distribution for minimum induced drag is elliptical. The conditions concerning the box wing configuration are more complex, since the span wise distribution and also the division of lift between both wings have to be considered.

2.2.1 Span Wise Lift Distribution

In [7] it is indicated that the lift distribution for minimum induced drag is a combination of an elliptical and a constant part for the horizontal wings and a linear and butterfly shaped part for the vertical winglets, as it is shown in Fig. 3.

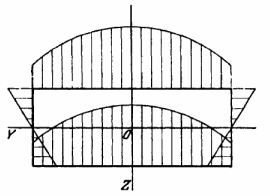


Fig. 3 Lift distribution of a box wing configuration [7]

As depicted in [8] the ratio of the constant to the elliptical part increases with higher h/b ratios. The investigations in [8] were performed for cantilever wing-winglet configurations, but the qualitative results should be transferable to box wing configurations.

2.2.2 Lift Division between Both Wings

According to biplane theory both wings have to generate the same amount of lift so that minimum induced drag is achieved [4]. With the help of Eq. (4) presented in [4] the induced drag of an arbitrary biplane can be calculated. It reads

$$(D_\mathrm{i})_\mathrm{bi} = \frac{1}{\pi q}\left(\frac{{L_1}^2}{{b_1}^2} + 2\sigma\frac{L_1 L_2}{b_1 b_2} + \frac{{L_2}^2}{{b_2}^2}\right) \qquad (4)$$

where q is the dynamic pressure, L_i the lift of the individual wings, b_i the span of the individual wings and σ is the wing interference factor which depends on the h/b ratio. In [4] the equation for calculating the minimum induced drag of a biplane is derived as well, being

$$(D_\mathrm{i})_{\mathrm{bi,min}} = \frac{L^2}{\pi q b^2}\cdot\frac{1+\sigma}{2} \qquad (5)$$

where L is the total lift and b the span of both wings. Equation (5) is only applicable if both wings generate $L/2$ and have equal wing spans.

Now the relation of the actual drag of a biplane with equal wing spans to the minimum induced drag can be formulated. It is assumed that the result is applicable to box wing configurations as well. Dividing Eq. (4) by Eq. (5) and introducing the lift ratio

$$x = \frac{L_1}{L_2} \qquad (6)$$

yields

$$\frac{D_\mathrm{i}}{D_{\mathrm{i,min}}} = 2\frac{x^2 + 2\sigma x + 1}{(x+1)^2(\sigma+1)}. \qquad (7)$$

Using

$$\sigma = 2\frac{(D_\mathrm{i})_\mathrm{box}}{(D_\mathrm{i})_\mathrm{ref}} - 1 \qquad (8)$$

which is concluded from [5] and taking Eq. (3) it is possible to plot the ratio of the actual to the minimum induced drag depending on the h/b ratio and the lift ratio (Fig. 4).

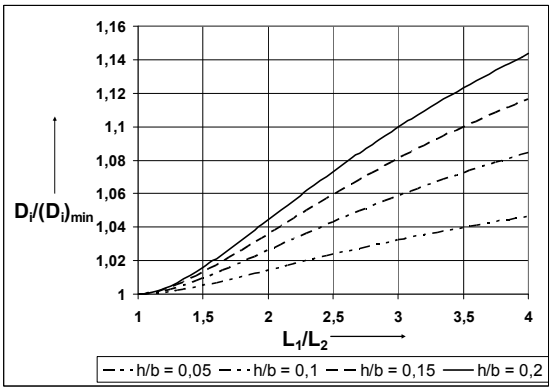


Fig. 4 Induced drag penalties according to biplane theory

The derivation of Eq. (7) can be found in [9]. In section 4.2 it will be shown that the lift ratio has a value about two. For *h*/*b* ratios of about 0,2, which is a reasonable value for actual aircraft, a 4 % increase of induced drag would occur.

The induced drag penalty results when applying biplane theory. But there also exists a different approach which takes account of the characteristics of a closed wing system. Amongst others it is presented in [10] and [11] and states that for a closed wing system minimum induced drag can also be achieved when both horizontal wings generate a different amount of lift. This is possible by adding a constant circulation loop to the distribution of circulation of the whole wing configuration. An example of the resulting lift distribution is depicted in Fig. 5. Adding a constant circulation loop does not change the elliptical parts, but the constant part of one horizontal wing is reduced while that of the other wing is increased by the same amount. Consequently the winglet loading changes as well.

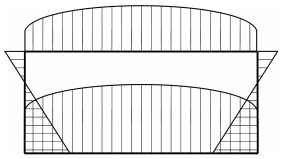


Fig. 5 Lift distribution with unequal lift of both wings

[2] and [5] apply the conditions according to biplane theory to the design of a box wing aircraft. This approach was taken for the medium range box wing aircraft presented in section 4 of this paper. The applicability of the proposals in [10] and [11] is part of ongoing studies.

3 Static Longitudinal Stability and Controllability of Box Wing Aircraft

3.1 Basics

For assessing static longitudinal stability and controllability of a box wing aircraft its CG envelope is taken as reference. It is the permissible region of the aircraft's center of gravity. The forward limit is defined by controllability requirements (control limit). The aft limit is the neutral point, which is defined by stability requirements (stability limit). For a stable and controllable aircraft the control limit has to be situated in front of the stability limit.

The condition for a controllable aircraft is a positive pitching moment about the center of gravity (M_{CG}) which counters the negative pitching moment generated by the wings. With the help of the pitching moment coefficient $C_{M,CG}$ this condition reads

$$C_{M,\mathrm{CG}} > 0\,. \qquad (9)$$

There are two conditions for a stable aircraft. The stability condition requires that the slope of the pitching moment coefficient with regard to the lift coefficient C_L is negative [12]:

$$\frac{\mathrm{d}\,C_{M,\mathrm{CG}}}{\mathrm{d}\,C_L} < 0\,. \qquad (10)$$

The trim condition requires that the pitching moment coefficient at zero lift is positive [12]:

$$\left(C_{M,\mathrm{CG}}\right)_{C_L=0} > 0\,. \qquad (11)$$

The conditions given with the Ineqs. (9), (10) and (11) are applied to the equilibrium of moments of the aircraft in horizontal and non-accelerated flight neglecting control surfaces or a horizontal stabilizer. This results in a formulation of the stability limit and the control limit of the CG.

3.2 Equilibrium of Moments in Horizontal and Non-Accelerated Flight Neglecting Control Surfaces

The formation of the equilibrium of moments is based on the scheme shown in Fig. 6.

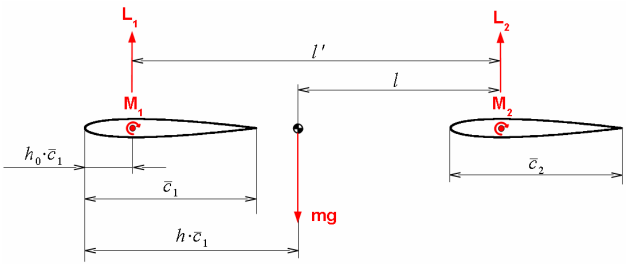


Fig. 6 Forces and moments acting on a box wing aircraft

The lift forces L_1 and L_2, the weight mg and the pitching moment of the wings about their aerodynamic centers M_1 and M_2 are considered. Drag and thrust are neglected, so that the vertical distance between both wings is irrelevant. The CG position h is expressed as multiple of the length of the mean aerodynamic chord (MAC) of the forward wing $\bar{c}_1$, measured from the leading edge of MAC_1. The position of the aerodynamic center of the forward wing h_0 is also expressed as multiple of $\bar{c}_1$, measured from the leading edge of MAC_1. For simplification h_0 is assumed to be 0,25. The distance between the aerodynamic centers of both wings is given with l', the distance between the CG and the aerodynamic center of the aft wing with l. $\bar{c}_2$ is the length of MAC_2. The aerodynamic centers are assumed to be at 25 % of the respective MAC.

The following equations only show results of the formulation of the equilibrium of moments. A detailed derivation is presented in [9]. According to Fig. 6 the equilibrium of moments reads

$$M_{CG} = L_1(h-h_0)\bar{c}_1 - L_2 l + M_1 + M_2 = 0. \qquad (12)$$

Introducing the total lift L and building the coefficient form of Eq. (12) results in

$$\begin{aligned} C_{M,CG} &= C_L(h-h_0)\bar{c}_1' - C_{L,2}\bar{V}' + \\ &+ C_{M,1}\bar{c}_1' s_1 + C_{M,2}\bar{c}_2' s_2 = 0 \end{aligned} \qquad (13)$$

with $\bar{c}_i'$ as the relative length of the respective MAC given by

$$\bar{c}_i' = \frac{\bar{c}_i}{\bar{c}}; \quad i = 1;2 \qquad (14)$$

where the length of the total MAC $\bar{c}$ is defined with

$$\bar{c} = \bar{c}_1 s_1 + \bar{c}_2 s_2 \qquad (15)$$

according to [9]. s_1 and s_2 are the relative reference areas of the individual wings:

$$s_i = \frac{S_i}{S_1 + S_2}; \quad i = 1;2. \qquad (16)$$

$\bar{V}'$ is the modified volume coefficient of the aft wing. It is defined with

$$\bar{V}' = \frac{l'}{\bar{c}} s_2. \qquad (17)$$

3.3 CG Envelope

Applying the conditions given by the Ineqs. (9), (10) and (11) to Eq. (13) results in a formulation of the control and the stability limit of the CG position. The condition for controllability reads

$$h > h_0 + \frac{C_{L,2}}{C_L} \cdot \frac{\bar{V}'}{\bar{c}_1'} + \frac{C_{M,1}}{C_L} s_1 + \frac{C_{M,2}}{C_L} s_2 \frac{\bar{c}_2}{\bar{c}_1}. \qquad (18)$$

Here all parameters are either defined by aircraft geometry or the flight state. The condition for stability is

$$h < h_0 + \frac{dC_{L,2}}{dC_L} \cdot \frac{\bar{V}'}{\bar{c}_1'} \qquad (19)$$

where the determination of the gradient $dC_{L,2}/dC_L$ requires special attention (see section

3.4). The other parameters are defined by aircraft geometry. The trim condition reads

$$C_{M,1}c_1's_1 + C_{M,2}c_2's_2 - \left(C_{L,2}\right)_{C_L=0} \cdot \overline{V}' > 0. \quad (20)$$

The lift coefficient of the aft wing at zero total lift $(C_{L,2})_{C,L=0}$ can be determined with the help of the gradient $dC_{L,2}/dC_L$. The appropriate equation is

$$\left(C_{L,2}\right)_{C_L=0} = C_{L,2} - \frac{dC_{L,2}}{dC_L} \cdot C_L, \quad (21)$$

assuming that the gradient $dC_{L,2}/dC_L$ is constant. All other parameters in Ineq. (20) are geometric or aerodynamic constants. The pitching moment $C_{M,i}$ of all airfoils is assumed to be -0,1 which is sufficient for the preliminary design of the medium range box wing aircraft.

3.4 Determination of $dC_{L,2}/dC_L$

[12] shows that it is possible to calculate the gradient $dC_{L,2}/dC_L$ with the help of the following expression:

$$\frac{dC_{L,2}}{dC_L} = \frac{a_2}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) \quad (22)$$

where a_2 is the lift curve slope of the isolated aft wing and a the lift curve slope of the whole aircraft. ε is the downwash angle and α the angle of attack with regard to the zero lift line of the aircraft. a_2 is estimated with the help of Eq. (23) given in [13].

$$\frac{dC_L}{d\alpha} = \frac{2\pi \cdot A}{2 + \sqrt{A^2\left(1 + \tan^2\varphi_{50} - M^2\right) + 4}} \cdot \quad (23)$$

A is the wing aspect ratio, φ_{50} the sweep angle at half chord and M the Mach number. The total lift curve slope a is determined according to section 4.5.1.1 in [14]. There the influence on the aft wing because of the vorticity generated by the front wing is taken account of. Interferences between the fuselage and the forward wing are considered as well. The average downwash gradient $d\varepsilon/d\alpha$ is estimated according to method two in section 4.1.1 of [14]. The related equations for calculating the lift curve slope of the whole aircraft as well as the downwash gradient are extensive and therefore not shown in this paper. The interested reader can find the relations in [9].

The evaluation of Eq. (22) is done with the help of a spreadsheet which is further referred to in section 3.5.

3.5 Evaluation

The conditions derived in section 3.3 are evaluated with the help of a spreadsheet. Input parameters are the geometry of the wing configuration and the aerodynamic parameters for cruise flight. Like this it is possible to assess different geometric and aerodynamic configurations. Flight phases other than cruise have yet to be investigated.

It is desired to have the greatest possible CG envelope, which means that the stability limit should be situated as far aft as possible while the control limit should be situated as far forward as possible.

3.5.1 Stability Limit

The stability condition [Ineq. (19)] depends on the gradient $dC_{L,2}/dC_L$ and on the modified volume coefficient $\overline{V}'$. A high value for $dC_{L,2}/dC_L$ is beneficial for static longitudinal stability. Under practical conditions $dC_{L,2}/dC_L = 1$ can already be considered to be a very satisfying value. $dC_{L,2}/dC_L$ mostly depends on the sweep of both wings. The sweep of the forward wing needs to be high. For example, reducing the sweep angle from 35° to 0° results in a reduction of about 10 % for $dC_{L,2}/dC_L$. The aft wing is swept forward. Here a lower amount of sweep is favorable. For example, a reduction of sweep from -25° to 0° yields an increase of about 6 % for $dC_{L,2}/dC_L$. These values apply for the medium range box wing aircraft presented in section 4.

The stability limit highly depends on the modified volume coefficient $\overline{V}'$. A high value improves stability. $\overline{V}'$ increases with increasing l' and hence when the wings are placed far apart from each other.

Looking at the trim condition [Ineq. (20)] it becomes obvious that the important parameters are the lift coefficient of the aft wing at zero total lift $(C_{L,2})_{C,L=0}$ and the modified volume coefficient $\overline{V}'$ which has a positive value. For attaining static longitudinal stability the aft wing has a lower lift coefficient than the forward wing, hence $(C_{L,2})_{C,L=0} < 0$ (see Fig. 7). This results in the left hand side of Ineq. (20) being positive, even if pitching moment coefficients are negative.

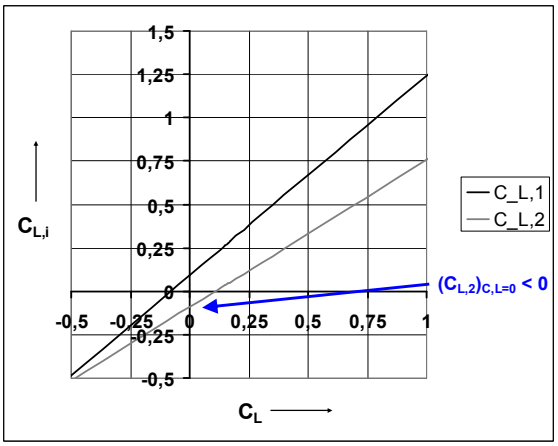


Fig. 7 Lift coefficients of the individual wings vs. total lift coefficient of the box wing aircraft

Manipulating the pitching moment coefficient of the wings (or possibly the fuselage) is not considered at this stage, but will be discussed in section 3.6. According to Eq. (21) $(C_{L,2})_{C,L=0}$ depends on the lift coefficient of the aft wing at the investigated flight state as well as the total lift coefficient at this flight state and the gradient $dC_{L,2}/dC_L$. A low value for the lift coefficient of the aft wing $C_{L,2}$ helps to comply with the trim condition. For a given total lift coefficient this means that the ratio $C_{L,1}/C_{L,2}$ needs to be high.

For conventional tail aft configurations the positive zero lift pitching moment is provided by the horizontal stabilizer which produces a downward force and thus lets the aircraft pitch up.

3.5.2 Control Limit

For having a stable and controllable aircraft the control limit has to be situated in front of the stability limit. The more forward the control limit, the larger the CG envelope. Thus the sum of the right hand side of Ineq. (18) has to have the smallest possible value. In the beginning it is assumed that both wings have equal mean aerodynamic chord lengths and wing areas. Thus $\overline{V}'$ can only be changed with the help of the longitudinal distance of both wings l'. This parameter is however influenced by the sweep of both wings, their integration with regard to the fuselage and the tail as well as a feasible winglet sweep. Consequently changing l' requires a complete redesign of the aircraft. So the only parameter to be freely manipulated in Ineq. (18) is the lift coefficient of the aft wing. All of the other parameters are defined by the flight state. Consequently the ratio $C_{L,2}/C_L$ needs to be low (or $C_{L,1}/C_{L,2}$ needs to be high), as it was already desired at the end of section 3.5.1.

3.5.3 CG Envelope Diagram

It is possible to plot the CG envelope depending on the modified volume coefficient of the aft wing. The relating diagrams are shown in Figs. 8, 9 and 10 for different values of $C_{L,2}/C_{L,1}$. They apply for the medium range box wing aircraft presented in section 4 under cruise conditions with maximum take off weight and maximum payload. The CG envelope is the region forward of the stability limit and aft of the control limit. In Fig. 8 the ratio $C_{L,1}/C_{L,2}$ is unity. It can be seen that the box wing aircraft is unstable because the control limit is aft of the stability limit. Having a value lower than one (Fig. 9) moves the control limit further aft which deteriorates stability. The diagram for a properly chosen ratio of $C_{L,1}/C_{L,2}$ is depicted in Fig. 10. The relating CG envelope is indicated with the green line and is about 48 % MAC, which equals 24 % MAC for a conventional reference aircraft. The CG of the aircraft is within the allowable limits taking account of a static margin of about 5 % MAC in terms of a conventional wing. The static margin is indicated with the yellow line. In Fig. 10 the CG envelope would increase with higher values for

$\overline{V}'$. Hence the longitudinal distance between both wings should be increased for a larger CG envelope.

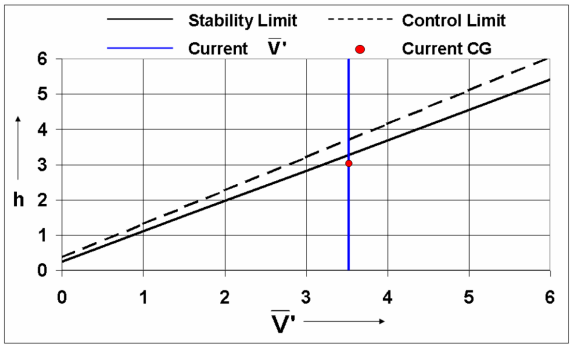


Fig. 8 CG envelope diagram for an unstable box wing aircraft ($C_{L,1}/C_{L,2} = 1$)

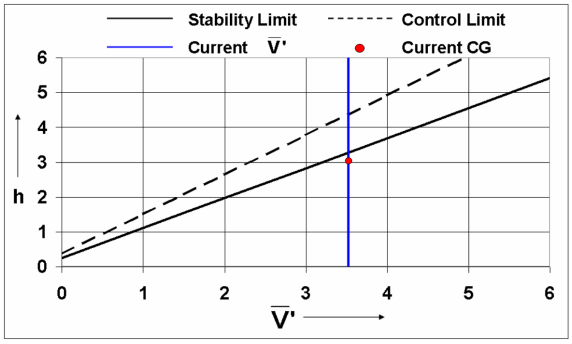


Fig. 9 CG envelope diagram for an unstable box wing aircraft ($C_{L,1}/C_{L,2} < 1$)

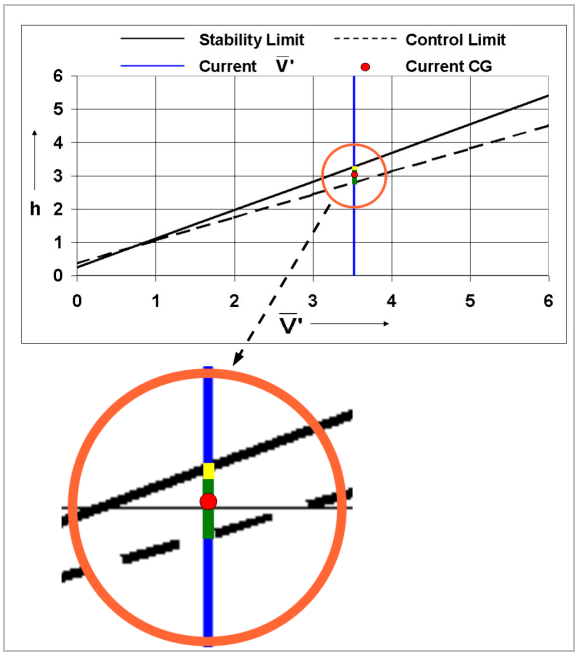


Fig. 10 CG envelope diagram for a stable box wing aircraft ($C_{L,1}/C_{L,2} > 1$)

For conventional tail aft aircraft the CG envelope diagram is used in order to determine the required $\overline{V}'$ for a desired CG envelope. Like this it is possible to size the horizontal stabilizer independently from the other aircraft components. This approach is not possible for a box wing aircraft. Here the value for l' and the resulting $\overline{V}'$ is the result of the designed configuration. Based on the design the CG envelope diagram is drawn and it is checked whether the aircraft's CG is within the allowable limits. If this is not the case, the aircraft has to be completely redesigned because the design of the wings, the fuselage, the tail and the resulting CG position all depend on each other. For conventional tail aft configurations it is possible to move the wing independently if it does not suit the CG position.

3.6 Alternative Ways of Attaining Stability

In section 3.5 the option of having unequal lift coefficients is favored for attaining static longitudinal stability. However, there also exist alternative options which have not been mentioned yet or which were intentionally excluded in order to explain the main idea in simple terms first.

It is possible to combine wing twist and sweep so that the wing has a positive pitching moment. A part of the wing configuration in front of the CG has to produce a high amount of lift (e.g. the center region of the forward wing) while a part aft of the CG produces less lift or even a downward force (e.g. center region of the aft wing). This principle is also applied to flying wings where the wing tips mostly generate a downward force. This approach is not considered for the design of the medium range box wing aircraft because it causes the span wise lift distribution to deviate from the optimum.

Another option is an adaption of the airfoils. With the help of reflexed airfoils whose rear sections are cambered upwards it is possible to have a positive pitching moment without applying twist and sweep as described above. However, it is unlikely that such airfoils are applicable to a transonic transport aircraft. In addition these airfoils will not provide a high

glide ratio. Designing the fuselage with a reflexed rear section is also possible, as it is shown in [15], but the question of applicability arises again.

Another possibility is a control surface at the tail of the aircraft. In [16] a surface at the rear end of the fuselage is proposed. This option seems to be promising. A new formulation of the equilibrium of moments would be necessary to evaluate the effects regarding stability and controllability. The inner part of the aft wing could be equipped with upward deflecting elevators. However, this is nothing else as airfoils whose rear sections are cambered upwards. Substantial additional drag would be the consequence. An additional horizontal stabilizer independent from the aft wing could be a possibility or using a V-tail with symmetrically upwards deflected elevators (see Fig. 11). Further investigations would need to show if it is better to accept resulting trim drag than to accept a higher than minimum induced drag due to a lift distribution where $L_1/L_2 > 1$ according to Fig. 4.

4 Application to a Medium Range Box Wing Aircraft

4.1 Design Approach

The design of the medium range box wing aircraft is based on a reference aircraft comparable to the Airbus A320. Both aircraft have the same design mission. Like this it is possible to compare the characteristics of both aircraft in order to assess the potential of the box wing configuration. The design study is presented in [9] and [17]. For having a valuable comparison both aircraft are supposed to have the same total wing area and the same wing span. Both wings of the box wing aircraft have the same reference area. Because most of the investigations are preliminary the found design still leaves room for optimization.

4.2 Implementation of Stability Requirements

As discussed static longitudinal stability and controllability is attained with the help of unequal lift coefficients of both wings. The ratio $C_{L,1}/C_{L,2}$ is about 1,74 which was determined for cruise conditions. If an average lift coefficient of about 0,75 is assumed for cruise, the lift coefficient of the forward wing would be 0,96 and that of the aft wing 0,55. The resulting CG envelope is about 48 % MAC, which is equivalent to an envelope of 24 % MAC for a conventional reference wing. This magnitude is only applicable for aircraft being well balanced with regard to their CG position. Aircraft with a different weight distribution require a larger CG envelope. For the box wing aircraft this would mean a further increase of the ratio $C_{L,1}/C_{L,2}$, which is not feasible taking account of the cruise lift coefficients of the individual wings. Consequently the box wing aircraft needs to be well balanced as well. This requires the individual CGs of the airframe, engines, fuel and payload all to be located approximately at the same position. Hence the CG shift is minimized for different payload and fuel quantities. The most forward position of the wing is limited by the required clearance to the front door. A forward swept vertical tail brings the wing in a symmetric position with respect to the fuselage. A compact fuselage also minimizes the CG shift for different loading conditions. Since the CG is close to the center of the cabin the CG shift proceeds almost symmetrically during loading. The resulting cabin layout includes seat rows with eight seats abreast and two main aisles. The increased cross sectional area allows for an accommodation of standard LD3 containers. The fuselage is formed as a drop which is an efficient aerodynamic shape. Detailed investigations of the optimum fuselage fineness ratio (fuselage length to diameter ratio) can be found in [18]. Fig. 12 of [18] shows a minimum drag of the fuselage relative to cabin surface for a fineness ratio of about 7.5. With a fineness ratio of 5,8 the proposed box wing aircraft is not too far from this optimum. The fuselage is lighter than the reference fuselage because of its higher

diameter and decreased length. The resulting aircraft layout is shown in Fig. 11.

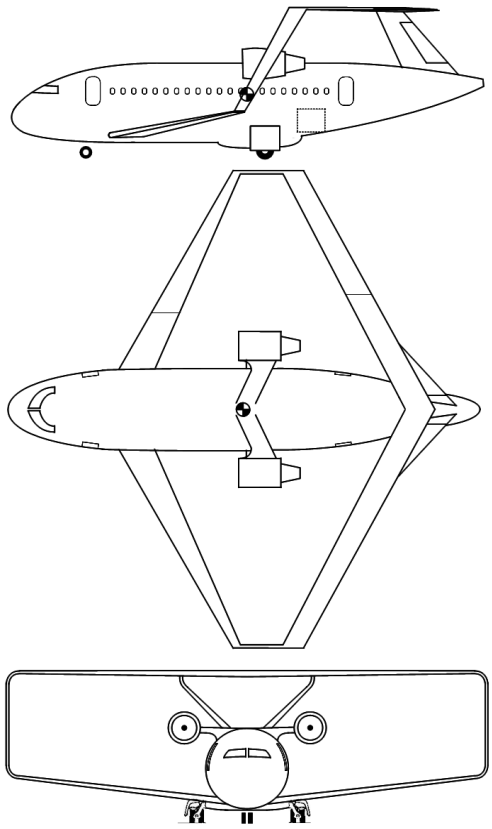


Fig. 11 Three view drawing of the medium range box wing aircraft

Another design feature of the aircraft is the V-tail which is also proposed in [5]. With its help the aft wing is supported by two surfaces what increases its stability. Mainly the V-tail is supposed to function as vertical stabilizer, but its angular surfaces also provide the possibilities of a horizontal stabilizer. It could help to achieve the positive zero lift pitching moment which is needed to comply with the trim condition [Ineq. (20)]. However, this would require a new analysis.

4.3 Decrease of Aerodynamic Efficiency

With the help of the ratio $C_{L,1}/C_{L,2}$ it is possible to determine the increase of induced drag according to biplane theory [Eq. (7)]. Since both wings of the box wing aircraft have the same reference area and it is assumed that they are exposed to the same dynamic pressure, the ratio of lift coefficients is the same as the ratio of total lift L_1/L_2. Hence it is possible to apply Eq. (7). With the given ratio of 1,74 (section 4.2) $D_i/D_{i,min} = 1{,}034$ is the result, so there is a 3,4 % increase of induced drag because of the unequal lift of both wings. According to Eq. (24) below this is equal to a 3,4 % decrease of the span efficiency factor, provided that all of the other parameters in Eq. (24) are constants.

$$D_i = \frac{L^2}{q \cdot \pi \cdot b^2 \cdot e}. \qquad (24)$$

Applying Eq. (25) for the determination of the maximum glide ratio E_{max} this results in a 1,7 % decrease of the maximum glide ratio.

$$E_{max} = \frac{1}{2}\sqrt{\frac{\pi \cdot A \cdot e}{C_{D,0}}}. \qquad (25)$$

$C_{D,0}$ is the zero lift drag coefficient of the aircraft. For the designed medium range box wing aircraft a drop from 20,7 to 20,4 results for E_{max} because of the unequal lift of both wings. Because of that the aircraft consumes about 110 kg more fuel for the reference mission than with minimum induced drag.

It needs to be emphasized that these numbers result when applying biplane theory, as it is also done in [2] and [5]. But from [10] and [11] it can be understood that biplane theory may not be applicable to box wings in terms of induced drag calculation because the theory does not take account of the special characteristics of closed wing systems. According to [10] and [11] there should be no increase of induced drag for ratios of L_1/L_2 being higher than one. So it seems that biplane theory gives an upper limit of the increase of induced drag. However, the efficiency loss according to biplane theory is small. Hence it is unlikely to have a significantly decreased induced drag with the options presented in section 3.6.

The determination of the induced drag of a box wing configuration with the help of simple equations may not produce results accurate enough. For a more detailed assessment of the induced drag of the final configuration it is

necessary to use computational fluid dynamics, preferably vortex lattice methods.

5 Conclusion

In this paper the characteristics concerning static longitudinal stability and controllability of box wing aircraft and the resulting effects regarding aerodynamic efficiency were investigated. It was shown that it is possible to determine the induced drag of an optimally loaded box wing aircraft with the help of literature data. According to biplane theory one condition for optimum loading is that both wings generate the same amount of lift. From the assessment of static longitudinal stability it was reasoned that this condition is most likely violated.

For assessing static longitudinal stability and controllability the theoretical foundations were presented. These were applied to the box wing configuration which resulted in a formulation of the allowable CG envelope, based on the negligence of thrust and drag forces as well as the assumption of horizontal and non-accelerated flight without an additional horizontal stabilizer or control surface deflection. It was discovered that having equal lift coefficients at both wings leads to an unstable aircraft provided that the zero lift pitching moment of the wings is negative. For a stable aircraft the ratio $C_{L,1}/C_{L,2}$ has to reach a certain value higher than one. The higher the value, the larger is the CG envelope of the aircraft. Depending on the flight state the lift coefficient of the forward wing could become problematic. It is also possible to attain stability without increasing the ratio $C_{L,1}/C_{L,2}$. In this case the combined zero lift pitching moment of the wings and the fuselage needs to be positive. A compromise between a manipulation of the wing/fuselage zero lift pitching moment and a lower ratio $C_{L,1}/C_{L,2}$ is also possible. Adapting the wing geometry for attaining stability might be in conflict with other design requirements.

The application of the found results was demonstrated with the help of a medium range box wing aircraft based on the Airbus A320. The wings were supposed to have a value of -0,1 for the zero lift pitching moment. The ratio $C_{L,1}/C_{L,2}$ was determined to be 1,74 for cruise which provides a CG envelope of 48 % MAC. This is equal to an envelope of 24 % MAC for a comparable and conventional reference aircraft. A larger envelope would require a higher value for $C_{L,1}/C_{L,2}$ which does not seem to be feasible. Therefore the layout of the aircraft is well balanced regarding the CG position. The empennage was chosen to be a V-tail which could act as additional horizontal stabilizer. In this way the CG envelope could be increased for a given $C_{L,1}/C_{L,2}$ ratio or the $C_{L,1}/C_{L,2}$ ratio could be decreased for a given CG envelope. A profound assessment of this possibility has yet to be made.

The decrease of the span efficiency of the medium range box wing aircraft because of stability requirements was determined according to biplane theory. Its span efficiency factor is decreased by 3,4 % which results in a 1,7 % decrease of the maximum glide ratio. This means an increase in fuel burn of about 1 % for the reference mission, which is about 110 kg. However, judging from the aerodynamics of closed wing systems there may not even be an increase in induced drag and hence in fuel consumption when the box wing aircraft is stabilized with a ratio $C_{L,1}/C_{L,2} > 1$.

Acknowledgement

The project underlying this report was funded by the German Federal Ministry for Education and Research (support code (FKZ) 03CL01G). The authors are responsible for the content of this publication.

References

[1] European Commission, *Flightpath 2050 : Europe's Vision for Aviation*, Publications Office of the European Union, Luxembourg, 2011.

[2] Lange R.H., et al., "Feasibility Study of the Transonic Biplane Concept for Transport Aircraft Application", The Lockheed-Georgia Company, 1974.

[3] European Aviation Safety Agency, *Certification Specifications for Large Aeroplanes : CS-25*, Amendment 10, EASA, Cologne, 2010, CS 25.173.

[4] Prandtl L., "Induced Drag of Multiplanes," National Advisory Committee for Aeronautics, 1924.
[5] Frediani A., "The Prandtl Wing", *VKI Lecture Series* : *Innovative Configurations and Advanced Concepts for Future Civil Transport Aircraft*, Rhode St-Genèse, 2005.
[6] Rizzo E., *Optimization Methods Applied to the Preliminary Design of Innovative, Non Conventional Aircraft Configurations*, Edizioni ETS, Pisa, 2007, Chap. 4.
[7] Durand W. (Ed.), von Kármán T., Burgers J.M., *Aerodynamic Theory Vol. 2* : *General Aerodynamic Theory – Perfect Fluids*, Julius Springer, Berlin, 1935.
[8] Mangler W., "The Lift Distribution of Wings With End Plates", National Advisory Committee for Aeronautics, 1938.
[9] Schiktanz D., "Conceptual Design of a Medium Range Box Wing Aircraft", Master Thesis, Department of Automotive and Aeronautical Engineering, Hamburg University of Applied Sciences, Germany, 2011.
[10] Kroo I., "Nonplanar Wing Concepts for Increased Aircraft Efficiency", *VKI Lecture Series* : *Innovative Configurations and Advanced Concepts for Future Civil Transport Aircraft*, Rhode St-Genèse, 2005.
[11] Demasi L., "Investigation on Conditions of Minimum Induced Drag of Closed Wing Systems and C-Wings", *Journal of Aircraft*, Vol. 44, No. 1, 2007, pp. 81-99.
[12] Young T., "Flight Mechanics", Lecture Notes, Department of Mechanical and Aeronautical Engineering, University of Limerick, Ireland, 2001.
[13] Scholz, D., "Skript zur Vorlesung Flugzeugentwurf", Lecture Notes, Department of Automotive and Aeronautical Engineering, Hamburg University of Applied Sciences, Germany, 1999.
[14] Finck R.D., "USAF Stability and Control Datcom", McDonnell Douglas Corporation, 1978.
[15] Frediani A., et al., "Development of ULM Prandtlplane Aircraft and Flight Tests on Scaled Models", *XIX Congresso Nazionale AIDAA*, Rome, 2007.
[16] Frediani A., "Swept-Wing Box-Type Aircraft with High Flight Static Stability", International Patent, WO 2004/074093 A1, 2004.
[17] Schiktanz D., Scholz D., "Box Wing Fundamentals – An Aircraft Design Perspective", *Deutscher Luft- und Raumfahrtkongress*, Bremen, 2011, publication pending.
[18] Niță M., Scholz D., "From Preliminary Aircraft Cabin Design to Cabin Optimization", *Deutscher Luft- und Raumfahrtkongress*, Hamburg, 2010.

NeoCASS+, A Conceptual Design and Simulation Framework for Morphing Aircraft

L. Cavagna, **A. De Gaspari**, **S. Ricci** and **L. Riccobene**
Dipartimento di Ingegneria Aerospaziale – Politecnico di Milano
Via La Masa 34, 20156, Milano, Italy

e–mail: sergio.ricci@polimi.it

Keywords: *Morphing aircraft, Conceptual design, Active camber*

Abstract

The interest in the design and implementation on modern aircraft of morphing concepts, able to guarantee the capabilities to adapt their shape based on the different flight missions request, is more and more growing. The paper summarizes the research activities performed by Politecnico di Milano in the field of morphing wings aiming at investigating possible implementation techniques for the active camber morphing concept. Special attention is focused on the development of a design framework, named NeoCASS+, that can be used right from the conceptual design phase to perform multi-fidelity analyses by modeling the interaction between three different design levels or tools. At first the aircraft conceptual design level, where both the aeroelastic analysis and optimization capabilities as well the possibility to consider and evaluate the benefits of morphing concepts are included. Second, a morphing airfoil shape optimization tool coupled with a compact airfoil geometry representation approach. Finally, a synthesis of compliant mechanisms tool able to support the designer in the definition of the structural configuration of morphing wing.

1 Introduction

The increasing interest in developing morphing aircraft is mainly based on the idea to adapt their shape to optimize specific performance indices during the mission. It allows to satisfy the more stringent environmental impact constraints, such as the ones established by the so called Strategic Research Agenda 2 [1] by ACARE. Indeed, looking at the medium and long term period, it is evident that significant steps forward to more efficient aircraft, able to meet direct and indirect greening requirements, will be achievable only by enhancing the aircraft capability to adapt its shape so to maximize the efficiency during all the flight envelope [2].

Nevertheless, the design of these kind of devices requires the availability of ad–hoc developed procedures able to tackle the conflicting requirements such as the high deformability requested to change the airfoil shape coupled to the load carrying capability. Despite the many morphing concepts proposed in the literature, few of them have been taken into account by civil aircraft industry till now due to many reasons, among the others the non–fully demonstrated efficiency of such devices, the certification related issues, the maintenance difficulties

and, in summary, the actual unclear advantage in terms of global direct operational costs.

One of the possible way to make morphing concepts more appealing for aerospace industry could be the development of design tools able to take into account of possible, realistic morphing concepts, and to evaluate their benefits, right from the conceptual design phase, where the final aircraft configuration has still to be frozen.

This paper presents a Design and Simulation Framework for Morphing Aircraft called NeoCASS+ that, starting from the previous experience gained during the recent EU–FP6 funded project named SimSAC, allows for a fast conceptual design loop, including aeroelastic and MDO analysis, of modern aircraft equipped with morphing concepts. One of the main advantage of the proposed framework is the capability to quickly generate, almost hand–off, the full aeroelastic model of free flying aircraft, so allowing to investigate many and different aircraft configurations. In particular, the emphasis is here directed to the active camber morphing concept and to the evaluation of the impact of continuous chord–wise and span–wise camber variation on trimmed configuration in terms of lift to drag ratio or load distribution.

2 The Morphing Aircraft Framework

The Morphing Aircraft Framework here proposed is depicted in Fig. 1. It is based on three main components: the NeoCASS (Next generation Conceptual Aero-Structural Sizing Suite) suite [3, 4], a tool for the aerodynamic optimization of morphing airfoils and finally a dedicated design tool for the optimal structural configuration of morphing wings, named MorfeO (MOrphing aiRFoil dEsign and Optimization). The aim of each component can be summarized as follows:

- NeoCASS is the kernel of aircraft conceptual design phase, allowing to both provide the structural sizing as well as the aeroelastic analysis and optimization capabilities. Here it is extended with the capability to quickly model variable camber airfoils.

- The aerodynamic optimization module is invoked by NeoCASS to investigate the effect of adopting morphing technologies in place of classical control surfaces, able to modify the airfoil shape. Indeed, it is used to optimize the airfoil shape to match pre-defined mission requirements specified by NeoCASS but maximizing the aerodynamic efficiency.

- Finally, once an optimal airfoil shape has been defined, MorfeO is used to design its internal structure, i.e., the adaptive rib, so to guarantee that the deformed shape of the airfoil matches as much as possible the optimal aerodynamic shape with the minimum of energy requested for actuation.

The entire procedure is linked to a compact airfoil representation technique (CST – Class/Shape function Transformation), based on the approach proposed by Kulfan in [5], which on one hand is used by the airfoil shape optimization and by MorfeO, on the other hand helps to introduce mesh deformations into the aircraft aerodynamic model where the morphing control surfaces are placed. The CST tool allows to deform the global shape of the airfoil without affecting its local regularity and it is based on merging a Shape Function with a baseline shape described by a Class Function. The Class Function is a non-linear function that mathematically defines a variety of basic airfoil shape, while the Shape Function is decomposed into scalable components represented by the Bernstein polynomials.

The whole framework, in the current version, includes the variable camber morphing concept only allowing to investigate the effect of seamless control surfaces, like ailerons and flaps, in place on traditional ones at conceptual design level. It adopts both typical low fidelity tools, for the aerodynamic and struc-

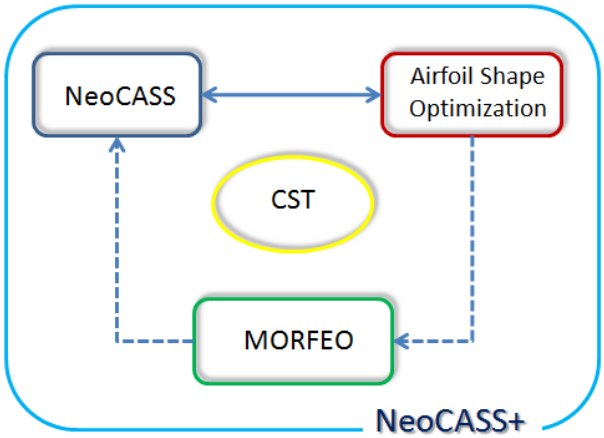


Figure 1: NeoCASS+scheme

tural modeling, as well as easy interfaces with medium to high fidelity aerodynamic tools.

In the following sections each component of NeoCASS+ framework will be briefly described.

3 NeoCASS

NeoCASS *(Next generation Conceptual Aero Structural Sizing)* is a suite of modules that combines state of the art computational, analytical and semi–empirical methods to tackle all the aspects of the aero-structural analysis of a design layout at conceptual design stage. It gives a global understanding of the problem at hand without neglecting any aspect of it: weight estimation, initial structural sizing, aerodynamic performances, structural and aeroelastic analysis from low to high speed regimes, divergence, flutter analysis and determination of trimmed condition and aerodynamic derivatives both for the rigid and deformable aircraft. More specifically, NeoCASS includes two main modules, respectively named GUESS *(Generic Unknowns Estimator in Structural Sizing)* and SMARTCAD *(Simplified Models for Aeroelasticity in Conceptual Aircraft Design)*. A third, external module, named W&B, is indeed necessary in order to have a first estimate of the non-structural masses and their location for the estimation of inertial loads.

The semi-analytical module named GUESS , based on a modified version of the AFaWWE code (*Analytical Fuselage and Wing Weight Estimation*) [6], is run to produce for the whole airframe a first-try stiffness distribution on the basis of user-defined sizing maneuvers. The first trial structural sizing is performed in a fully stressed design condition and introduces structural instability limits for compressed panels and stiffeners. However, no aeroelastic effect is considered during this sizing phase, due to the implicitness of the problem. Thus, once the initial structural sizing and the first stiffness distribution is determined by GUESS , a structural and aerodynamic mesh is automatically generated to allow for the successive numeric aeroelastic assessment and optimization, run in SMARTCAD .

GUESS estimates only structural primary masses, starting from W&B results, which then are refined by means of a regression to account for secondary structural weight; in the morphing simulation framework the masses related to conformal control surfaces and high lift devices can be updated.

At the conceptual level the main goal is focused on determining and representing at least structural/non-structural mass and stiffness distribution to satisfy strength, stiffness and stability requirements. Few simple lumped structural elements capable of giving equivalent structural behavior can be used for this purpose such as a linear equivalent plate, linear beam or non-linear beam to introduce geometry non-linear effects [7] in case, for example, large displacements have to be considered. These models lead to low–order algebraic problems, keeping the computational cost very low and allowing several configurations to be examined quickly.

The time for the preparation of the aeroelastic model is minimized: indeed, GUESS automatically generates both the structural as well the aerodynamic meshes as well all the necessary information for the fluid-structure coupling. A change in a geometry

parameter describing the aircraft at hand, is automatically reflected to the numerical model by means of automatic grid generation, allowing the whole process to be restarted for a successive analysis on the updated model, avoiding the need of detailed structural data.

SMARTCAD, on the other hand, is the aeroelastic kernel, dedicated to aero–structural analysis, which allows static and dynamic analyses to be carried out. It can be used as a stand-alone application once the structural and aerodynamic meshes are provided. Different kind of analysis can be carried out:

- Static analysis, linear buckling.
- Vibration mode calculations.
- Linearized flutter analysis.
- Linear/nonlinear static aeroelastic analysis, trimmed calculation for a free-flying rigid or deformable aircraft.
- Steady and unsteady aerodynamic analysis to extract derivatives for flight mechanics applications.
- Structural optimization.

Two low-fidelity aerodynamic methods are available in SMARTCAD , depending on whether steady or unsteady analysis is carried out: Vortex Lattice Method (VLM) [8] with camber contribution on normalwash once the airfoil description is provided and Doublet Lattice Method (DLM) [9] for the prediction of the generalized forces due to harmonic motion in the subsonic regime. Both methods are based on potential flow theory leading, under the hypothesis of irrotational, isentropic and inviscid flow to a Laplace's equation respectively for the linearized velocity or acceleration potential. The same geometric discretization of the aerodynamic surfaces as a flat lifting surface with singularities and collocation points located respectively at $\frac{1}{4}$ and $\frac{3}{4}$ of each panel chord is used so that Kutta condition is satisfied a priori. Thus, the same mesh is used (see Fig. 2), with the only exception that

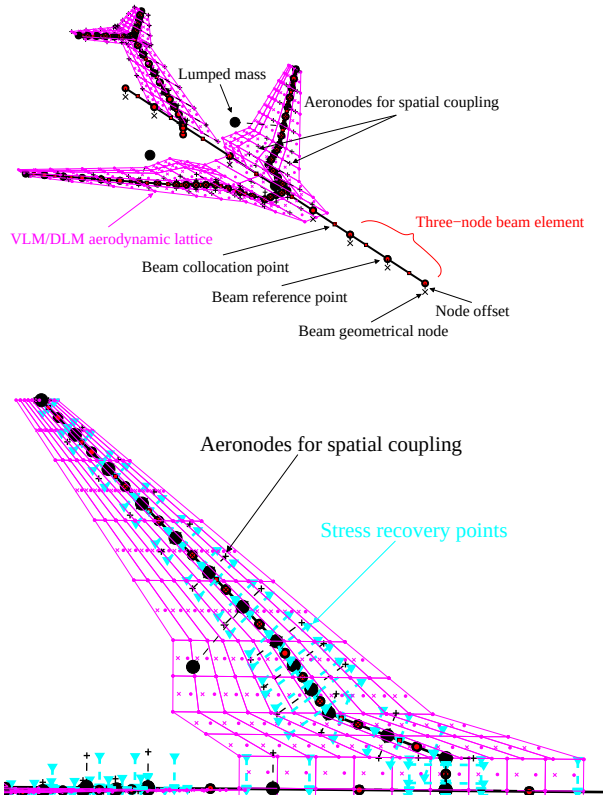


Figure 2: Aerodynamic meshes

the VLM takes into account camber contributions through boundary conditions and requires trailing vortexes for wake modeling.

Finally, an interface scheme to exchange displacements and loads between the structural grid and the aerodynamic boundary surface is required. Two class of methods are available for this purpose within SMARTCAD : a scheme based on Moving Least Squares (MLS) method [10] and the Radial Basis Function (RBF) method [11]. Both methods ensure the conservation energy transfer between the fluid and the structure and they are suitable for the treatment of complex configurations. To avoid interpolating rotations and using the same algorithms in a straightforward way also for CFD meshes, extra points are added along the wingbox through rigid arms. Aerodynamic loads are distributed along both master beam nodes and these additional nodes. Loads are then reduced to lumped forces and moments on the former set of nodes.

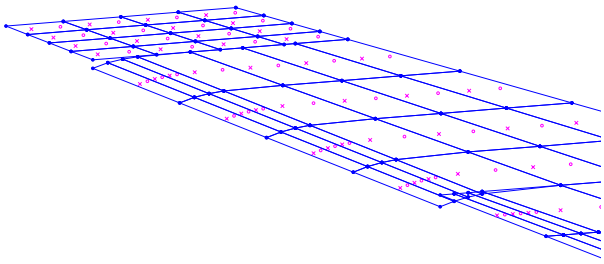


Figure 3: Active camber variation applied to the aerodynamic mesh.

3.1 Active Camber Concept Implementation

The performance requirements coming from NeoCASS+ is used to determine the aerodynamic characteristics of the airfoil providing the appropriate C_L for each mission segment, which reduces the feasible solution field, i.e. morphing shapes set, where the shape optimization algorithm described in Section 4 will work.

Once known the deformed shape, it can be exported as a CST object so that NeoCASS+ could manage active camber. In the vortex lattice method the aerodynamic is referred to a planar surface, including mean line effects by changing normal orientation accordingly. Control surfaces are managed rotating, physically, the related mesh portion, thus imposing the correct boundary conditions.
The active camber is thus implemented at aerodynamic mesh level: it acts as a control surface, deforming a portion of the mesh where the conformal surface is located. An *airfoil* shape change is translated in a *mean line* change, section by section, as sketched in Fig. 3. Aerodynamic loads change accordingly and are transferred to structural grid. The camber variation law used to deform the aerodynamic grid is the result of the airfoil shape optimization, which will be described in the following section.

Results on a civil aircraft equipped with conformal control surfaces undergoing a roll maneuver [12] showed that, despite relevant performance gains on the rigid aircraft, a thorough aeroelastic analysis should be pursued: when aeroelastic effects are introduced, it can be noted that the adaptive solution in fact worsen roll performance as speed increases. This is due to pressure distribution on wing section, as depicted in Fig. 4: the classical aileron C_p displays a peak of pressure next to the hinge axis and then quickly decreases until trailing edge; this peak shifts distribution resultant toward trailing edge inducing a nose-down moment which twists the airfoil. When deploying the conformal surface, the peak is lower due to smooth deflection but the pressure remains high until trailing edge, shifting even further aerodynamic force and thus generating a higher nose–down moment compared to classical aileron; this induces a higher twist of the section, lessening surface effectiveness.

The gain could become a loss when aeroelastic effects are introduced. A possible counter–measure is change the active camber concept, not relying only on the trailing edge.

It was observed [13] that a leading edge control surface basically acts on aerodynamic moment while keeping unchanged the force: an opposite deflection of leading and trailing edge thus allows to control wing twist.

4 Morphing Airfoil Shape Optimization

In a first step, the airfoil requirements are derived by NeoCASS taking into account for realistic morphing concepts and evaluating their impact on the trimmed configuration in terms of complete aircraft lift and load distribution. Then, the optimal morphing airfoil for each flight condition is obtained by using a shape optimization solver. It is based on medium high–fidelity 2D aerodynamic codes, and deals with finding the best airfoil shape change able to meet the local 2D lift coefficient under specific structural constraints related to the role played by the skin. The wing region where the morphing device is applied is obtained as a blended surface connecting a family of optimized morphing airfoils.

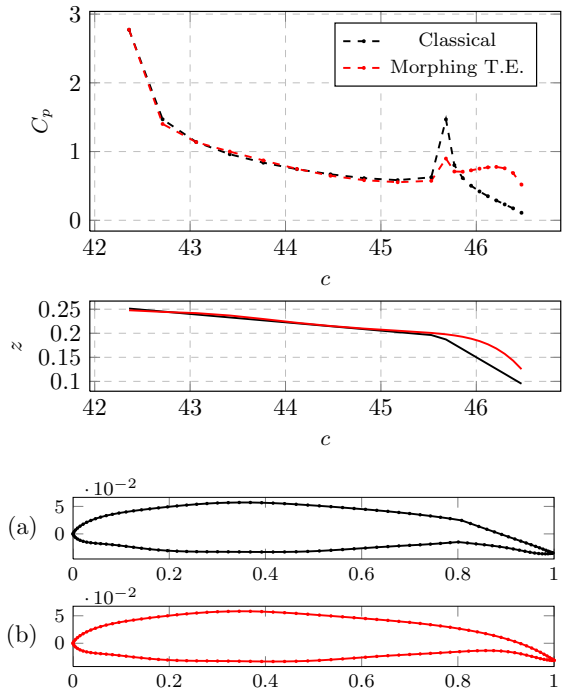


Figure 4: Comparison between chordwise distribution of the conventional and morphing control surface lift coefficient

The results of this airfoil shape optimization are used in two different ways. They represent the feedback for the NeoCASS aircraft full model that can be used to update the shape variation law which introduces into the aircraft model the aerodynamic mesh deformation corresponding to the shape changes of the morphing control surfaces. On the other hand, the best morphing airfoil shapes can be passed to MorfeO where they are used as target shapes in order to design the internal compliant structure which must be able to produce the desired shape change under the the external aerodynamic load corresponding to the flight conditions defined during the aeroelastic analysis performed on the free flying aircraft model. Once the morphing mechanism is synthesized and the power actuation is estimated, they can be used in order to update the morphing devices masses into the full aircraft model.

The airfoil optimization can be viewed as a high–fidelity optimization step that provides the best airfoil change able to meet the requested local 2D lift coefficient [14]. Moreover, the shape optimization of morphing airfoils is a critical task because it requires to find the most suitable morphing strategy able to maximize the aerodynamic efficiency and at the same time to limit the deformation energy of the skins and the power of actuator necessary to control the shape change. Special attention must be devoted to the definition of the constraints and to the optimization problem formulation.

4.1 Constraints

The definition of morphing airfoils is strongly influenced by the presence of the skins which, through the wing joints, are an obstacle to the motion to achieve the shape change required by a specific mission. The shape optimization here presented is able to guarantee the most efficient aerodynamic shape as well as the minimum or a limited level of axial and bending stresses in the skin. This is possible thanks to the adoption of the CST technique, here extended to both leading and trailing edge continuous control surfaces (see [14, 15] for more details about the adopted mathematical formulation).

The CST parametrization technique is well suited to represent the shape changes of morphing airfoils but also to describe the structural behavior of the skin. Due to the fully analytical description of airfoil and to the reduced number of parameters requested to describe its morphing shape, the geometric calculation of the length $L(x)$ and curvature $\kappa(l)$ of the skins is immediate. Moreover, the length and curvature of airfoil upper and lower surfaces are geometrical quantities which could be strictly related to the structural properties of the morphing airfoil skin. The maximum stress consists of two term: the stress due to axial tension or compression σ_{axial} and the stress due to bending σ_{bend}.

When the airfoil shape changes due to a morphing process, it is easy to compute the length of both undeformed and morphing airfoil surfaces and to estimate the axial stress

that is required to stretch or compress the airfoil skin. The bending stress is computed calculating the curvature difference between the initial and the final airfoil shape with respect to the two normalized arc lengths l, as follow:

$$\Delta\kappa(l) = \kappa_m\left(x(l)\right) - \kappa_u\left(x(l)\right) \tag{1}$$

where κ_u and κ_m are the curvature functions of the undeformed and morphing airfoil, $x(l)$ is the inverse of normalized arc length function that can be computed starting from $L(x)$. According to Euler–Bernoulli beam theory, the maximum bending stress along the skin can be calculated from the curvature difference function as:

$$\sigma_{bend} = \frac{Et}{2}\left(\Delta\kappa(l)\right) \tag{2}$$

where E is the Young's Modulus and t is the skin thickness. Since Equation 2 is based on the difference between the final and initial curvature distribution, it is general and suitable to compute stress values due to shape changes corresponding to large displacements. The available length and curvature variations of the airfoil's upper and lower surfaces, strictly related to the axial and bending stresses generated when the skin is forced to assume the morphing airfoil shape, can be used as explicit constraint function to add to the lift coefficient constraint $C_L^{desired}$ coming out from the conceptual simulation performed by NeoCASS.

4.2 Problem Formulation

For the flight conditions considered in NeoCASS, the airfoil shape is obtained by solving a minimum drag airfoil optimization problem subject to lift coefficient requirements and skin structural constraints. Thus the shape optimization here presented is used to guarantee that the morphing airfoil has optimal aerodynamic characteristics over different flight conditions and under structural constraints based on the CST geometrical quantities.

Although the shape optimization procedure is general and any CFD solver can be used, at the moment the shape optimization is coupled to the 2D viscous solver named XFOIL [16]. The optimization algorithm used here is the Sequential Quadratic Programming algorithm (SQP).

In this optimization problem, fixed parameters are the structural box front C_{FS} and rear C_{RS} spar positions, the desired morphing leading edge ΔZ_{LE} and trailing edge ΔZ_{TE} deflection with respect to the undeformed airfoil. As these informations the aerodynamic parameters required to perform the airfoil analysis (Angle of Attack AoA of trim, altitude h, speed V_∞, and Reynolds number Re), corresponding to the different flight conditions, are directly derived by the NeoCASS conceptual design tool.

Two different optimization problems for airfoil trailing and leading edge can be defined. In both cases the objective function consists of two terms: the first one acts on the airfoil region which must be kept undeformed, while the second one is the aerodynamic drag C_D. While the aerodynamic term controls the shape of the morphing trailing or leading edge, the other one works to fit the undeformed airfoil region corresponding to the structural box defined by the front and rear spar position. Since the two terms affect different airfoil regions, they do not represent conflicting design objectives. One simple way to combine these two terms is using a simple multi–criteria approach based on the weighted sum of the two terms. In addition to the constraints previously described, an inflatable term ΔA, representing the maximum allowable airfoil internal area variation, with respect to the undeformed one, is introduced to avoid the generation of unrealistic morphing shapes with a too much stretched thickness.

The general optimization problem which allows to determine the morphing leading or trailing edge shape change can be stated as follows:

Minimize:

$$w \cdot \|z_m(x_{box}) - z_u(x_{box})\|_2^2 + (1-w) \cdot C_D \quad (3)$$

such that:

$$C_L \geq C_L^{desired}$$
$$(\sigma_{bend,\,max}) < \sigma_{allowable}$$
$$|L_m - L_u| \leq \Delta L \cdot c$$
$$|A_m - A_u| \leq \Delta A \cdot A_u$$

where w is the weighting coefficient, the first term of the objective function represents the Least Square Error (LSE) between the undeformed and the deformed airfoil shape, x_{box} is the structural box region abscissa, while the z_u and z_m are the undeformed and morphing airfoil ordinate. The first constraint represents the minimum lift coefficient requirement, next constraint allows to define the maximum allowed structural stress, while the third constraint represents the maximum allowed length variations, in chord percentage, of the leading edge skin or the upper and lower surface skins. Finally, the fourth term represents the inflatable term, where A_m and A_u are the internal areas of the deformed and undeformed airfoil.

Fig. 5 shows an example of leading and trailing edge shape change together with the morphing bending stress distribution along the skins obtained by applying the CST based shape optimization to a $NACA\ 63_2215$ and placing the front spar at $C_{FS} = 0.2\,c$ and the rear spar at $C_{RS} = 0.6\,c$. The aerodynamic parameters correspond to a flight condition with $Re = 3.E{+}6$, $AOA = 9$ degrees for the leading edge and $Re = 3.E + 6$, an Angle of Attack of $AOA = 6$ degrees for the trailing edge.

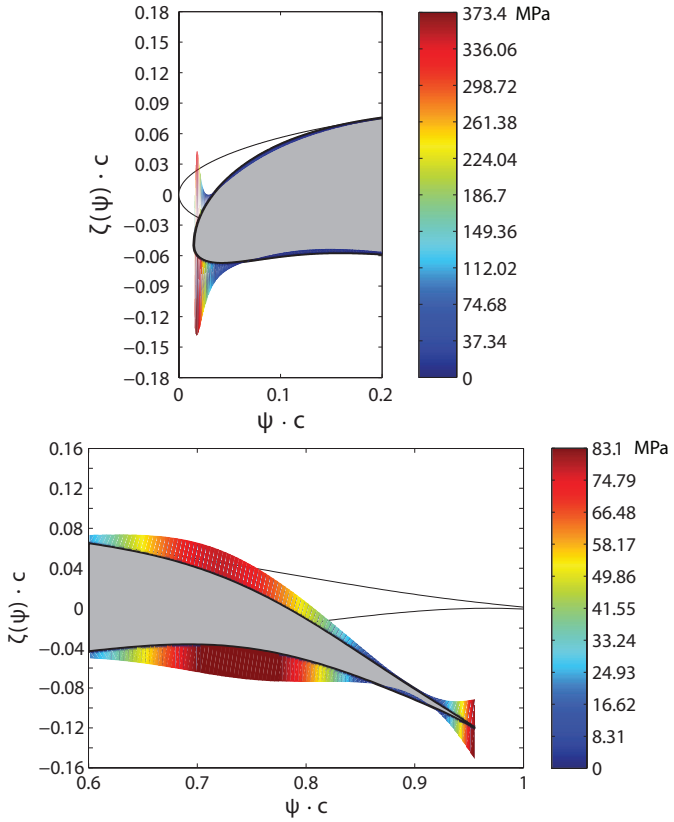


Figure 5: Bending stress distribution of the morphing airfoil coming out from the shape optimization by imposing a leading edge deflection of $-0.05\,c$ and a trailing edge deflection of $-0.12\,c$.

5 MorfeO

MorfeO *(MOrphing aiRFoil dEsign and Optimization)* is the part of a more general tool for the synthesis of compliant mechanisms [14, 17], dedicated to the design of morphing airfoils. A morphing airfoil includes the skin, a structure and a mechanism. Any process to design a morphing airfoil must simultaneously include these elements.

MorfeO allows to conciliate the conflicting requirements of deformability, load–carrying capability and low weight design systems by means of the synthesis of compliant mechanisms based on the *distributed compliance* concept, in which only flexible elements are employed instead of rigid links connected by flexure hinges. These lightweight systems are optimized to spread elastic strain over the entire structure so that all its elements share the actuator input load to produce large deformation. A compliant structure without stress concentrations is suitable for load bearing applications, as requested by aircraft structures.

Once the morphing airfoil shape optimization is completed, MorfeO allows to the design of the best airfoil internal structural configuration able to match, once actuated, the optimal

shape defined in 4.2.

5.1 Load Path Representation for the Design of Morphing Airfoils

MorfeO is based on the load path representation which is a particular design *parametrization of the stiffness tensor* employing load paths to represent different structural topologies, cross–sectional beam areas for the sizing problem and the position of essential points for the shape problem [15]. This unified approach is based on beam element models so that the results are steered to solutions based on distributed compliance concept. This approach is valid for the synthesis of Single–Input Single–Output (SISO) compliant mechanisms [18, 19] and it can be successfully applied to the design of Single–Input Multi–Output (SIMO) compliant mechanisms for the structural shape control [20].

When the load path representation is applied to the shape control problems like the morphing airfoils, it allows to optimize a compliant structure able to transfer the input actuator force to a set of so–called *active output points* placed along the boundary of the structure, in order to change its shape and minimize a functional, under several constraints. Load paths are physical connection sequences between three different types of points: the input actuator, structure constraints and active output points. The three types of essential points define as many load paths connecting load input and active output points ($InOut$ paths), load input and constraint points ($InSpc$ paths), and constraint points and active output points ($SpcOut$ paths). A fourth type of characteristic points is represented by the structure internal points which are the load path intermediate connections.

When the load path representation is applied to shape control problems, the design variables include path sequence (Seq), binary path existence variable (Top), internal point coordinates ($InterLoc$) and cross sectional load path sizes (Dim). Moreover, the design variables also include load path output destinations ($Dest$) and structure boundary sizes ($hBound$).

A number of points placed along the airfoil skin contour, greater or equal to the number of active output points, is used to minimize the Least Square Error (LSE) between the deformed curve and the target shape which comes out from the morphing airfoil shape optimization level described in Section 4. The optimization problem tries to minimize the least square error under size constraints for the load path beam elements and structure boundary elements, internal point boundaries, two global connectivity equations, stress and buckling constraints and the elastic equilibrium equation. In order to calculate the deformed curve every set of load path is transformed into a sequence of Finite Volume Beams [21].

The Finite Volume C^0 Beam is a particular type of beam, usually adopted as deformable connection component in the multibody applications. It is a three–node non–linear beam which proved to be intrinsically shear–lock free. The finite–volume approach leads to the collocated evaluation of internal forces and moments, as opposed to usual variational principles which require numerical integration on a one–dimensional domain. Since the use of non–linear analysis can be fundamental to describe the behavior of mechanisms subject to large displacements, the equilibrium equation is represented by the residual convergence. Finally, the structural solver incorporates modal, buckling, static linear, non–linear analysis modules.

When the output active point number and locations are defined, the designer can specify the minimum discretization mesh size of the skin finite volume beam model, in order to improve the resolution of the results. A number of points placed along the airfoil skins model, greater or equal to the number of active output points, can be used to minimize the LSE. The additional points introduced along the boundary allow to spread the pressure load, coming out from the aerodynamic analysis, on the airfoil skins. Moreover the designer can define, in

an graphical interactive way, both the internal point locations and a set of load paths.

MorfeO is coupled to the CST tool in in order to assist the engineer in the preparation of optimization problem. Besides providing the target curve, the CST tool defines the design domain and analytically imposes the structure internal point boundaries included in the optimization constraints. Moreover, it allows to identify the optimal position of the active output points and helps to compute the external aerodynamic load. Morphing airfoil meets the load–carrying requirement if they are able to adapt their shape and to maintain it under the external aerodynamic load corresponding to the same flight conditions defined in NeoCASS. MorfeO includes the interaction between the mechanism/structure and the fluid by means of a technique based on the Radial Basis Function (RBF) in order to build a coupling between structural and aerodynamic models.

5.2 Genetic Optimizer

Because the deformation of the airfoil skins is influenced by both the topology and the dimensions of the compliant mechanism, it is important to unify and simultaneously address the topology and size design of the airfoil structure. For this purpose, MorfeO incorporates a customized Genetic Optimizer where the individuals making up the population, are composed by mixed–type design variables and the generation of each new population is produced by selection, crossover and mutation dedicated strategies which represent the kernel of the whole design tool. Writing dedicated selection, crossover and mutation subroutines allows to combine the topology synthesis, the size and shape optimization into the same process, by imposing connectivity, stress and buckling constraints simultaneously.

5.2.1 Multiobjective Approach

The synthesis of morphing airfoil is defined as the design of a structure able to convert a displacement or a force in order to efficiently transfer the deformation work from an input point (actuator) to the output points placed along the skins. This requirement is only kinematic, and additional structural requirements are necessary in order to maintain the undeformed or deformed shape under external loads. Considering the distributed compliance concept implies an indirect control of the overall stiffness of the structure. In order to design a compliant mechanism able to meet both kinematic (motion) and structural (load–carrying) requirements, the design is decomposed into several parts considering the mechanism design and the structure design, respectively, for a number of load conditions corresponding to the flight condition analyzed in NeoCASS . These design problems are combined by incorporating a multiobjective optimization, based on the concept of Pareto domination, in the genetic optimizer. The Pareto–optimal set of resultant fitnesses is a set of solutions such that it is not possible to improve any objective without, at the same time, worsening another; all the other solutions are *dominated* by these better ones and can be discarded.

In general, more than one load conditions corresponding to the undeformed shape and more than one target shapes corresponding to different flight conditions can be considered. On one hand, the morphing airfoil internal configuration which minimize the LSE is one that maximize the virtual work required because the displacements of the output points allow to match the desired airfoil shape change [22], under the input actuation load and the external aerodynamic loads corresponding to a given flight condition. On the other hand, the structural requirement can be met by solving a system where the input point is fixed and the external aerodynamic loads are applied along the airfoil skins. Minimizing the LSE between the deformed curve and the undeformed airfoil allows to account for the resistance of the workpiece.

Fig. 6 show an example of leading and trailing edge design, obtained by applying MorfeO to the $NACA\,63_2215$ airfoil and using

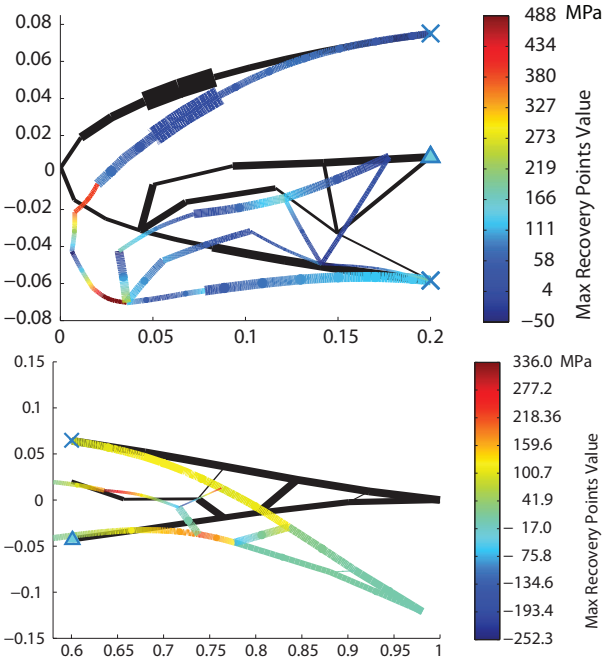


Figure 6: Finite Volume Beam based leading and trailing edge model with aerodynamic loads and corresponding maximum morphing normal stress distribution.

the target shapes of Fig. 5.

6 Conclusions

The paper describes a framework for the conceptual design of aircraft equipped with morphing devices. It is composed by three main components. The first one, i.e., NeoCASS, is a consolidated tool that has been already proven and validated since many years. The other two, for the aerodynamic optimization and for the optimal structural design based on compliant structures concept, i.e., MorfeO, have been recently developed. An exhaustive validation of the complete proposed framework is not easy due to the high number of tools available inside the framework but also because of the difficulty to find public data and reference results concerning morphing aircraft. Nevertheless, the validation of the single components of the framework is under way and some results showing the successfully application of the proposed methods have been included in the paper.

References

[1] Personalities, *European Aeronautics: A Vision for 2020*, 2001, ISBN 92-894-0559-7.

[2] McGowan, A. R., Horta, L. G., Harrison, J. S., and Raney, D. L., "Research Activities within NASA's Morphing Program," RTO Meet Proc 36 RTO–AVT Meet On Structural Aspects of Flexible Aircraft Control, 1999.

[3] Cavagna, L., Ricci, S., and Travaglini, L., "Structural Sizing and Aeroelastic Optimization in Aircraft Conceptual Design Using NeoCASS Suite," *AIAA 2010–9076*, 13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference, Fort Worth, TX, USA, 13–15 September 2010, pp. 1–26.

[4] Cavagna, L., Ricci, S., and Travaglini, L., "NeoCASS: an Integrated Tool for Structural Sizing, Aeroelastic Analysis and MDO at Conceptual Design Level," *AIAA 2010–8241*, AIAA Atmospheric Flight Mechanics Conference, Toronto, Canada, 2–5 August 2010, pp. 1–30.

[5] Kulfan, B. M., "Universal Parametric Geometry Representation Method," *Journal of Aircraft*, Vol. 45, No. 1, January–February 2008.

[6] Ardema, M., Chambers, A., Hahn, A., Miura, H., and Moore, M., "Analytical Fuselage and Wing Weight Estimation of Transport Aircraft," Tech. Rep. 110392, NASA, Ames Research Center, Moffett Field, California, May 1996.

[7] Cavagna, L., Masarati, P., Ricci, S., and Mantegazza, P., "Development and Validation of an Investigation Tool for Nonlinear Aeroelastic Analysis," *Proceedings of the* 26th *ICAS Congress*, Anchorage, Alaska, 14–19 September 2008.

[8] Moran, J., *An Introduction to Theoretical and Computational Aerodynamics*, John Wiley & Sons, 1984.

[9] Albano, E. and Rodden, W. P., "A Doublet–Lattice Method for Calculating the Lift Distributions on Oscillating Surfaces in Subsonic Flow," *AIAA Journal*, Vol. 7, No. 2, 1969, pp. 279–285.

[10] Quaranta, G., Masarati, P., and Mantegazza, P., "A Conservative Mesh-free Approach for Fluid-Structure Interface Problems," *International Conference on Computational Methods for Coupled Problems in Science and Engineering*, edited by M. Papadrakakis, E. Oñate, and B. Schrefler, CIMNE, Santorini, Greece, 2005.

[11] Beckert, A. and Wendland, H., "Multivariate Interpolation for Fluid-Structure-Interaction Problems Using Radial Basis Functions," *Aerospace Science Technology*, Vol. 5, 2001, pp. 125–134.

[12] Riccobene, L., *Methods and tools for the conceptual analysis of adaptive aircraft configurations*, Ph.D. thesis, Politecnico di Milano, Milano, 2010.

[13] De Gaspari, A., Ricci, S., Riccobene, L., and Scotti, A., "Active Aeroelastic Control over a Multisurface Wing: Modeling and Wind-Tunnel Testing," *AIAA Journal*, Vol. 47, No. 9, 2009, pp. 1995–2010.

[14] De Gaspari, A., *A Two Levels Approach for the Optimal Design of Morphing Wings Based On Compliant Structures*, Ph.D. thesis, Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, 2010.

[15] De Gaspari, A. and Ricci, S., "A Two Levels Approach for the Optimal Design of Morphing Airfoils," *AIAA 2010–9388*, 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Fort Worth, Texas, USA, 13 – 15 September 2010, pp. 1 – 15.

[16] Drela, M. and Youngren, H., *XFOIL 6.94 User Guide*, Cambridge, MA, 10 December 2001.

[17] De Gaspari, A. and Ricci, S., "Using Multilevel Optimization for the Design of MorphingWings Based on Compliant Mechanisms," *Applied Vehicle Technology (AVT-173) Panel Workshop on Virtual Prototyping of Affordable Military Vehicles Using Advanced MDO*, NATO Research and Technology Organisation (RTO), Sofia, Bulgaria, 18 – 18 May 2011, pp. 1 – 33.

[18] Lu, K.-J. and Kota, S., "Topology and Dimensional Synthesis of Compliant Mechanisms Using Discrete Optimization," *Journal of Mechanical Design*, Vol. 128, September 2006, pp. 1080–1091.

[19] Hetrick, J. A. and Kota, S., "An Energy Formulation for Parametric Size and Shape Optimization of Compliant Mechanisms," *Journal of Mechanical Design*, Vol. 121, No. 6, June 1999, pp. 229–234.

[20] Lu, K.-J. and Kota, S., "An Effective Method of Synthesizing Compliant Adaptive Structures using Load Path Representation," *Journal of Intelligent Material Systems and Structures*, Vol. 16, April 2005, pp. 307–317.

[21] Ghiringhelli, G. L., Masarati, P., and Mantegazza, P., "Multibody Implementation of Finite Volume C^0 Beams," *AIAA Journal*, Vol. 38, No. 1, January 2000, pp. 131–138.

[22] Shield, R. T. and Prager, W., "Optimal Structural Design for Given Deflection," *Journal of Applied Mathematics and Physics, ZAMP*, Vol. 21, 1970, pp. 513–523.

CEAS 2011

Modelling and Simulation 1

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

On Numerical Modeling of DBD Multi-Actuator System

S. Chernyshev, A. Kuryachii, D. Rusyanov, V. Skvortsov
Central Aerohydrodynamic Institute n.a. prof. N.E. Zhukovsky (TsAGI)
1, Zhukovsky str., Zhukovsky, 140180 Moscow region, Russia

Keywords: *dielectric barrier discharge actuators, modeling*

Abstract

The spatially periodic system of the dielectric barrier discharge actuators intended for laminar flow control on a swept wing is considered. The original scheme of the system ensuring the discharge formation only near one edge of every exposed electrode is proposed. The mathematical model of the dielectric barrier discharge in air is used for calculation of such main characteristics of every actuator as generated volumetric force and heat release. Numerical simulation carried out at one set of parameters of the formulated mathematical problem demonstrates the utility of the proposed scheme of multi-actuator system.

1 Introduction

The method of laminar flow control (LFC) on a swept wing based on an attenuation of the cross-flow-type instability by means of electrogasdynamic (EGD) influence on three-dimensional boundary layer in the vicinity of a wing leading edge was proposed at TsAGI more than 20 years ago [1]. Practical realization of this method seems to become realistic owing to the usage of EGD-actuators operating on the base of the dielectric barrier discharge (DBD) [2]. The concept of the proposed LFC method is illustrated on Fig. 1.

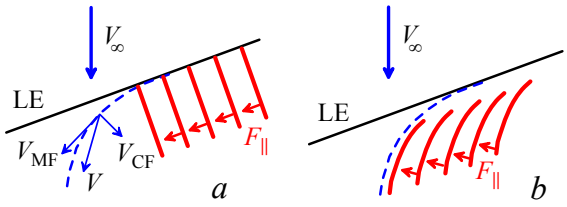


Fig. 1 The sketch of the LFC method

Here $\boldsymbol{V}_\infty$ is the free-stream velocity vector, LE is the swept wing leading edge, $\boldsymbol{V}$ is the gas velocity vector in some point inside a boundary layer, $\boldsymbol{V}_{MF}$ and $\boldsymbol{V}_{CF}$ are the main-flow and cross-flow components of the velocity vector, blue dashed curve shows the external inviscid streamline, red solid curves show the exposed electrodes of DBD-actuators, $\boldsymbol{F}_\parallel$ is the vector component of the volumetric force generated by every actuator and directed parallel to a wing surface.

The simplest configuration of DBD-actuator system is shown on Fig. 1, *a*. The exposed electrodes are placed continuously from lower to upper surfaces of a wing perpendicular to its leading edge. Volumetric force impact directed along the leading edge will produce a gas velocity component decreasing $\boldsymbol{V}_{CF}$. Moreover one can expect that the "filling" of the velocity profile in viscous flow directed along a critical line due to gas acceleration in this direction will

result in a rise of the stability of this flow and, hence, an increase in the critical Poll Reynolds number.

The more complex configuration shown on Fig. 1, *b* seems to be more effective and less energy consuming. The arrangement of the curvilinear exposed electrodes parallel to the external streamline will produce the volumetric force $\boldsymbol{F}_{\parallel}$ directly against $\boldsymbol{V}_{CF}$ thereby intensifying the effect of volumetric force on cross-flow. In this case the DBD-actuators may be placed at some distance from the critical line, namely, beginning at the line of the cross-flow-type instability origin. It will also reduce total power consumption. In both cases an attenuation of the cross-flow velocity $\boldsymbol{V}_{CF}$ results in a decrease of increments of spatial growth of steady-state modes of the cross-flow-type instability [3].

Practical realization of this LFC method demands development and optimization of the multi-actuator system spatially periodic along a wing leading edge. The optimal system must create a volumetric force $\boldsymbol{F}_{\parallel}$ large enough for prevention the laminar-turbulent transition caused by the cross-flow-type instability or the leading edge contamination at moderate electric power consumption.

Since 1998 several tens articles devoted to both theoretical and experimental study of DBD-actuators are published annually [4, 5]. Numerous numerical modeling, as a rule, deal with a single actuator [6-11]. The most detailed experimental investigations are executed with single actuator too [12, 13]. Experimental optimization of multi-actuator systems is more complex because of a presence of additional parameters, for example, the phase shift of the applied voltage on adjacent actuators, the distance between them, the widths of the exposed and insulated electrodes [14, 15]. Moreover the optimal system must ensure a discharge ignition only at one side of every exposed electrode and exclude it at the other side.

To minimize the volume of necessary experiments it seems urgent to carry out preliminary numerical research with the aim to reveal the most important parameters of the multi-actuator system and to estimate their influence on its main characteristics such as generated volumetric force and heat release (Joule dissipation).

2 Multi-Actuator System Arrangement

The scheme of the proposed spatially periodic system of DBD-actuators is shown on Fig. 2. Here *A* is the electroconductive structure (the wing envelope), the electric potential of this structure is taken as zero, *B* are the insulated electrodes under zero potential too, *C* are the exposed electrodes under alternating electric potential, *1* and *2* are the dielectric layers, *D* are the additional insulated electrodes under the same potential as the exposed electrodes. If the alternating potential has large enough amplitude, the barrier discharge turns on in the vicinity of the right edges of the exposed electrodes *C*.

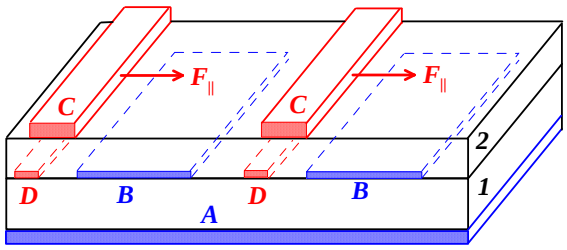


Fig. 2 The scheme of DBD-actuator system

The alternating volumetric electrostatic force arises in these regions due to electric charge separation in the discharge gap. The time-averaged horizontal component $F_{\parallel}$ of this force is directed from left to right and accelerates boundary layer flow in this direction. The additional electrodes *D* permit to attenuate drastically the electric field strength near the left edges of the exposed electrodes thereby to avoid the discharge ignition here and to increase the effectiveness of the system. One can hope also that the usage of the electrodes *D* will permit to reduce significantly the width of the exposed electrodes, the distance between them, and the thickness of the internal dielectric layer *1*.

Note, that the width of the exposed electrode must be large enough or one its edge must be covered with insulating film to avoid the formation of discharge at that edge in usual schemes of DBD-actuator with one buried electrode [15].

The miniaturization of the considered system of DBD-actuators is important for practical realization of the mentioned LFC method since the boundary layer thickness in the vicinity of a swept wing leading edge of civil airplane is very small (about 0.5 mm). Note, that the exposed electrodes may be flush mounted to minimize boundary layer perturbations. The example of successful application of low-sized DBD multi-actuator system for transonic shock control is presented in [16].

Main characteristics of various variants of the considered system may be evaluated owing to numerical simulation on the base of proper mathematical model.

3 Boundary Value Problem Statement

Numerical modeling of one spatial period of the described system is fulfilled in the calculation domain shown on Fig. 3 and consisting of subregions *1-4*. Here C_r is the right half of some exposed electrode and C_l is the left half of the next exposed electrode. The thickness of the insulated electrodes *B* and *D* is supposed to be zero.

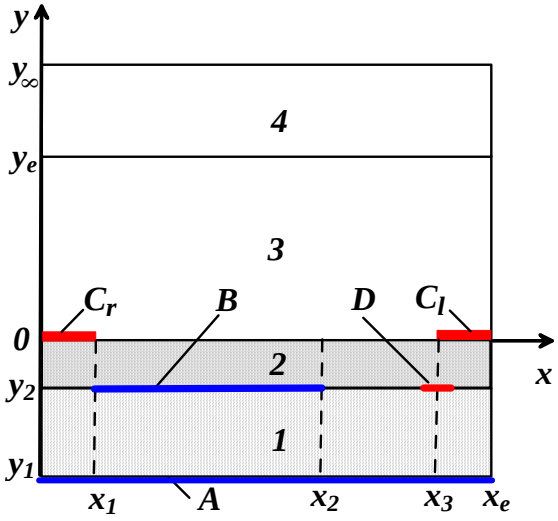


Fig. 3 Calculation domain

The electric potential φ and the electric field strength $\mathbf{E} = -\nabla\varphi$ are defined by the Laplace's equation in the dielectric layers *1* and *2* and in the external subregion *4* (indices 1, 2, 4), and by Poisson's equation in gas (index 3)

$$\Delta\varphi_i=0\ (i=1, 2, 4),\ \varepsilon_0\Delta\varphi_3=-e(n_p-n_n-n_e) \qquad (1)$$

Here ε_0 is the permittivity of vacuum, e is the elementary charge, n_p, n_n, n_e are the volumetric concentrations of positive ions, negative ions, and electrons, respectively. Relative dielectric constant of air is taken to be equal 1. The external subregion *4* is introduced to satisfy the condition of zero potential at infinity. The following transformation of y-coordinate is used here to move the external boundary $y=y_\infty$ at infinity: $y'=y_e(y_\infty-y_e)/(y_\infty-y)$ [17].

The periodicity conditions are stated on the left and the right boundaries of the whole calculation domain. The conditions of continuity of the potential and the electric induction are stated on the boundaries between the dielectric layers *1* and *2* and between the subregions *3* and *4*. The potential continuity condition and the relation between the discontinuity of the electric induction and the superficial charge are stated on the upper surface of the external dielectric layer *2*. Taking into account zero potential of the electrodes *A* and *B* and specified potential of the electrodes *C* and *D* the boundary conditions for the electric potential take the form

$$\begin{gathered}
\varphi(0,y,t)=\varphi(x_e,y,t)\\
\partial\varphi/\partial x(0,y,t)=\partial\varphi/\partial x(x_e,y,t)\\
(x,y)\in\Omega_{12}:\ \varphi_1(x,y,t)=\varphi_2(x,y,t)\\
\varepsilon_1\partial\varphi_1/\partial y(x,y,t)=\varepsilon_2\partial\varphi_2/\partial y(x,y,t)\\
(x,y)\in\Omega_{23}:\ \varphi_2(x,y,t)=\varphi_3(x,y,t) \qquad (2)\\
\varepsilon_2\partial\varphi_2/\partial y(x,y,t)-\partial\varphi_3/\partial y(x,y,t)=\sigma/\varepsilon_0\\
\varphi(x,y_1,t)=\varphi(x_1\le x\le x_2,y_2,t)=\varphi(x,y_\infty,t)=0\\
(x,y)\in\Omega_C,\ \Omega_D:\ \varphi(x,y,t)=-V_0(t)\sin(2\pi ft)
\end{gathered}$$

Here Ω_{12} denotes the boundary between the dielectric layers *1* and *2* except the electrodes *B* and *D*, Ω_{23} denotes the open surface of the dielectric layer *2*, Ω_C and Ω_D denote the surfaces of the exposed electrodes C_r, C_l and the insulated electrode *D*, respectively, ε_1 and ε_2 are the dielectric constants of the layers *1* and *2*, σ is the superficial charge on the surface Ω_{23}, V_0 and f are the amplitude and the frequency of the applied voltage.

The dynamics of charged particles is governed by three continuity equations for

positive ions, negative ions, and electrons without taking into consideration a convective transfer [8]

$$\begin{gathered}
\partial n_p/\partial t+\nabla \mathbf{\Gamma}_p=k_i N n_e-k_{dr} n_e n_p-k_r n_p n_n+S_{ph} \\
\partial n_n/\partial t+\nabla \mathbf{\Gamma}_n=0.2k_{at} N n_e-0.8k_{dt} N n_n- \\
-k_r n_p n_n+\alpha_{at}(0.2N)^2 n_e \qquad (3) \\
\partial n_e/\partial t+\nabla \mathbf{\Gamma}_p=k_i N n_e-k_{dr} n_e n_p-0.2k_{at} N n_e+ \\
+0.8k_{dt} N n_n-\alpha_{at}(0.2N)^2 n_e+S_{ph} \\
\mathbf{\Gamma}_i=z_i\mu_i n_i-D_i\nabla n_i,\ z_p=1,\ z_n=z_e=-1 \\
k_i=0.8k_{i\mathrm{N}}+0.2k_{i\mathrm{O}} \\
\lg k_{i\mathrm{N}}=-8.09-40.29/\gamma \\
\lg k_{i\mathrm{O}}=-8.31-28.57/\gamma \\
k_{dr}=2\cdot 10^{-7}(300/T_e)^{0.7} \\
k_r=2\cdot 10^{-6}(300/T_p)^{3/2} \\
\lg k_{at}=-10.21-5.7/\gamma \\
k_{dt}=9.2\cdot 10^{-13},\ \alpha_{at}=4.45\cdot 10^{-31}
\end{gathered}$$

Here $\mathbf{\Gamma}_i$ ($i=p, n, e$) are the vectors of the drift-diffusion fluxes of charged particles, N (cm^{-3}) is the volumetric concentration of air molecules, k_i is the ionization rate coefficient taking into account an origin of positive ions of both nitrogen and oxygen, $k_{i\mathrm{N}}$ and $k_{i\mathrm{O}}$ are the individual coefficients of nitrogen and oxygen ionization, k_{dr} is the rate coefficient of the dissociative recombination of electron with positive ion, k_r is the rate coefficient of the ion-ion recombination, k_{at} and α_{at} are the rate coefficients of two body and three body attachment of electrons to oxygen molecules, k_{dt} is the coefficient of detachment of electrons from negative ions, $\gamma=10^{16}E/N$ (V·cm^2) is the function of the reduced field, E (V/cm) is the absolute value of the electric field strength, T_e and T_p are the temperatures of electrons and ions, S_{ph} is the source of the background ionization in absence of the applied voltage.

The reactions rate coefficients $k_{i\mathrm{N}}$, $k_{i\mathrm{O}}$, k_{dr}, k_r, k_{at}, k_{dt} are measured in cm^3/s, α_{at} is measured in cm^6/s. The coefficient of the electron mobility μ_e and the electron temperature T_e depending on the reduced field function γ are calculated by the method described in [18]. The mobility coefficients of ions are defined as $\mu_{p,n}=C_{p,n}(N_0/N)(300/T_{p,n})^{0.3}$ (cm^{2}V^{-1}s^{-1}), $C_p=2.1$, $C_n=3.2$, $N_0=2.69\cdot 10^{19}$ cm^{-3}. A distinction of ions $T_{p,n}$ and neutrals T temperatures is neglected. The diffusion coefficients D_i ($i=p, n, e$) are calculated on the base of the Einstein relation using the mobility coefficients. The numbers 0.8 and 0.2 correspond to the contents of the nitrogen and oxygen molecules in the total concentration of neutrals N.

The equation system (3) is solved in the subregion *3*. The background ionization source S_{ph} is introduced in (3) only for unambiguous determining of the initial concentrations of electrons and negative ions at a given background concentration of positive ions n_0. As $k_i=0$ and $k_{at}=0$ at $E=0$, then setting the right sides of the first two equations (3) equal to zero and solving the obtained equations one can find

$$S_{ph}=n_p(k_{dr}n_e+k_r n_n)=n_0 n_e(k_{dr}+\xi k_r)$$

$$\xi=\frac{\alpha_{at}(0.2N)^2}{0.8k_{dt}N+k_r n_0}\cong 0.05\frac{\alpha_{at}N}{k_{dt}}$$

Hence the initial conditions for the equation system (3) take the form

$$\begin{gathered}
n_p(x,y,0)=n_0=10^3\ \mathrm{cm}^{-3} \\
n_n(x,y,0)=\xi n_0/(1+\xi) \qquad (4) \\
n_e(x,y,0)=n_0/(1+\xi)
\end{gathered}$$

The boundary conditions on the left and the right borders are the periodicity conditions

$$n_i(0,y,t)=n_i(x_e,y,t),\ \mathbf{\Gamma}_i(0,y,t)=\mathbf{\Gamma}_i(x_e,y,t) \qquad (5)$$

The boundary conditions for fluxes of ions and electrons Γ_{wi} ($i=p,n,e$) on both dielectric and metallic surfaces are deduced in [18] with the use of the relaxation approximation of Boltzmann equation for the velocity distribution function of charged particles and the approximated Maxwell approach instead of more strict consideration of the Knudsen layer. The physical model of interaction between charged particles and dielectric surface in nitrogen [19] has been completed for air and used for a statement the equations governing the accumulation of ions and electrons on dielectric surface. But in the present work a simplified form of the boundary conditions on the dielectric surface without taking into account a

decelerating influence of the electric field is used. The superficial charge σ is calculated using the approximated model of instantaneous recombination of electrons and ions on the surface [6]. So the boundary conditions for Γ_{wi} and the equation of superficial charge variation take the form

$$(x,y)\in\Omega_{23}:$$
$$\Gamma_{wp}=-0.5c_pn_p,\ \Gamma_{wn}=-0.5c_nn_n$$
$$\Gamma_{we}=-0.5c_en_e+\zeta_dc_pn_p \qquad (6)$$
$$\partial\sigma/\partial t=-0.5(c_en_e+c_nn_n)+(0.5+\zeta_d)c_pn_p$$
$$t=0:\ \sigma(x,t)=0$$

Here ζ_d is the coefficient of the secondary ion-electron emission on the dielectric surface.

The boundary conditions for fluxes Γ_{wi} on the open surfaces Ω_C' of the external electrodes C_r and C_l depend on the electric field strength on these surfaces E_w where it reaches extreme values. Positive E_w decelerates incident positive ions and secondary electrons, negative E_w decelerates incident negative ions and electrons. These conditions take the form [18]

$$(x,y)\in\Omega_C',\ E_w\le 0:$$
$$\mu_p|E_w|\le 0.5c_p:\ \Gamma_{wp}=-0.5c_pn_p$$
$$\mu_p|E_w|>0.5c_p:\ \Gamma_{wp}=\mu_pn_pE_w$$
$$\Gamma_{wn}=-0.25c_nn_nA_n(1-0.5B_n)^{-1}$$
$$\Gamma_{we}=-(0.25c_en_eA_e+\zeta_m\Gamma_{wp})(1-0.5B_e)^{-1}$$
$$(x,y)\in\Omega_C,\ E_w>0: \qquad (7)$$
$$\Gamma_{wp}=-0.25c_pn_pA_p(1-0.5B_p)^{-1}$$
$$\mu_nE_w\le 0.5c_n:\ \Gamma_{wn}=-0.5c_nn_n$$
$$\mu_pE_w>0.5c_n:\ \Gamma_{wn}=-\mu_nn_nE_w$$
$$\Gamma_{we}=-0.5c_en_e-2\zeta_m\Gamma_{wp}A_{e\zeta}$$
$$A(X)=\exp(-X^2)$$
$$B(X)=1-\mathrm{erf}(X)+2\pi^{-1/2}XA(X)$$
$$X_i=(e|E_w|\lambda_i/k_BT_i)^{1/2},\ c_i=(8k_BT_i/\pi m_i)^{1/2}$$

Here ζ_m is the coefficient of the secondary ion-electron emission on the electrode surface, λ_i, c_i, m_i ($i=p,n,e$) are the mean free path, the thermal velocity, and the mass of corresponding particles, k_B is the Boltzmann constant. The functions A and B reflect the decelerating effect of the electric field, A_i and B_i ($i=p,n,e$) relate to incident particles, $A_{e\zeta}$ relates to secondary electrons. It is supposed that the secondary electrons have the mean energy of 0.5 eV.

The mean free paths of ions and electrons presenting in the argument of A and B are estimated according to formulas [18]

$$\lambda_{p,n}=10^{15}/(4.1N)\ (\mathrm{cm})$$
$$\lambda_e=5.45\cdot10^{14}/(\theta_e^{1/3}N),\ \theta_e=8.62\cdot10^{-5}T_e$$

The external boundary conditions at $y=y_e$ imply that the diffusive fluxes through this boundary are much less than the drift fluxes far enough from actuator surface. So these conditions take the form

$$0\le x\le x_e,\ y_1=y_e:\ \Gamma_{wi}=z_i\mu_in_iE_y\ (i=p,n,e) \qquad (8)$$

The formulated boundary value problem (1)-(8) has been solved by the finite element method with the use of the first order Lagrangians and strongly nonuniform triangular mesh. The total number of the finite elements for considered geometry was about 150000.

4 Results of Calculations

Results of calculations carried out for one set of the stated problem parameters are presented and discussed below. The geometric parameters indicated on Fig. 3 are the following: x_1=0.25, x_2=1.5, x_3=2.75, x_e=3, y_1=-0.3, y_2=-0.1, y_e=2, y_∞=2.5 cm. The width of the electrode D is 0.1 cm. The height of the exposed electrodes equals 55 μm and they are deepened in the dielectric layer *2* up to 5 μm. Both upper and lower verges of these electrodes are rounded with the radius of curvature 5 μm to eliminate electrostatic singularity arising on sharp corners.

The following physical parameters were used in calculations: the pressure and the temperature of air p=760 torr and T=300 K, the amplitude and the frequency of the applied voltage V_0=7 kV and f=5 kHz, the dielectric constants of the insulation layers ε_1=2.1 (Teflon) and ε_3=3.5 (Kapton), the coefficients of the secondary ion-electron emission on the dielectric and the electrode surfaces $\zeta_d=\zeta_m$=0.01.

The discharge behavior strongly depends on electrostatic situation on the dielectric surface. The superficial charge distribution becomes quasi-periodic in time after several cycles of the applied voltage [18, 20]. The calculation of every cycle is very time-consuming. Therefore the time dependences of the horizontal volumetric force and heat release (Joule dissipation) are shown on Fig. 4 for the second cycle of the applied voltage. The potential of the exposed electrodes is shown by dashed curves. The volumetric force and the Joule dissipation are calculated as the integrals over the area S_3 of the subregion *3*

$$F_{||} = \int_{S_3} eE_x\left(n_p - n_n - n_e\right)dxdy$$

$$Q = \int_{S_3} e\mathbf{E}\left(\mathbf{\Gamma}_p - \mathbf{\Gamma}_n - \mathbf{\Gamma}_e\right)dxdy$$

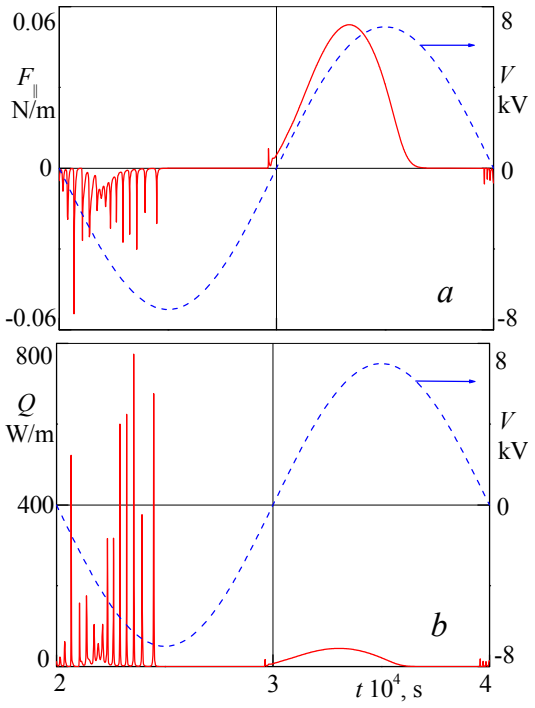


Fig. 4 Volumetric horizontal force and Joule dissipation during second cycle of discharge

During the negative-going potential of the exposed electrodes (so-called the forward stroke of DBD) the discharge consists of a series of micro-discharges turning on near the right edge of the exposed electrode C_r with the repetition time of 2-3 μs. The extreme values of the conduction current in micro-discharges achieve -(0.05-0.35) A/m. The current between them is very small. It is in a qualitative agreement with the numerical results of [6].

The forward stroke of DBD is accompanied by an intensive precipitation of electrons on the dielectric surface. Arising negative superficial charge weakens the absolute value of the electric field strength in the vicinity of the right edge of the electrode C_r. If the additional electrode D is absent, the electric field strength near C_r after several micro-discharges becomes less than the field strength near the left edge of the electrode C_l, therefore the next micro-discharge turns on near C_l. At the presence of the electrodes D the discharge turns on only near C_r.

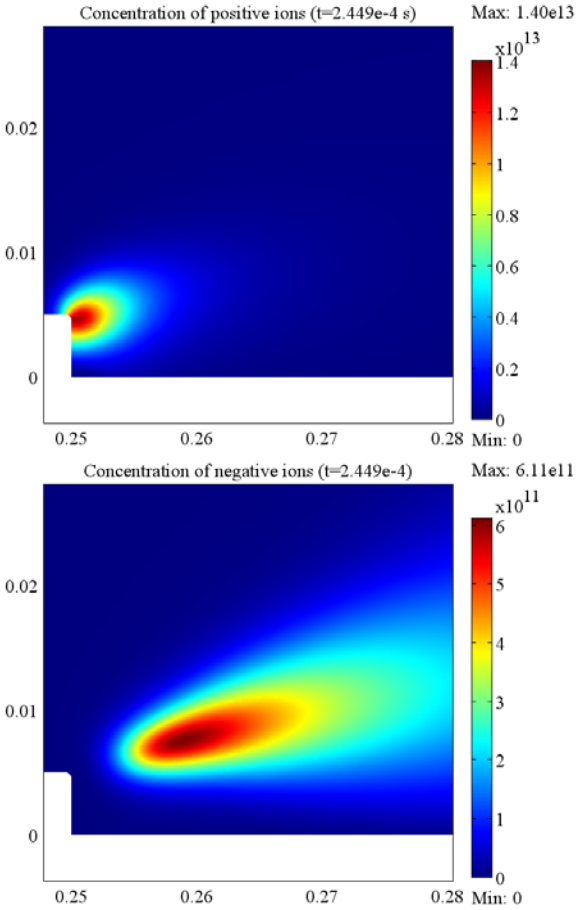


Fig. 5 Concentrations of positive (above) and negative (below) ions at t=2.449·10^{-4} s (coordinates in cm)

The space integrated horizontal force is negative during the forward stroke of DBD. Positive ions make a main contribution in force generation since their concentration exceeds significantly the concentrations of negative ions and electrons. The distributions of concentrations of ions of both signs in cm^{-3} are shown on Fig. 5 at the time moment

corresponding to maximal conduction current in the last micro-discharge.

Positive ions moving to the exposed electrode concentrate near its edge. Their maximal concentration $n_{p\,max}=1.4\cdot10^{13}$ cm^{-3} is achieved on the electrode surface. The maximal concentration of negative ions is less more than 20 times and achieved at about 0.1 cm from the electrode edge. The maximal concentration of electrons $n_{e\,max}=1.18\cdot10^{11}$ cm^{-3} is even less.

However the horizontal force during the forward stroke of the discharge may become positive at some conditions, namely, if this stroke is long-term and consists of large enough number of micro-discharges. This condition realizes at low enough frequency and/or large amplitude of the applied voltage. Figure 6 demonstrates the behavior of ion concentrations at fixed height y=0.01 cm and various distances from the electrode C_r. It is evident that at large enough distance the peaks of positive ions decrease in time whereas the peaks of negative ions remain almost constant. Therefore with time the space-integrated positive volumetric force generated by negative ions may exceed the absolute value of the force generated by positive ions.

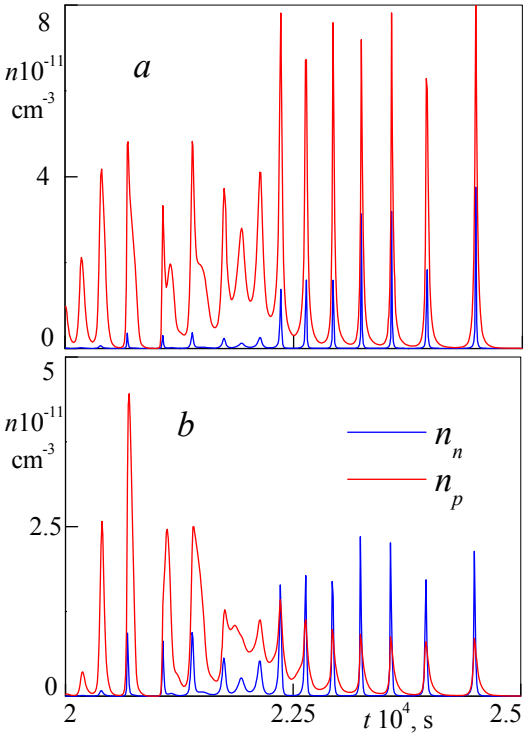


Fig. 6 Concentrations of positive and negative ions at the distance x=0.26 (*a*) and 0.28 (*b*) cm

Comparison of our results with the results of numerical simulation in [6] shows that the concentration of negative ions relaxes after micro-discharge much slowly in [6] where the relaxation of negative and positive ions was taken into account but the detachment of electrons from negative ions was neglected. However according to our calculations the values of the right-side terms of the second equation (3) determining the source and the sink of negative ions $S_{at}\equiv0.2k_{at}Nn_e$, $S_{dt}\equiv0.8k_{dt}Nn_n$, $S_r\equiv k_r n_p n_n$ in the point corresponding to maximum concentration of negative ions on Fig. 5 equal $S_{at}=1.182\cdot10^{19}$, $S_{dt}=1.098\cdot10^{19}$, $S_r=0.228\cdot10^{19}$ cm^{-3}s^{-1}. That is the intensity of detachment is comparable with the attachment intensity and more significant than ion-ion relaxation.

During the backward stroke of DBD (when the potential of the exposed electrode increases) the discharge seems like initial single spike and subsequent smooth nonmonotonic variation of the electric current, the volumetric force and the heat release (see Fig. 4).

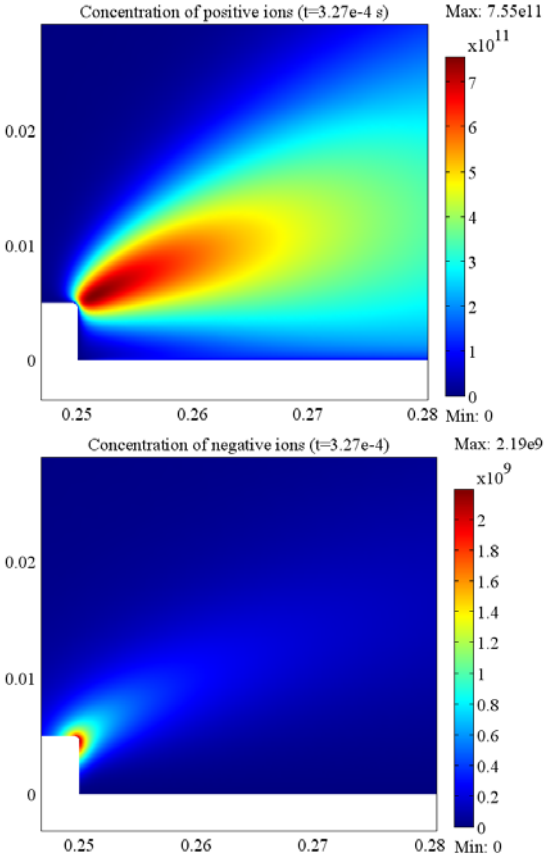


Fig. 7 Concentrations of positive (above) and negative (below) ions at t=3.27·10^{-4} s

The distributions of ion concentrations in the time moment corresponding to the maximal force during the backward stroke are shown on Fig. 7. It is evident that positive ions determine the horizontal force during this stage of DBD, therefore this force is positive.

The main characteristics of DBD-actuator are the volumetric horizontal force $<F_{||}>$ and the heat release $<Q>$ integrated over space and averaged in time during one cycle of the applied voltage and the relation $E=<F_{||}>/<Q>$ called the coefficient of energy efficiency [21]. Integration on time of the distributions $F_{||}$ and Q shown on Fig. 4 and dividing on the cycle period $T=2\cdot10^{-4}$ s resulted in $<F_{||}>=9.17\cdot10^{-3}$ N/m, $<Q>=16$ W/m and $E=5.74\cdot10^{-4}$ s/m. Note that about a half of Joule dissipation (7 W/m) liberates during the forward stroke of DBD when the generated horizontal force is negative. Nevertheless the obtained value of the energy efficiency coefficient is large enough. According to theoretical estimations presented in [1], a high energy efficiency of DBD-actuators is very important for effective attenuation of the cross-flow-type instability in compressible boundary layer.

The effect of DBD-actuators on boundary layer stability can be estimated theoretically by using the spatial distributions of time-averaged volumetric force and heat release in calculations of laminar boundary layer and characteristics of its stability. Joule dissipation concentrates in narrow region near the edge of the exposed electrode C_r since both the flux densities of all charged particles and the electric field strength are maximal here. But it is necessary to take into consideration that about a half of this power enters initially into vibrational degrees of freedom of air molecules which relax then in translational degrees of freedom during finite time. Therefore convective transfer in boundary layer will "smear" volumetric heat source. Besides intense heating of the exposed electrodes owing to a relaxation of high kinetic energy of incident ions must be taken into consideration too.

The distribution of time-averaged volumetric force is rather complex as compared to Joule dissipation because of appreciably different distributions of instantaneous force during forward and backward strokes of DBD. The density distributions of horizontal and vertical components of the volumetric force time-averaged over the second cycle are shown on Fig. 8 in N/cm^3. Vectors of the total volumetric force with length proportional to its magnitude are shown here too. The color scales in both pictures are restricted from negative values for clearness of representation. Extreme negative values of both components are achieved on the electrode surface and equal -0.421 N/cm^3 for horizontal component and -0.273 N/cm^3 for vertical one. Negative values of both components are determined by volumetric force generated during the forward stroke of the discharge and concentrated near the electrode edge.

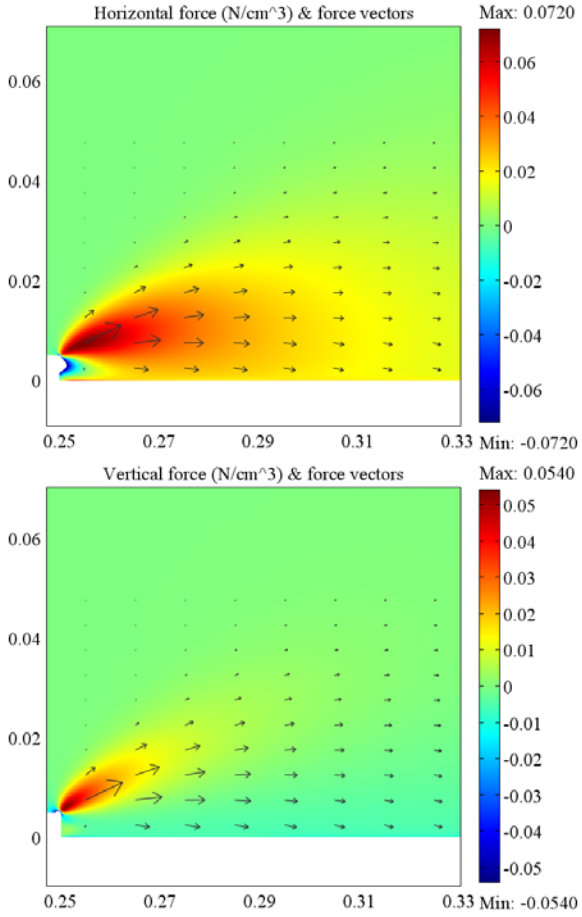


Fig. 8 Distributions of time-averaged volumetric horizontal (above) and vertical (below) force

One can see on Fig. 8 that the characteristic horizontal and vertical sizes of the area of remarkable horizontal force achieve 0.6 and 0.2 mm, respectively.

5 Conclusion

Presented results on numerical modeling of DBD multi-actuator system demonstrates that the discharge formation at upstream edges of the exposed electrodes can be excluded. It can be achieved by installation of additional buried electrodes beneath that edges with the same electric potential as the exposed ones. The estimation shows that fairly large coefficient of energy efficiency can be obtained in the considered scheme of the multi-actuator system. Further parametric research and optimization of this DBD system is of interest.

References

[1] Chernyshev S.L., Kiselev A.Ph., Kuryachii A.P. "Laminar flow control research at TsAGI: Past and present", *Progress in Aerospace Sciences*, Vol. 47, No. 3, 2011, pp. 169-185.

[2] Roth J.R., Sherman D.M., Wilkinson S.P. "Electrohydrodynamic flow control with a glow-discharge surface plasma", *AIAA Journal*, 2000. Vol. 38, No. 7, 2000, pp. 1166-1172.

[3] Mack L.M. "On the stability of the boundary layer on a transonic swept wing" AIAA Paper, No. 264, 1979, 11 p.

[4] Moreau E. "Airflow control by non-thermal plasma actuators", *J. Phys. D: Appl. Phys.*, Vol. 40, No. 3, 2007, pp. 605-636.

[5] Corke T.C., Post M.L., Orlov D.M. "SDBD plasma enhanced aerodynamics: concepts, optimization and applications", *Progress in Aerospace Sciences*, Vol. 43, No. 7-8, 2007, pp. 193-217.

[6] Boeuf J.P., Lagmich Y., Callegari Th., Pitchford L.C., Unfer Th. "New insights in the physics of DBD plasma actuators for flow control", *AIAA Paper*, No. 1376, 2008, 17 p.

[7] Bogdanov E.A., Kudryavtsev A.A., KuranovA.L., Kozlov I.E., Tkachenko T.L. "2D Simulation and scaling of DBD plasma actuator in air", *AIAA Paper*, No. 1377, 2008, 18 p.

[8] Soloviev V.R., Krivtsov V.M. "Features of a surface barrier discharge modeling", *AIAA Paper*, No. 842, 2009, 16 p.

[9] Likhanskii A., Semak V., Shneider M., Opaits D., Miles R., Macheret S. "The role of photoionization in the numerical modeling of the DBD plasma actuator", *AIAA Paper*, No. 841, 2009, 14 p.

[10] Huang P.G., Shang J.S., Stanfield S.A. "Periodic electrodynamic field of dielectric barrier discharge", *AIAA Paper*, No. 961, 2010, 12 p.

[11] Mamunuru M., Kortshagen U., Ernie D., Simon T. "Plasma actuator simulation: Force contours and dielectric charging characteristics", *AIAA Paper*, No. 1221, 2010, 9 p.

[12] Enloe C. L., Font G. I., McLaughlin T. E., Orlov D. "Surface potential and longitudinal electric field measurements in the aerodynamic plasma actuator", *AIAA Journal*, Vol. 46, No. 11, 2008, pp. 2730-2740.

[13] Font G.I., Enloe C.L., Newcomb J.Y., Teague A. L., Vasso A.R., McLaughlin T.E. "Effects of oxygen content on the behavior of the dielectric barrier discharge aerodynamic plasma actuator", *AIAA Paper*, No. 545, 2010, 16 p.

[14] Roth J.R., Madhan R.C.M., Yadav M., Rahel J., Wilkinson S.P. "Flow field measurements of paraelectric, peristaltic, and combined plasma actuators based on the one atmosphere uniform glow discharge plasma (OAUGDP^{TM})", *AIAA Paper*, No. 845, 2004, 11 p.

[15] Thomas F.O., Corke T.C., Iqbal M., Kozlov A., Schatzman D. "Optimization of dielectric barrier discharge plasma actuators for active aerodynamic flow control", *AIAA Journal*, Vol. 47, No. 9, 2009, pp. 2169-2178.

[16] Leonov S., Yarantsev D., Kuryachii A., Yuriev A. "Study of friction and separation control by surface plasma", *AIAA Paper*, No. 512, 2004, 8 p.

[17] Zienkiewicz O.C., Emson C., Bettess P. "A novel boundary infinite element", *Int. J. Num. Meth. Eng.*, Vol. 19, No. 3, 1983, pp. 393-404.

[18] Kuryachii A.P., Rusyanov D.A., Skvortsov V.V. "Features of numerical modeling of a dielectric barrier discharge", *TsAGI Sci. Journal*, Vol. 42, No. 1, 2011, pp. 37-52.

[19] Golubovskii Yu.B., Maiorov V.A., Behnke J., Behnke J.F. "Influence of interaction between charged particles and dielectric surface over a homogeneous barrier discharge in nitrogen", J. Phys. D: Appl. Phys., Vol. 35, 2002, pp. 751-761.

[20] Enloe C.L., McLaughlin T.E., Gregory J.W., Medina R.A., Miller W.S. "Surface potential and electric field structure in the aerodynamic plasma actuator", *AIAA Paper*, No. 1103, 2008, 11 p.

[21] Porter C.O., Baughn J.W., McLaughlin T.E., Enloe C.L., Font G.I. "Temporal force measurements on an aerodynamic plasma actuator", *AIAA Paper*, No. 104, 2006, 15 p.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Re-Analysis Procedure for Substructured Dynamics Finite Element Models

F. Macri and S. Ricci
Politecnico of Milan, Italy

Keywords: Helicopter Vibration, Optimization, Craig-Bampton, Re-analysis

Abstract

This work deals with Finite Element Model reduction, in particular with substructured models via the Craig-Bampton method. One of the drawback of adopting model substructuring and reduction during structural optmization is the need to compute again all transformation matrices as a consequence of a change in some design variables. This could be very expensive computationally.

This work analyzes several techniques that allows to update transformation matrices in a faster way. Those techniques are compared each other and with Craig-Bampton reduction, in order to compare their accuracy in terms of capability of natural frequency calculation and reconstruction of mode shapes.

1 Introduction

The helicopter is a quite complex machine whose optimal design involves several disciplines such as aerodynamics, structural dynamics, controls, etc.

For these reasons the problem of helicopter design and optimization, from a multidisciplinary point of view, has received significant attention in past decades and the availability of high level tools, able to merge and take into account of different and interactive disciplines, appears as a must.

This is especially true for what concerns the problem of the reduction of the vibration levels inside the helicopter cabin, still representing one of the most important design targets.

As a basis of a more general framework for helicopter's structural optimization, the proposed procedure is based on the structure decomposition technique: the dynamics behavior of each helicopter component is defined by a set of few physical parameters, such as panel thickness, beam cross section dimensions, etc...

The main advantage of the substructuring approach during structural optimization is that the changes in one parameter is confined in a well defined reduced part of the whole structure. Another obvious advantage is the lowering of the set of parameters defining the dynamics of each subsystem.

The main difficulty, on the other hand, is to define how to transfer the information across the different structural components.

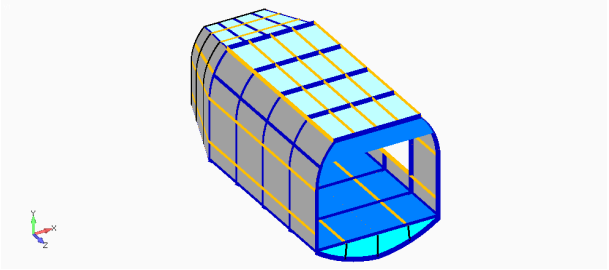


Figure 1: Example of substructure: FE model of an helicopter fuselage part

2 The Craig-Bampton Method

The solution of the interface problem comes out directly from the physics of the problem: each part is connected to the others in a very low number of point, so that it is easy to define and solve the interface problem. The substructuring is achieved with a Fixed Interface Component Mode Synthesis technique, the Craig-Bampton reduction method.

Craig-Bampton reduction method represents a very well known approach that allows to solve each substructure alone, and then, imposing the interface force and displacement, the dynamics of the complete system [5]. For each substructure, a reduced order model is computed, then the model is assembled. Craig-Bampton method is one of the reduced basis approaches, that achieve model reduction by employing a few preselected basis vectors which span the solution space of the approximation.

The Craig-Bampton scheme imply the partitioning of substructures matrices in 4 parts: one containing the terms relatives to the effects of the boundary Degrees of Freedom (DOFs) to boundary DOFs, X_{bb} (where X is a general FE model matrix), one containing terms relatives to effect of boundary DOFs to internal DOFs, X_{bi}, (which, transposed, contains terms relatives to the effect of internal DOFs to boundary DOFs), and the last containing terms relatives to the effects of internal DOFs to internal DOFs, X_{ii} .

The reduction basis is obtained by the use of fixed interface normal modes and constraint modes.Fixed interface normal modes are obtained solving the subsequent eigenvalue problem (one for each substructure, N_c is the total number of substructure):

$$\begin{aligned} &\left(-\omega_{j_c}^2 \left[M_{ii}\right]_c + \left[K_{ii}\right]_c\right) \left\{\phi_j\right\}_c = \{0\} \\ &c = 1 : N_c \end{aligned} \tag{1}$$

Consequently, the part of the reduction basis relative to fixed interface normal modes is:

$$\left[\Phi\right]_c = \begin{bmatrix} \{\phi_1\}_c & \{\phi_2\}_c & \cdots & \{\phi_N\}_c \\ \{0\}_c & \{0\}_c & \cdots & \{0\}_c \end{bmatrix} \tag{2}$$

They give informations relative to the dynamics of interior DOF.

Instead, constraint modes give informations relatives to the dynamics of the part consequent to motion of the boundary nodes, forces to the interface and rigid body motion.

They could be deduced from the assignment of unitary forces applied to boundary nodes, and they come out directly from the set of equations:

$$\begin{bmatrix} \left[K_{ii}\right]_c & \left[K_{ib}\right]_c \\ \left[K_{bi}\right]_c & \left[K_{bb}\right]_c \end{bmatrix} \begin{bmatrix} \left[\Psi_{ii}\right]_c \\ \left[I_{bb}\right]_c \end{bmatrix} = \begin{bmatrix} \left[0_{ib}\right]_c \\ \left[R_{bb}\right]_c \end{bmatrix} \tag{3}$$

So, the part of the reduction base relative to constraint mode is:

$$\begin{bmatrix} \left[\Psi_{ii}\right]_c \\ \left[I_{bb}\right]_c \end{bmatrix} = \begin{bmatrix} \left[K_{ii}\right]_c^{-1} \left[K_{ib}\right]_c \\ \left[I_{bb}\right]_c \end{bmatrix} \tag{4}$$

Obviously, when reduced models are used during an optimization procedure, the models as well as the responses should be computed several times during the iterations till convergence. So, the drawback of this procedure is that the computation of the reduction base, considering that it should be done for each iteration, could be very computationally demanding. This fact, depending on the number of iterations and how much the original structure is changed, could reduce the convenience in using substructuring.

2.1 Updating Fixed Interface Normal Modes by the Enriched Craig-Bampton Approach

Due to the possible high computational costs related to the updating of Craig-Bampton method (CB method), an alternative approach for model reducing has been studied, that requires a partial updating of the original component models and transformation matrices without a complete (exact) recalculation, reducing the total computational costs.

For what concerns fixed interface normal modes, their updating consists of an extension of the previously computed basis, adding a term related to the effect of the variation on the fixed interface normal modes; this procedure is called *Enriched Craig-Bampton method* [1], [6]. Fixed interface normal modes of the modified structure derives from the solution of the subsequent

eigenvalue problem, subscript c is omitted for readability:

$$\left(-\omega_j^2\left(\left[M_{ii}^1\right]\right)+\left(\left[K_{ii}^1\right]\right)\right)\left\{\phi_j^1\right\}=\{0\} \quad (5)$$

$$\left[M_{ii}^1\right]=[M_{ii}]+[\Delta M_{ii}] \quad (6)$$

$$\left[K_{ii}^1\right]=[K_{ii}]+[\Delta K_{ii}] \quad (7)$$

where the terms ΔM_{ii} and ΔM_{ii} are the variation in stiffness and mass matrices due to modification of the design parameters and the superscript 1 of ϕ_j^1 means that the eigenvector is relative to the new configuration. The set of equation could be rearranged in the following way:

$$\left(-\omega_{j_1}^2\left[M_{ii}\right]+\left[K_{ii}\right]\right)\left\{\phi_j^1\right\}=\left\{f_\Delta(\omega_{j_1})\right\} \quad (8)$$

$$\left\{f_\Delta(\omega_{j_1})\right\}=\left(-\omega_{j_1}^2\left[\Delta M_{ii}\right]+\left[\Delta K_{ii}\right]\right)\left\{\phi_j^1\right\} \quad (9)$$

where the f_Δ are the forces acting to the initial substructure due to the design modifications.

Those forces are used for defining a correction to the displacement field which can be employed to update the initial normal mode set.Assuming that inertia forces are negligible compared to the resisting elastic forces, the matrix containing the displacement correction is computed in the following way:

$$[R_L]^c=\left[\{f_\Delta(\omega_1)\}^c \quad \dots \quad \{f_\Delta(\omega_N)\}^c\right] \quad (10)$$

$$[R_D]^c=[K_{ii}]^{c^{-1}}[R_L]^c \quad (11)$$

The updated eigenvectors are not available, so the matrix containing residual forces could not be computed exactly, but they could be approximated using the eigenvectors computed previously.

$$\left[\tilde{R}_D\right]^c=[K_{ii}]^{c^{-1}}\left[\tilde{R}_L\right]^c \quad (12)$$

$$\left[\tilde{R}_L\right]^c=\left[\left\{\tilde{f}_\Delta(\omega_1)\right\}^c \quad \dots \quad \left\{\tilde{f}_\Delta(\omega_N)\right\}^c\right] \quad (13)$$

$$\left\{\tilde{f}_\Delta(\omega_{j_1})\right\}=\left(-\omega_{j_1}^2\left[\Delta M_{ii}\right]+\left[\Delta K_{ii}\right]\right)\left\{\phi_j^0\right\} \quad (14)$$

The extended fixed interface normal mode set assumes the form:

$$[\phi_{ex}]=\begin{bmatrix}[\phi_i]^c & \left[\tilde{R}_D\right]^c\\ [0]^c & [0]^c\end{bmatrix} \quad (15)$$

Summarizing, the method is based on the following two hypothesis: the negligibility of inertia forces with respect to elastic forces, and the capacity of residual forces due to original eigenvectors to approximate actual configuration residual forces.

The procedure makes use of Singular Value Decomposition for reconditioning matrix $\tilde{R}_L$.

3 Updating Constraint Modes by the Combined Approximation Procedure

For what concerns constraint modes, the following equations system derives from a modification of design parameters (superscript c is omitted for readability):

$$\begin{bmatrix}\left[K_{ii}^1\right]_c & \left[K_{ib}^1\right]_c\\ \left[K_{bi}^1\right]_c & \left[K_{bb}^1\right]_c\end{bmatrix}\begin{bmatrix}\left[\Psi_{ii}^1\right]_c\\ \left[I_{bb}\right]_c\end{bmatrix}=\begin{bmatrix}\left[0_{ib}\right]_c\\ \left[R_{bb}^1\right]_c\end{bmatrix} \quad (16)$$

$$\left[K_{ii}^1\right]=[K_{ii}]+[\Delta K_{ii}] \quad (17)$$

$$\left[K_{ib}^1\right]=[K_{ib}]+[\Delta K_{ib}] \quad (18)$$

$$\left[K_{bb}^1\right]=[K_{bb}]+[\Delta K_{bb}] \quad (19)$$

$$\left[\Psi_{ib}^1\right]=[\Psi_{ib}]+[\Delta\Psi_{ib}] \quad (20)$$

$$\left[R_{bb}^1\right]=[R_{bb}]+[\Delta R_{bb}] \quad (21)$$

Consequently, due to the modification of stiffness matrix and reaction forces, the changes in constraint modes base are obtained solving the following system of equations:

$$\begin{aligned}&\left([I_{bb}]+[K_{ii}]^{-1}[\Delta K_{ii}]\right)[\Delta\Psi_{ib}]+\\&+[K_{ii}]^{-1}\left([\Delta K_{ii}]\left[\Delta\Psi_{ib}^0\right]+[\Delta K_{ib}]\right)\end{aligned} \quad (22)$$

Using the binomial series expansion for the previous equation:

$$([I]+[A])\approx\left([I]-[A]+[A]^2-\dots\right) \quad (23)$$

the correction of the constraint modes base takes the form:

$$\begin{aligned}[\Delta\Psi_{ib}]&\approx\left([I]-[B]+[B]^2-\dots\right)[\Delta r_1]=\\&=[\Delta r_1]+[\Delta r_2]+\dots\end{aligned} \quad (24)$$

where the inserted terms are:

$$[\Delta r_1] = -[K_{ii}]^{-1}\left([\Delta K_{ii}]\left[\Delta\Psi^0_{ib}\right] + [\Delta K_{ib}]\right) \quad (25)$$

$$[\Delta r_k] = -[B]\,[\Delta r_{k-1}] \quad (26)$$

$$[B] = [K_{ii}]^{-1}\,[\Delta K_{ii}] \quad (27)$$

The Combined Approximations approach exploits the idea of conditioning the binomial series of each residual constraint mode, in order to banning them to get high value and preserving the good conditioning of the problem [4].

So, each residual mode is approximated in the space spanned by the vectors extracted from the series expansion

$$\{\Delta r_1\}_t = -[K_{ii}]^{-1}\{R\}_t \quad (28)$$

$$\{\Delta r_k\}_t = -[B]\{\Delta r_{k-1}\}_t \quad (29)$$

in which:

$$[R] = \left[\{R\}_1 \quad \dots \quad \{R\}_{N_b}\right] = \\ = [\Delta K_{ii}]\left[\Delta\Psi^0_{ib}\right] + [\Delta K_{ib}] \quad (30)$$

Storing residual modes in the matrix H_t:

$$[H_t] = \left[\{\Delta r_1\}_t \quad \dots \quad \{\Delta r_{N_b}\}_t\right] \quad (31)$$

the correction to the t-th constraint mode is approximated as:

$$\{\Delta\Psi_{ib}\}_t \approx \{\Delta r_1\}_t\, y_{t,1} + \dots + \{\Delta r_{N_b}\}_t\, y_{t,N_b} \quad (32)$$

and coefficients $y_{t,k}$ could be obtained solving the system:

$$[H_t]^T\left([K_{ii} + \Delta K_{ii}]\right)[H_t]\{y_t\} = \\ = [H_t]^T\left(-[\Delta K_{ii}]\left\{\Psi^0_{ib}\right\}_t - \{\Delta\Psi_{ib}\}_t\right) \quad (33)$$

The updating of the constraint modes do not provide terms in addiction, as explained in the final form of them.

$$[\Delta\Psi_{ib}] = [\{\Delta\Psi_{ib}\}_1 \dots \{\Delta\Psi_{ib}\}_{Ns}] \quad (34)$$

$$[\Psi] \approx \begin{bmatrix}\left[\Psi^0_{ib}\right] + [\Delta\Psi_{ib}] \\ [I_{bb}]\end{bmatrix} \quad (35)$$

3.1 Evaluation of computational costs for Combined Approximations Approach

The computation of the exact correction of the Constraint mode base requires (without considering matrices sparsity and symmetry) a number of Floating Point Operation (FLOP) of the order of N_i^3; instead, the computational costs deriving from combined Approximations Approach are depending either to the number of internal DOF, the number of boundary DOF and the number of terms of the series expansion.

Starting from the fact that a few terms of the series expansion are needed for a good approximations of the constraint mode correction, it could be seen that in those case in which the ratio between boundary nodes number and interior nodes number is high (near to one), it is convenient to compute the exact correction (equation (1)) instead of the combined approximations procedure.

4 Procedure for obtaining matrices from new combination of parameters

Reanalysis techniques need a procedure for evaluating mass and stiffness matrices after the variation of at least one of the design parameters.

A procedure for computing (faster than recompiling stiffness and mass matrices) matrix has been implemented. This procedure needs the use of a FEM solver (like MSC Nastran) only at the first step of the procedure, then it is possible to obtain updated matrices only with a multiplication of previously computed influence coefficients with the new set of design parameters.

The procedure considers a polynomial dependence of the (stiffness of mass) matrix terms X_{ij} to the parameters θ:

$$X_{ij} = A^1_{ij}\theta_1 + \dots + A^{np}_{ij}\theta_{n_p} + A^{n_p+1}_{ij}\theta_1^2 + \\ \dots + A^{3n_p}_{ij}\theta^3_{n_p} + A^{3n_p+1}_{ij}\theta_1\theta_2 + \dots \\ + A^N_{ij}\theta^3_{n_p-1}\theta^3_{n_p} + A^{N+1}_{ij} \quad (36)$$

where n_p is the number of design parameters, the matrix A^k_{ij} is the matrix of influence coeffi-

cients of the k-th term of the polynomial envelope of X_{ij}. The term A_{ij}^{n+1} is the constant term (not dependent to design parameters); $N+1$ is the total number of terms of the polynomial envelope and it is equal to:

$$N = \frac{3}{2}\left(n_p^2 + n_p\right) \tag{37}$$

Using a full factorial sampling, it is possible to obtain each influence coefficient, solving the following set of equations:

$$\{X_{ij}\} = [\Theta]\{A_{ij}\} \tag{38}$$

in which:

$$\{X_{ij}\} = \begin{Bmatrix} X_{ij}^1 \\ \vdots \\ X_{ij}^{N+1} \end{Bmatrix} \quad \{A_{ij}\} = \begin{Bmatrix} A_{ij}^1 \\ \vdots \\ A_{ij}^{N+1} \end{Bmatrix} \tag{39}$$

$$[\Theta] = \left[\{\theta_1\} \dots \{\theta_{n_p}\} \{\theta_1\}^2 \dots \{\theta_{n_p}\}^3 \{1\}\right] \tag{40}$$

where $\{\theta_q\}^k$ is a vector containing the k powers of a collection of randomly generated numbers used as θ_q design parameter and the dimension of Θ are $(N+1)\cdot(N+1)$, and X_{ij}^h is the term of the matrix evaluated with the h combination of randomly generated design parameters.

The influence coefficients of all the X_{ij} are obtained with the inverse of the matrix Θ, that could be computed only one time. In this way, for obtaining all the influence coefficients, the number of needed FLOP is

$$N_c = (N+1)\left((N+1)^2 + N_T\right) \tag{41}$$

where N_T is the total number of nonzero terms of the matrix $[X]$. In this way, with the knowledge of these influence coefficients, it is possible to obtain new mass and stiffness matrices terms only multiplying the new set of design parameters (and their powers) for the vector of influence coefficient, at the computational cost of $N_T(N+1)$ FLOP.

As an example, consider a symmetric matrix of 20000 DOF, with bandwidth of 9 and consider 3 design parameters for the substructure in exam: N is equal to 18, NT is equal to 100000, so the computational cost for obtaining all the influence coefficients is $19(361 + 100000) = 1906859$ FLOP. The convenience in using this procedure lies in the fact that FEM solver is used only one time, and then it is possible to compute all the needed matrices with the use of high level programming language, like MATLAB.

5 Results

In this section, the first part shows the results of mass and stiffness matrices reconstruction obtained with the MATLAB full factorial sampling procedure. The second part, instead, presents the comparisons between the several reduction procedure presented previously.

In order to compare the accuracy of the matrices reconstruction, a simple case test has been studied: normal modes of a simplified finite element model of an helicopter tail, showed in Fig. 2 have been compared.

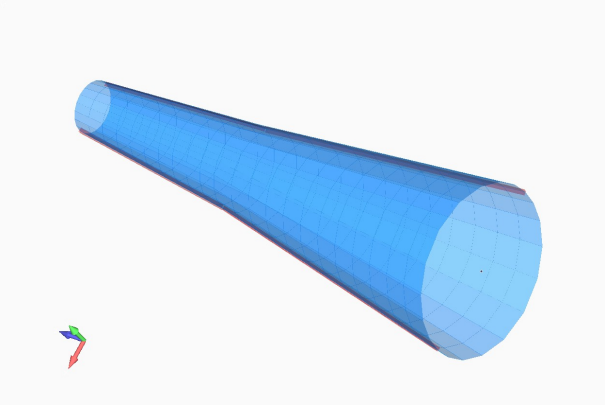


Figure 2: Simplified Finite Element Model of an heicopter tail

Tail natural frequency deriving from the exact assembly of mass and stiffness matrices, computed with FEM solver MSC.Nastran, have been compared with natural frequencies deriving from matrices reconstructed with the previously explained procedure.

To avoid differences deriving from different numerical packages, only mass and stiffness matrices have been computed with FEM solver.

Physical parameters, such as beam flange lenght or panel thickness, have been changed in order to verify matrices reconstruction in the range of physical parameters. Figures 3 - 4 tes-

tify the accuracy of the procedure, in terms of robustness with respect to the changes in the design parameters.

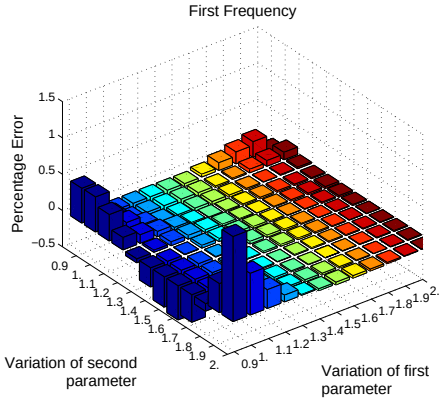


Figure 3: Percentage error on the first flexible mode frequency: variation of the parameter has to be considered as relative variation

For what concerns mode shapes reconstruction, results are almost the same: reconstructed eigenvectors match very well eigenvectors deriving from exactly computed matrices.

The re-analysis procedures explained before have been compared with the same FE model used for the previous comparison. Their accuracy, in terms of natural frequency and correlation between shape modes, has been analyzed for changing physical parameters.

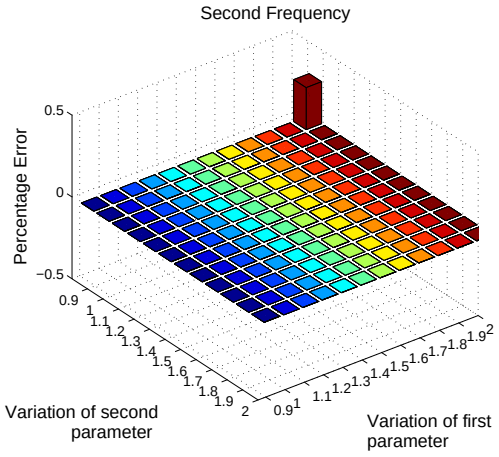
The reduction method studied are (first it is written the acronym used in the figures):

1. **CB**: Craig-Bampton method reduction

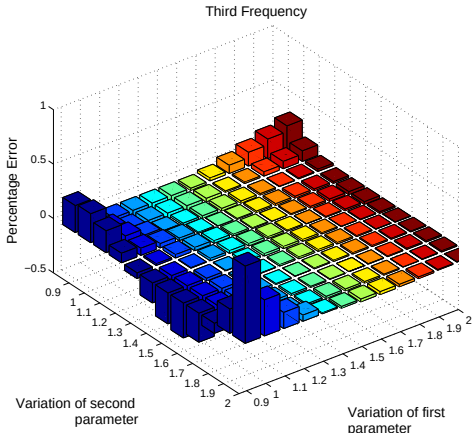
2. **ECB**: the procedure exploits the Enriched Criag-Bampton method for updating fixed interface normal modes; constrint modes are computed exactly, as in CB method

3. **CA**: the procedure computes exactly fixed interface normal modes, as in the CB method; instead, constraint modes are computed with the Combined Approximations Approach

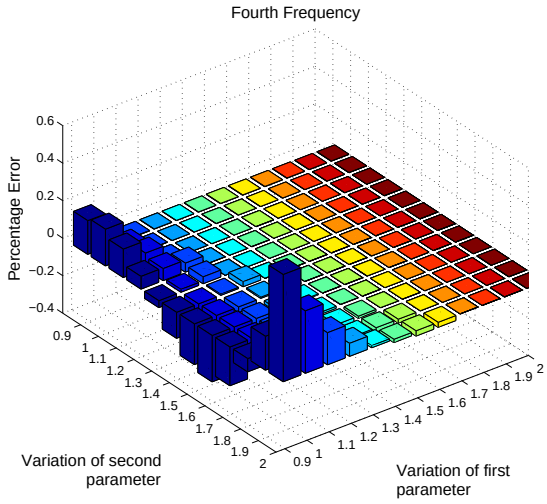
4. **ECBCA**: the procedure exploit both the reanalysis techniques explained before



(a)

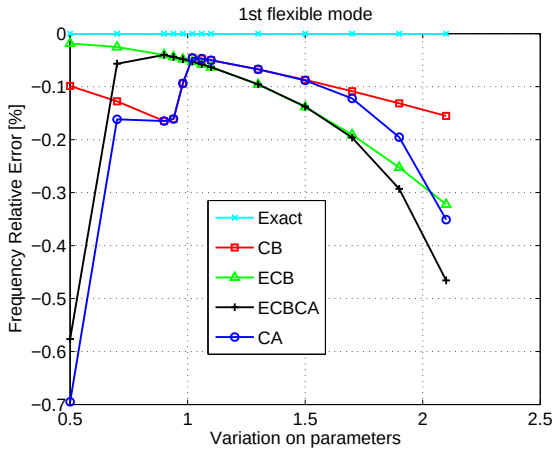


(b)

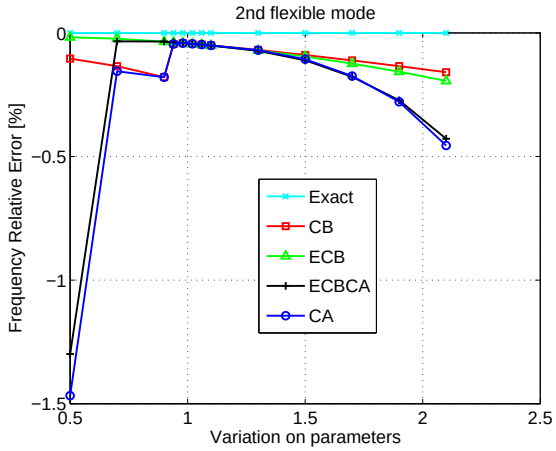


(c)

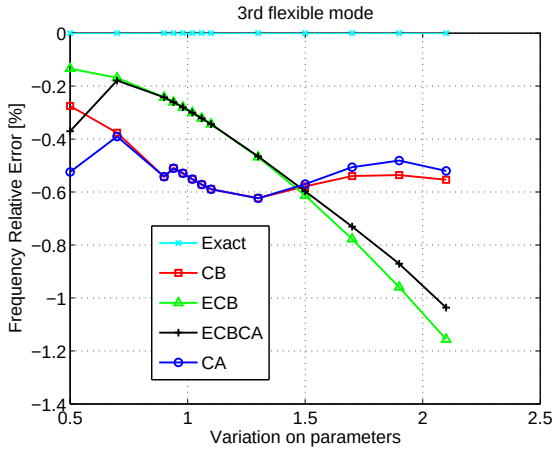
Figure 4: Percentage error on natural frequencies: a) second flexible, b) third flexible, c) fourth flexible mode



(a)

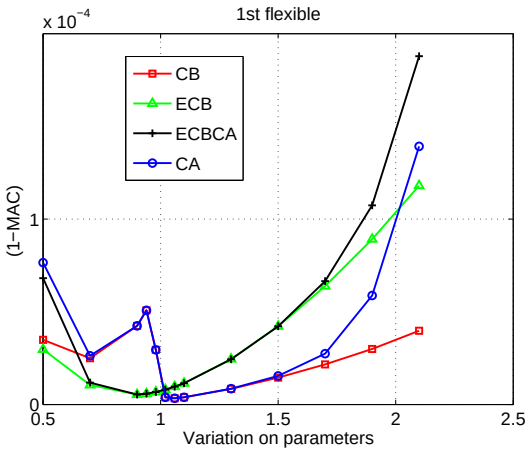


(b)

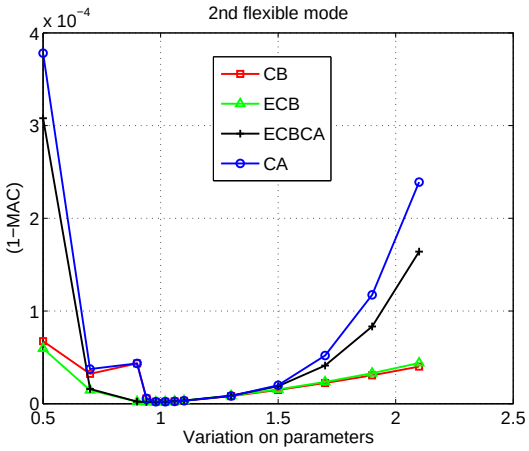


(c)

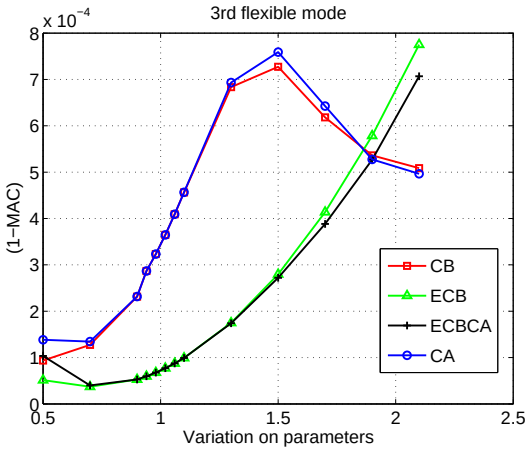
Figure 5: Percentage error on natural frequencies: a) first flexible, b) second flexible, c) third flexible mode



(a)



(b)



(c)

Figure 6: Error in the correlation between mode shapes: a) first flexible, b) second flexible, c) third flexible mode

In Fig. 5, percentage relative error on natural frequency, Eq. 42, are shown.

$$\epsilon = 100\frac{\Delta f}{f_{exact}} \tag{42}$$

In Fig. 6, error of MAC (Mode Accuracy Criterion) between reduction method eigenvector and exact eigenvector:

$$MAC = \frac{\{\phi_k\}^T \{\theta_k\}}{\left(\{\phi_k\}^T \{\phi_k\}\right)\left(\{\theta_k\}^T \{\theta_k\}\right)} \tag{43}$$

in which $\{\phi_k\}$ is the k-th eigenvector computed with one procedure and $\{\theta_k\}$ is the k-th eigenvector with the other procedure.

6 Conclusions

Results have testified the accuracy of reconstrucuted matrices by the computed influence coefficients. Natural frequency and mode shapes are almost the same.

The accuracy of reduced model performances by the updating procedures are shown in Fig. 5 and 6, so the procedure using Enriched Criag-Bampton and Combined Approximations Approach (ECBCA) could be used in place of a new Craig-Bampton reduction, saving computational time.

A final remarks should be done on the reconstruction matrices procedure: in the case of great model dimensions, the computation of influence coefficients could be very time consuming, in terms of model preparing and in terms of number of analysis. For these reasons, tt would be convenient only in those cases in which it is expected a great number of iterations during the optimization process.

References

[1] Masson G., Ait Brik B., Cogan S. and Bouhaddi N., "Component mode synthesis (cms) based on an enritched ritz approach for efficient structural optimization", *Journal of Sound and Vibration*, Vol. 296, No. 4-5, 2006, pp. 845-860.

[2] Bouazzouni A., "Selecting a ritz basis for the reanalysis of the freqeuncy response functions of modified structures", *Journal of Sound and Vibration*, Vol. 198, No. 2, 1997, pp. 309-322.

[3] Bindolino G., Mantegazza P. and Ricci S., "Integrated Servostructural Optimization in the Design of Aerospace Systems", *Journal of Aircraft*, Vol. 36, No. 1, 1999, pp. 167-175.

[4] Kirsch U., *Design Oriented Analysis of structures, A unified approach*, Vol. 95 of *Solid Mechanics and its Applications.*, Springer, 2008.

[5] Rixen D. and Gerardin M., *Mechanical Vibrations and Structural Dynamics*, Wiley, 1993

[6] D. Ackay Perdahcioglu, "Optimizing the Dynamic Behaviour of Structure", Ph.D. Dissertation, Department of Applied Mechanics, University of Twente, Country, 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A procedure for establishing an approximate helicopter performance mode

Magnus Sethson
Fluid and Mechatronic Systems, Department of Management and Engineering, The Institute of Technology, Linköping University, SE-58183 Linköping, Sweden

Keywords: rotorcraft,helicopter,rotor,mathematical model, UAV

Abstract

The main performance characteristics of a helicopter is its fuel consumption. This paper reports the current state in an ongoing project seeking a fast and approximative implementation of mathematical models for helicopter performance predictions, primarily in terms of fuel consumption. The model is limited to a subset of the normal operation modes, only including hover, climb and forward. Also, the focus is on smaller rotors, i.e. sub-scale rotors in the range of 1.5m *to* 4.0m *in diameter as used in unmanned airborne vehicles (VTOL-UAV). The presented work consists of three parts; the helicopter performance mathematical model and the numerical optimization method used for dealing with its implicit nature, the approximation grid scheme and the simulation of flight paths. The overall objective of the project is to establish an computational efficient approximative model usable for onboard condition monitoring of the helicopter performance and fuel consumption.*

1 Introduction

The main performance characteristics of a helicopter on patrol is its fuel consumption. This work present a classical model of an lifting rotor as configured in most of todays helicopters. The model follows closely textbook examples like [6], [1] and [4]. Most of the work presented here follows the model outlined in the book of Leishman, with some modifications. The modeling is focused on sub-scale rotors in the range of 1.5m to 4.0m in diameter as often found in unmanned vehicles, VTOL-UAV. Such vehicles need a fuel monitoring functionality supporting their ability to return home or finding a safe landing place in case of command linkage failure. The presented work addresses that fuel monitoring functionality. In normal operation, a fuel consumption calculation is performed beforehand by the operator before take-off using the ground control station software. Other checks are also performed like, way-point precision estimates, maximum banking angle checks and retardation distances. Predictions are possible based upon recent weather forecasts. Newly uploaded missions are also subject to similar checks before uploading, possible less accurate within the timeframe of operation. Different tools and techniques for analysis mathematical rotorcraft models exist and have been studied, [9][2], where the procedure is analogous to this work even if the focus is more to control engineering and stability analysis. There are also examples of scientific work that takes the approximative approach on a more thorough level, [7].

Consider the following example scenario (see Fig. 1): An VTOL-UAV (1) take-off from a land based helipad and starts its (2) route in a

pre-programmed way-point based mission. (3) The command uplink fails. The UAV performs the link recover (4) climb flight pattern and also starts to evaluate return path targeting different stationary or moving helipads in the area. If all evaluations fail it may turn into an escape route mission for self-termination. During the planning of the flight paths, the fuel consumption and the reachability of the target waypoint is essential. The onboard computer computational power does not allow for any detailed calculations or extensive path-planning simulations. Still, it is desirable that the fuel prediction model on ground and in the vehicle is consistent. This is achieved by having an in-air sufficiently good approximation of the model used on ground.

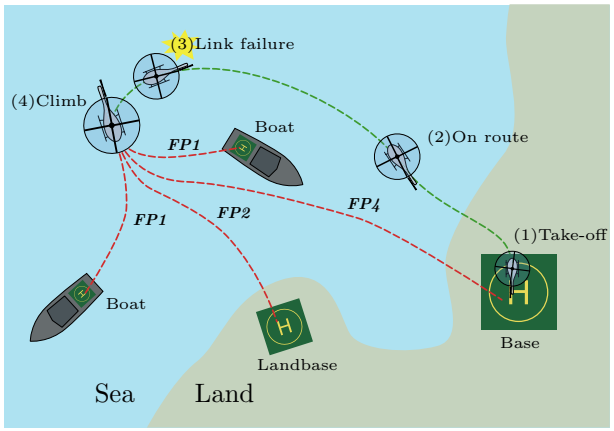


Figure 1: The path-planning scenario after a command link failure.

The work fall into three natural parts, see figure 2; The helicopter performance model, the approximation scheme and the path-planning fuel consumption prediction.

The first part of the work, the helicopter performance model, is based on classical assumptions in helicopter theory. The model uses momentum theory and blade element analysis as described in the well-known textbooks of Leishman and Bramwell. The model is for natural reasons implicit in that it makes no assumption of the quantity of induced velocity through the rotor disc. A linear inlet model of the induced velocity has been used, forming the shape of the

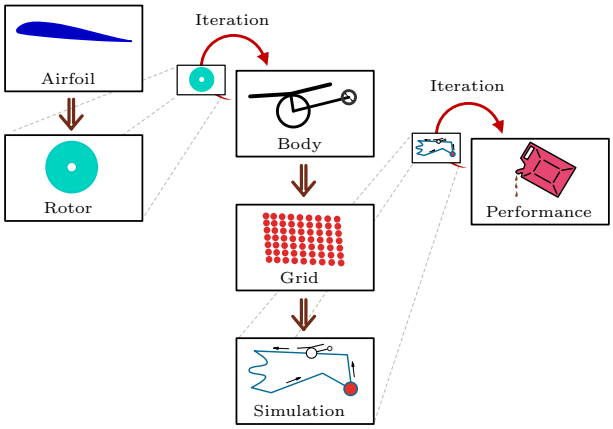


Figure 2: The outline of the procedure to establish an approximative fuel consumption model.

inlet flow and parameterized by λ. The model is extended from a single rotor to include a tail rotor and a spherical body.

Secondly, an approximation is built up by precalculated model performance values indexed by an operating state vector like: $[m, V, \rho, V_c]^T$, that is helicopter mass, forward flight speed, air density and climb speed. The interpolation routine then provides an estimated power required P. Finally, the in-air prediction routine and model storage is out-lined.

2 Rotorcraft model

The rotorcraft model consist of a main rotor bale element model, an approximative spheric body shape and a tail rotor, see figure 3. This work focus on the generality of the model structure and the procedure to calculate performance measures for preliminary design parameter studies. Therefore the work has strive to accomplish a working set of numerical solvers and generality in the model formulation instead of accomplish a good match to measurement data. A number of assumptions have been made on the way to that goal: First, no advanced inlet flow model of the main rotor disk have been used. Instead a constant average inflow model has been used. Secondly, the body shape approximation is a sphere and only provide a force balance for the forward flight drag. Pitch or roll torques are neglected.

Thirdly, the tail rotor assumes the same scheme as the main rotor but always in hover condition. No forward airspeed or main rotor slipstream is affecting the tail rotor. Still, induced power, profile drag loss, body losses and tail rotor losses together with required climb power is included. In that way the model includes all five terms in the performance model, as described by Leishman, [4]:

$$\underbrace{C_P}_{\text{Required}} = \underbrace{\frac{\kappa {C_T}^2}{2\sqrt{\lambda^2+\mu^2}}}_{\text{Induced}} + \underbrace{\frac{\sigma C_{d0}}{8}(1+K\mu^2)}_{\text{Profile}} + \underbrace{\frac{1}{2}\left(\frac{f}{A}\right)\mu^3}_{\text{Body}} + \underbrace{\lambda_c C_W}_{\text{Climb}} + \underbrace{C_{P_{TR}}}_{\text{Tail}} \tag{1}$$

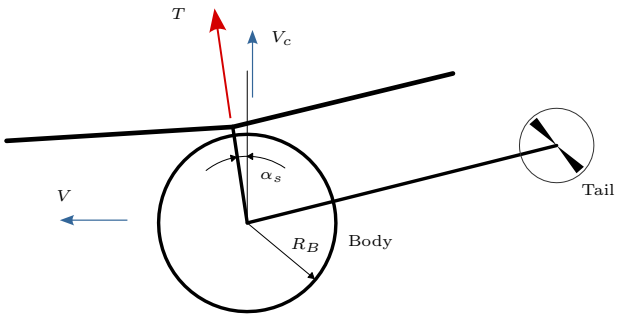


Figure 3: The idealized model of a helicopter used in the performance mathematical model in this work.

2.1 Main rotor model

The main rotor is studied as a separate element of propulsion. No airstream analysis is performed around the rotor and there are not interaction with tail or body. The applied assumptions are classical and described in many textbooks on helicopter performance analysis. Momentum theory and blade element analysis are used. The induced velocity v_i is required to define the whole of this model and it also reveals the implicit nature of rotorcraft modeling by the mutual relation between trust T and induced velocity v_i. The rotor is defined by its radius R and the air is represented by its density ρ.

$$v_i = \frac{1}{B}\sqrt{\frac{1}{2}\frac{T}{\rho\pi R^2}} \tag{2}$$

The tip-loss effect has been introduced by the compensation factor B. The applied blade element analysis and momentum theory is straight forward, however, three different offset linear inflow models has also been studied so far in this work but has been excluded from presentation in this paper.

2.1.1 Blade element model

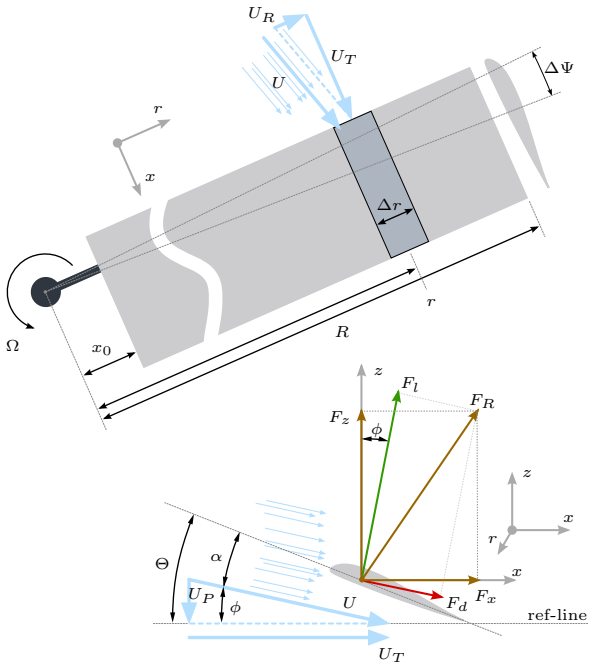


Figure 4: Definition of element quantitates.

The blade element method takes the local angle-of-attack, $\alpha(r,\Psi)$, into the local forces applied to the element. In this work limited to lift dL and drag dD. This results in:

$$\begin{aligned} dL =& L(\alpha(r,\Psi),M,U) \approx \\ & \frac{1}{2}\rho U^2 c C_l(\alpha(r,\Psi)) \\ dD =& D(\alpha(r,\Psi),M,U) \approx \\ & \frac{1}{2}\rho U^2 c C_d(\alpha(r,\Psi)) \end{aligned} \tag{3}$$

Often the airfoil coefficients are approximated like $C_l(\alpha) \approx C_l\alpha$. In this work, most linear simplifications has not been introduced as an readiness for the future progress of the model within project. U is the local airspeed and c is local blade chord. The effects of near speed of sound is neglected and the air is only represented by its density, ρ.

For the helicopter rotor-head the local $\alpha(r, \Psi)$ then becomes:

$$\alpha(r,\Psi) = (\Theta - \phi) = [\Theta_0 + \Theta_1 r - A_1 \cos\Psi - B_1 \sin\Psi] - \left[\tan^{-1}\frac{U_P}{U_T}\right] \quad (4)$$

Again, the common approximation $\phi = U_P/U_T$ is not introduced. The local airspeeds U_P and U_T is defined in Fig. 4. Also, following classical textbooks on helicopter modeling, the local forces dL and dD are replaced by the rotor disc lifting force dF_z and the counter rotating drag force dF_x. From this follows that these forces are defined as:

$$dF_z = dL\cos\phi - dD\sin\phi \quad (5)$$
$$dF_x = dL\sin\phi - dD\cos\phi \quad (6)$$

From that, given the number of blades N_b, follows the elements contribution to rotor lift (dT), torque (dQ) and power (dP) requirement:

$$dT = N_b dF_z \quad (7)$$
$$dQ = N_b dF_x r \quad (8)$$
$$dP = N_b dF_x \Omega r \quad (9)$$

The overall properties of the rotor disk is given by summing up all element contributions numerically. In this work the local radial $\Delta r_{i,j}$ length and angular coverage $\Delta\Psi_{i,j}$ has been used. A non equal discretization of the rotor disk may be used.

$$T = \frac{N_b}{2\pi}\sum_{i=1}^{N_\Psi}\sum_{j=1}^{N_R}\Delta\Psi_{i,j}\Delta r_{i,j} dF_{z_{i,j}} \quad (10)$$
$$Q = \frac{N_b}{2\pi}\sum_{i=1}^{N_\Psi}\sum_{j=1}^{N_R}\Delta\Psi_{i,j}\Delta r_{i,j} dF_{x_{i,j}} r \quad (11)$$
$$P = \Omega\frac{N_b}{2\pi}\sum_{i=1}^{N_\Psi}\sum_{j=1}^{N_R}\Delta\Psi_{i,j}\Delta r_{i,j} dF_{x_{i,j}} r \quad (12)$$

Notice the quasi-static nature of these results, the summation is not covering an area but instead a number (N_Ψ) of positions along the Ψ-coordinate, hence the 2π factor in the denominator. These three quantities (T, Q, P) forms the main objective of the modeling so far in the project. Also, the local airspeeds $U_P(r, \Psi)$ and $U_T(r, \Psi)$ are needed:

$$U_P(r,\Psi) = V_c + v_i \quad (13)$$
$$U_T(r,\Psi) = \Omega r + V\sin\Psi \quad (14)$$
$$U_R(r,\Psi) = V\cos\Psi \quad (15)$$

They represent the current state of the airflow on the inlet side of the rotor. The radial airspeed U_R along the blade is neglected. However, several models has been presented to include this effect as an time dependent state, [3],[5].

2.1.2 Hover flight iterative scheme

The classical approach for solving this in quasi-static models is to assume an required lifting force T_{req}, use that to calculate the average inflow airspeed (v_i) into the disk and from that estimate the local airspeeds at each element. The sum of all local contributions to the forces on each blade, the overall lifting force may be calculated. This iterative scheme is repeated until the assumed $T^{[n-1]}$ and calculated $T^{[n]}$ lifting forces converge to the same value, that is minimizing:

$$\min_{\Theta_0}\left(T^{[n]} - T^{[n-1]}\right)^2 \quad n \in [1..N_{ih}] \quad T^{[0]} = T_{req} \quad (16)$$

The variable to minimize upon is Θ_0. In all of this work, the Simplex method of Nelder and Mead [8] is used as implemented in GNU Scientific Library, GSL[1]. Also making assumptions on the Lock number will include flapping and coning as well. It is not included here.

2.2 Body model

The body is highly simplified and approximated with sphere with a rough surface. In general the rotor-body interactions are a most delicate problem in helicopter analysis both in terms of vibrations, performance and noise. Reducing this comprehensive set of development and research areas into one single sphere might seem odd, but fills its purpose well in this application.

2.2.1 Body drag

The drag force F_B from the body is then defined by the drag coefficient $C_d \approx 0.47$ and the forward speed V.

$$F_B = \frac{\rho C_d A_{ref}}{2} V^2 \tag{17}$$

The reference area A_{ref} is selected as the circular area described by R_B in figure 3. This drag force is only balanced by the forward trust F_{Dx} generated by the main rotor and described as:

$$F_{Dx} = mg \sin \alpha_s \tag{18}$$

Where the mass m and gravity g contributes to the force component in the forward direction. See the definition of α_s in Fig. 3. Following the iterative scheme described in section 2.1.2 above we get that the required lifting force can be set to:

$$T_{req} = mg \cos \alpha_s \tag{19}$$

T_{req} is balanced by the lifting force T of the rotor, see Eq. 10.

[1] http://www.gnu.org/software/gsl

2.2.2 Forward flight iterative scheme

Given a forward velocity V the scheme for finding the power consumed by the main rotor, P, and the torque, Q is given by 12 and 11 above. By combining Eqs. 2, 10, 17, 18 and 19 we get a two-dimensional optimization problem like the following:

$$\begin{aligned} &\min_{(\Theta_0,\alpha_s)} (e_F) \\ &n \in [1..N_{is}] \\ &T^{[0]} = T_{req} \end{aligned} \tag{20}$$

Where:

$$e_F = \frac{{e_T}^2}{s_T} + \left(K_D \frac{e_D}{s_T} \right)^2 \tag{21}$$

$$e_T = T^{[n]} - T_{req} \tag{22}$$

$$e_D = F_B - F_{Dx} \tag{23}$$

$$s_T = T^{[n]} + T_{req} \tag{24}$$

The factor K_D is then chosen to roughly describe the ratio between the main and tail rotor trust in "worst case strained operation". In this work $K_D = 20$ have been used with good numerical result. K_D affects the convergence speed and the stability of the iterative scheme.

2.3 Tail rotor

The power consumed by the tail rotor is easily estimated once the torque of the main rotor (Eq. 11) is identified. The required trust from the tail rotor is given by:

$$T_{req,tail} = \frac{Q}{L_{bom}} \tag{25}$$

The length L_{bom} is between the axes of the main and tail rotors. It is assumed that the tail rotor operate in no wind. That is of course a significant approximation. However, for the overall power estimate it is acceptable. The hovering condition of operation of the tail rotor is solved in an scheme analogous to the one described above for the main rotor in section 2.1.2.

2.4 Example

The model has been evaluated for several helicopters available on the market. One of witch has been available for some verification is presented here along with some results on the hovering performance. It is the jet-turbine powered Lama SA315B scale model RC-helicopter from Vario Helicopter[2]. It has a MTOW of 27kg. The mathematical model include the main rotor and the tail rotor.



Figure 5: Lama SA315B scale model helicopter from Vario Helicopter used as example. Photo by author.

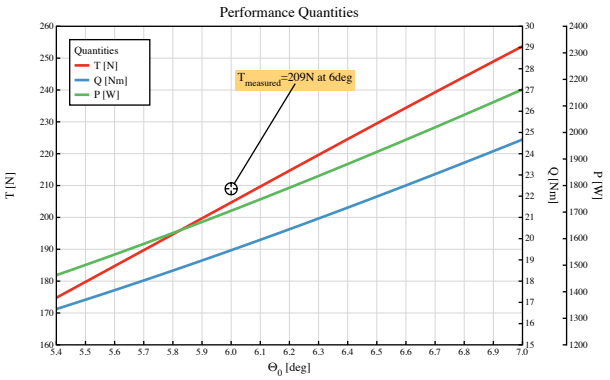


Figure 6: The diagram show the hovering performance, $V = 0$, at different collective pitch settings. Weight is 21.3kg. There is one measurement point, 209N at 6 degrees. The model lift (red line) slightly underestimates the lifting force, T.

3 Approximation grid

The described model in section 2 have been applied to a fixed grid of calculation points where each point is described by a 4-dimensional vector $[m, V, \rho, V_c]^T$. For each point, a number of quantities has been calculated and stored in a value vector, $[\Theta_0, \alpha_s, T, Q, P]^T$. Typically, a grid for one rotorcraft design have 10K to 200K grid points. The interpolation scheme in use is the simplest of them all; nearest-neighbor selection. The current state in the simulation model is used a search key in the 4D-matrix of grid points. The point with the shortest distance to the key-point is selected by making use of indices from truncating standard functions. No smoothing is applied. The search scheme then become almost

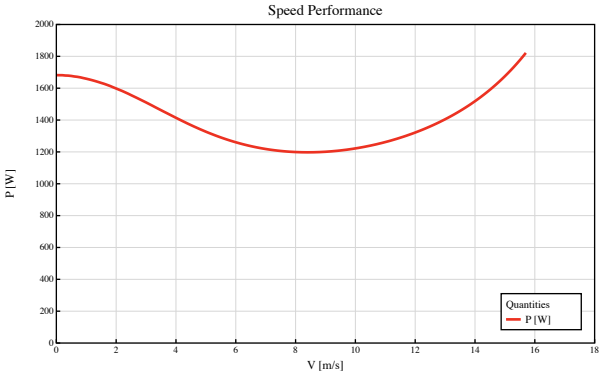


Figure 7: The diagram show the forward flight speed performance at a speed range up to $16m/s$. Total weight is 21.3kg. None of this has been verified.

[2]http://www.vario-helicopter.de

independent of the size of the grid. For example, the search index, I_V, for the velocity V is defined using the C-runtime library function *round()* as:

$$I_V = round\left(\frac{V - V_{min}}{V_{inc}}\right) \tag{26}$$

Any limits included int the rotorcraft model are included using a validity parameter stored along each grid point. Defined as 1.0 for a completely model correct point and less than 1.0 for grid points restricted by limits like; engine power limit (P_{max}), Θ_0 restrictions and main gear box torque limit (Q_{max}).

3.1 Example

An approximation grid for the RC-helicopter from section 2.4 above has been generated at different sizes. The numerical performance results can been seen here.

Size [Kpoint]	CPU-time [min]
12	11
34	42
192	124

Table 1: CPU time for different grid sizes on a Intel i7 3.4GHz CPU using a 8-core parallel implementation.

4 Fuel performance prediction

By making use of the approximative model one can predict the current power requirement and fuel consumption. To demonstrate and test this scheme, a simple flight path description have been used, where a series of waypoints are visited along with a requested velocity. The simulation updates the current position of the helicopter in a stepwise fashion without any dynamics involved. There are no acceleration or retardation zones close to the waypoints. This technique may be used for several applications. To demonstrate, the problem of calculating the fuel amount needed for a mission given the landing weight and flight path has been studied. Again an iterative scheme has been used for solving the task. Simulation is repeated using different settings on the take-off weight and the final landing weight is matched towards a decired weight set by the user. The iterative scheme minimizes the difference between the estimate of the landing weight $m_{land}^{(n)}$ and the desired landing weight m_{safe} by adjusting the take-off weight m_{to} which includes the fuel to be used. The minimizing problem looks like:

$$\min_{m_{to}} \left(m_{land}^{(n)} - m_{safe}\right)^2 \qquad n \in [1..N_{ss}] \tag{27}$$

4.1 Example

Again the example from section 2.4 is used. The problem of calculating the required fuel level before flight includes many of the possible application of the approximative model. An flight path according to Fig. 8 has been studied. The results for different grid sizes is given in table 2. All simulations are performed using a single-core implementation on an Intel i7 3.4GHz CPU. The time-step of the simulations are $\Delta t = 0.2s$. A simulation is done in a fraction of a second, typically $70ms$.

Size [Kpoint]	CPU [s]	Fuel [kg]
12	1.51	4.19
34	1.52	4.33
192	1.60	4.34

Table 2: Comparison between different grid sizes for the fuel estimation routine.

5 Conclusion

The procedure presented is part of an ongoing project. It's characteristics focus more on the inclusion of phenomenas in helicopter analysis than qualitative measures. The focus in the work has been to establish an framework in terms of a C++ implementation, partly parallel, of the described procedure. The implementation is also the focus on the future work. It

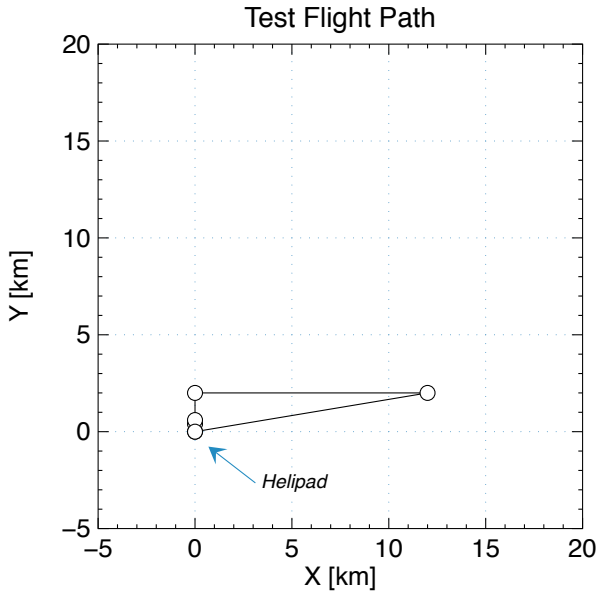


Figure 8: The test flight path used for studying the effect of grid sizes.

makes use GSL[3], STL[4] and C++ templates and is a very compact design of the procedure, only 2000 rows in size. From the tests performed during the development one can state that making use of this procedure using the approximative model reduces the simulation time for a flight path from about $90s$ to $70ms$. The cost for that is the establishment of an approximative grid model that takes about $2h$ to calculate in an parallel implementation. It can also be concluded that the use of iterative minimizing algorithms, especially the Simplex method, may be applicable for the solution of the induced inflow problem of rotorcraft rotor analysis, as outlined in section 2.2.2. The goal-function of such a minimizing task is subject to scaling parameters needed for a good numerical characteristics. Such parameters, exemplified by K_D, may have a interpretation in terms of helicopter performance operation measures.

[3]http://www.gnu.org/software/gsl
[4]http://www.sgi.com/tech/stl

References

[1] A R S Bramwell, George Taylor Sutton Done, and David Balmford. *Bramwell's helicopter dynamics*. American Institute of Aeronautics and Astronautics, 2 edition, 2001.

[2] Ronald A Hess. Analytical Assessment of Performance, Handling Qualities, and Added Dynamics in Rotorcraft Flight Control. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 39(1):262–271, January 2011.

[3] Kathryn B Hilbert. A Mathematical Model of the UH-60 Helicopter. *NASA, Ames Research Center*, pages 1–46, April 1984.

[4] J Gordon Leishman. *Principles of Helicopter Aerodynamics. 2nd.* Cambridge University Press, 2nd edition, April 2006.

[5] JR Majhi and R Ganguli. Helicopter blade flapping with and without small angle assumption in the presence of dynamic stall. *Applied Mathematical Modelling*, 34(12):3726–3740, 2010.

[6] Peter Rowland Payne. *Helicopter Dynamics and Aerodynamics. 1959.* The Macmillan Company, 1959.

[7] Jihong Zhu Shuo Zeng. Adaptive Compensated Dynamic Inversion Control for a Helicopter with Approximate Mathematical Model. Technical report, Dept. of Computer Science & Technology, Tsinghua University, Beijing, China, August 2006.

[8] Peter C Wang and Terry E Shoup. Parameter sensitivity study of the Nelderâ"Mead Simplex Method. *Advances in Engineering Software*, 42(7):529–533, July 2011.

[9] W Xiaoqing and H Yimin. Model system design for small-size unmanned helicopter based on moment theory, blade element theory and FlightLab software. *Systems and Control in Aerospace and Astronautics, 2008. ISSCAA 2008. 2nd International Symposium on*, 2009.

Development of Neural Networks for Air Data Estimation: Training of Neural Network Using Noise-Corrupted Data

Manuela Battipede, Piero A. Gili, Angelo Lerro
Politecnico di Torino, C.so Duca degli Abruzzi 24, 10129 Torino, Italy

Silvio Caselle, Paolo Gianardi
Alenia Aeronautica, C.so Marche 41, 10146, Torino, Italy

Keywords: *air data, estimation, neural network, real, noise*

List of symbols

ADS	Air Data System
ADU	Air Data Unit
FCC	Flight Control Computer
MALE	Medium Altitude Long Endurance
MLP	Multi Layer Perceptron
NN	Neural Network
NNxC, NNxN	Neural network x trained with clean (C) and noisy (N) signals
UAV	Unmanned Aerial Vehicle
a_x, a_y, a_z	Acceleration along x, y and z directions [m/s^2]
C_Z	F_z coefficient
$C_{Z,\tau}$	Derivative of F_z coefficient respect with the generic non dimensional τ
F_x, F_y, F_z	Total external force along x, y, z directions [N]
g_0	Gravity acceleration [m/s^2]
n_x, n_y, n_z	g-acceleration
pi	Number of past inputs used for the neural network
p_s	Static pressure [Pa]
q_c	Dynamic pressure [Pa]
p, q, r	Angular velocity along x, y, and z directions of body axes [rad/s]
u, v, w	True air speed component in the body reference frame along x, y and z directions [m/s]
t	Time [s]
V	True air speed magnitude [m/s]
α	Angle of attack [deg]
β	Angle of sideslip [deg]
δ_e, δ_r, δ_a	Elevator, rudder and ailerons deflection [deg]
ϕ	Roll attitude angle [deg]
ϑ	Pitch attitude angle [deg]
$\dot{(\,)}$	Time derivative

Abstract

In this paper several virtual air data sensors have been modeled using neural networks in order to estimate the angle of attack. These virtual sensors have been designed and tested using the aircraft mathematical model of the De Havilland DHC-2 "Beaver". The aim of the work was to evaluate the degradation of neural network performance, which is supposed to occur when real flight data (affected by noise) are used instead of simulated ones. Each input and output signals have been characterized with a realistic noise level according to a model derived from Alenia Aeronautica Sky-Y flight test campaigns. By the analysis of the results hereinafter presented it is clear that the neural network developed for virtual signals cannot be used with real signals, as very low estimation accuracy may result. The neural network implemented as virtual air-data sensor for real applications needs to be trained with noise-corrupted signals in order to keep the accuracy bounded within acceptable limits.

Introduction

Modern UAVs are equipped with numerous electronic devices and sensors, like Air Data probes, to measure flight variables. Without them the remote piloting and automatic flight could not be realized.

For Air Data Systems (ADS), where redundancy is requested for compliance to safety regulations, a voting and monitoring capability that cross-compares the signals from different sensors is needed: this feature is therefore used to detect and isolate ADS failures down to the single sensor, depending on the level of redundancy requested [1], [2]. The increasing need of modern UAVs to reduce cost and complexity of on-board systems has encouraged the practice to substitute, whenever feasible, expensive, heavy and sometimes even voluminous hardware devices with executable software code. A practical example, which is referred to as *analytical redundancy* in the current literature, is the process of replacing some of the actual sensors with virtual sensors, which can be used as *voters* in dual-redundant or simplex sensor systems, to detect inconsistencies of the hardware sensors and can be eventually employed to accommodate the sensor failures.

More generally, analytical redundancy identifies with the functional redundancy of the system. The idea of using software algorithms to replace physical hardware redundancy was introduces as soon as digital computers started being used in the 1970's to perform redundancy management. Approaches developed to detect and isolate sensor failures were ultimately to become important parts of later control reconfiguration schemes. An example is the Sequential Probability Ratio Tests that were flight-tested on the F-8 Fly-by-Wire demonstrator in the late 1970's [3]. Throughout the 1970s and 1980s many papers appeared describing various algorithms for managing redundant systems and redundant sensors. Many attractive advanced algorithmic solutions have been proposed especially throughout the last two decades, mostly related to model-based techniques but capable of taking into account some robustness requirements with respect to exogenous disturbances and model uncertainties. The challenge is to answer the following questions: could these techniques, developed within the academic community, be applied to existing industrially-developed UAVs and can they fit into current certification processes?

The air-data virtual sensor which is being developed through the collaboration between *Alenia Aeronautica* and the *Politecnico di Torino* is based on neural networks (NN) to overcome some of the shortcomings related to model-based methods. In the aerospace field, NNs are often used as system identification devices to estimate aerodynamic coefficients [4], angle of attack [5], [6], and sideslip [7] from data derived from other sources that are not vanes, differential pressure sensors or modern multifunction probes. This kind of virtual sensor, however, needs to be fed at least

by an actual value of dynamic pressure, as observed in all the cases presented in previous works [8], [9], [10]. As far as prediction is concerned, it is very well known that NNs can work as human brain: starting from the observation of physical phenomena (the training stage), the NN should be able to mimic the underlining physics of the observed event and predict its future developments. Notoriously, in implementing a NN, the training phase plays a crucial role: although the training techniques are all widely experimented, documented and substantiated, in fact, it is not trivial to collect a training data set that is adequately representative of the model dynamics we want the NN to mimic. Training data might come from flight tests, wind tunnel data or from a mathematical model. Using a mathematical model, however, only a preliminary design of a virtual sensor can be developed, because of the discrepancies inherently introduced in the modelization process, which reintroduces the main drawback of the majority of the model-based techniques. The wind tunnel data are usually not suitable in this circumstances because it is very hard and expensive to perform dynamic maneuvers. The best way to design a virtual sensor is, thus, through the use of real flight data. This kind of procedures has the disadvantage that the training signals are affected by several sources of noise, such as air turbulence, structure vibrations and electronic noise.

The aim of this work is evaluating the difference in neural network performance (or virtual sensor accuracy) between trainings performed using aircraft mathematical model results and flight test data. To perform this kind of evaluation a well-known mathematical model has been considered: training is performed using both smooth, noise-free signals and noise-corrupted signals where noise is artificially introduced using a model derived and validated from Alenia flight tests. The final goal of this activity is to assess the feasibility of the method and estimate the NN complexity necessary to maintain the virtual sensor angle of attack accuracy within the threshold of ±1 deg.

Description of the non-linear Beaver model

The FDC toolbox for Matlab was developed starting from the De Havilland DHC-2 “Beaver” aircraft [11] and it was designed to analyze aircraft non-linear dynamics and flight control systems. The aircraft is modeled as a rigid body with constant mass value, flying in still air. The Beaver command system is very basic and only the primary control surfaces are modeled whereas, for example, flap deflection effects are neglected.

The relationship between the wind-axis and body-axis velocities can be expressed as follows

$$\begin{Bmatrix} u \\ v \\ w \end{Bmatrix} = V \begin{bmatrix} \cos\alpha\cos\beta \\ \sin\beta \\ \sin\alpha\cos\beta \end{bmatrix} \qquad (1)$$

or, conversely:

$$\begin{aligned} V &= \sqrt{u^2 + v^2 + w^2} \\ \alpha &= \tan^{-1}\frac{w}{u} \\ \beta &= \tan^{-1}\frac{v}{\sqrt{u^2 + w^2}} \end{aligned} \qquad (2)$$

Differentiating Eq. 2, the resulting system, constituted by three ordinary differential equations,–needs to be solved in order to find the analytical expressions for V, α, and β. For simplicity, attention is focused on the angle of attack, so differentiating the second of Eq. 2 it follows:

$$\dot{\alpha} = \frac{u\dot{w} - w\dot{u}}{u^2 + w^2} \qquad (3)$$

Substituting u and w of Eq. 1 and rearranging the terms Eq.3 yields to:

$$\dot{\alpha} = \frac{\dot{w}\cos\alpha - \dot{u}\sin\alpha}{V\cos\beta} \qquad (4)$$

According to [11], the relationships between axial acceleration, gravity acceleration and angular velocities in the body reference frame can be written as

$$\begin{aligned}\dot{u} &= rv - qw - g_0 \sin\vartheta + F_x \\ \dot{v} &= -ru + pw + g_0 \sin\phi \cos\vartheta + F_y \\ \dot{w} &= qu - pv + g_0 \cos\phi \cos\vartheta + F_z\end{aligned} \qquad (5)$$

where the contribution of external force ($F_i = ma_i$) can also be indicated in terms of number of g acceleration $n_i = a_i/g$ ($i = x, y, z$). Introducing an expression for F_z as a function of pilot commands, or, more in general, as a function of the control surface deflections, it is possible to obtain the relationships between the normal acceleration n_z and the command actions, such as in Eq. 6, purposely written for the Beaver model [11]:

$$\begin{aligned}n_z = \frac{F_z}{m} = \frac{1}{2}\rho V^2 (C_{z0} + C_{z,\alpha}\alpha + C_{z,\alpha^3}\alpha^3 \\ + C_{z,q}\frac{qc}{V} + C_{z,\delta_e}\delta_e + C_{z,\delta_e\beta^2}\delta_e\beta^2)\end{aligned} \qquad (6)$$

Therefore, considering Eq. 5, Eq. 4 can be rearranged as follows

$$\begin{aligned}\dot{\alpha} = \\ \frac{(qu - pv + g_0 \cos\phi \cos\vartheta + n_z)\cos\alpha}{V\cos\beta} \\ - \frac{(rv - qw - g_0 \sin\vartheta + n_x)\sin\alpha}{V\cos\beta}\end{aligned} \qquad (7)$$

Then, substituting Eq. 1, we can obtain the final expression for angle of attack ODE:

$$\begin{aligned}\dot{\alpha} = \frac{1}{V\cos\beta}\{ \\ [V(q\cos\alpha\cos\beta - p\sin\beta) + n_z \\ + g_0 \cos\phi\cos\vartheta]\cos\alpha \\ -[V(r\sin\beta - q\sin\alpha\cos\beta) - g_0 \sin\vartheta \\ + n_x]\sin\alpha\}\end{aligned} \qquad (8)$$

Eq. 8 is essential to understand the independent variables which α depends on. According to Eq.8, thus, it is possible to write the following functional relationship:

$$\alpha = f_\alpha(q_c, n_x, n_z, \beta, \vartheta, \phi, p, q, r) \qquad (9)$$

where the velocity V has been substituted by dynamic pressure because it is the source from which the velocity is derived from. Following a similar procedure the functional relationship for β can be written as follows:

$$\beta = f_\beta(q_c, n_x, n_y, n_z, \alpha, \vartheta, \phi, p, q, r) \qquad (10)$$

So, substituting β in Eq. 9 with its expression written in Eq. 10, collecting all the independent variables and explicitating the dependency on pilot demands, Eq. 9 can be further expanded as:

$$\begin{aligned}\alpha = F_\alpha(q_c, n_x, n_y, n_z, \vartheta, \phi, \\ p, q, r, \delta_e, \delta_r, \delta_a)\end{aligned} \qquad (11)$$

From Eq. 11, it follows that the angle of attack is a function of the dynamic pressure, body axis acceleration, Euler angles and pilot commands. The advantage of using a neural network is that the information expressed by Eq.11 is already adequately detailed to implement a virtual sensor, whereas an exact or approximated expression for the F_α function is not strictly necessary.

Maneuvers for training and testing stages

As stated above, the training stage requires a collection of maneuvers that should be thoroughly representative of the flight envelope, so the choice can be oriented towards a flight containing a wide selection of command excitations, in terms of shape and frequency, and preferable where the same excitations reach their entire span range. An example of the maneuvers performed for data gathering is given in Fig. 1.

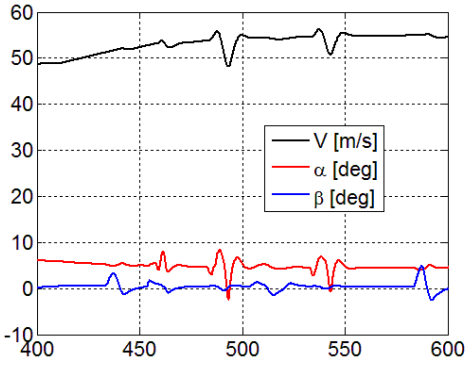


Fig. 1 Detail of aircraft dynamic of the training stage

These maneuvers have been obtained using an autopilot which commands the aircraft to follow a prescribed signal or surface control deflections. In the present case, the longitudinal aircraft dynamics are controlled by acting on the elevator surface, whereas the asymmetrical aircraft motion is obtained following ϕ reference signals. The simulated flight is composed by a basic set of maneuvers performed at different flight speed condition in order to have a training set that covers the most part of flight envelope. This basic set consist of yaw doublets, ϕ-hold maneuvers and maneuvers in the longitudinal plane. The flight envelope considered here extends from 35 to 55 m/s as suggested in [11].
Fig. 2 shows the maneuver that have been used to test, or validate, the neural network.

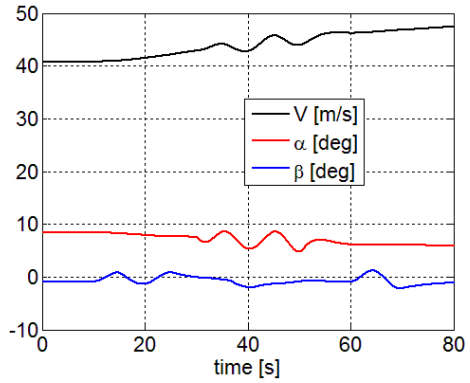


Fig. 2 Aircraft dynamics of the test stage

First the velocity is increased by acting on the throttle, and at the same time a beta sweep is performed. Then a double-repeated pitch sweep overlaps to a 40 deg heading turn at constant ϕ=10 deg. The last maneuver is a beta sweep. It can be noted the in this kind of validation maneuvers there are many superimpositions of motion which are not present in the training set. Based on the experience gained flying the "Beaver" model, this kind of maneuver is one of the worst that could be used as validation case.
As a fail/pass criterion the maximum error from nominal value has been selected rather than the classical concept of performance. Generally speaking, in fact, the *neural network performance* refers to the root mean square error of estimated quantities over an entire flight. As the virtual sensor is a discriminator source we are most interested in the maximum deviation between estimated and true data.

Noise characteristics

If compared to simulated signal, actual signals are always affected by noise deriving from external sources, such as air turbulence, and internal sources, mostly related to electronic devices. The discrepancies between the mathematical model and the real aircraft can be seen as another source of noise, especially when the sensor sampling is close to the first structural mode frequencies, such as in slender bodies. As it is always quite complicated to keep separate each source of noise, therefore any single signal has been marked with its own noise level.
To develop a realistic noise model, data have been gathered during a flight test activity performed on Alenia Sky-Y. The Sky-Y is a propeller-driven MALE unmanned air vehicle designed by Alenia as a demonstrator for surveillance and patrolling purposes (see Fig. 3).



Fig. 3 Alenia Sky-Y

In the course of the aforesaid campaign the Sky-Y was equipped with two 5-hole nose boom probes, each of them able to provide a complete set of air data $(\alpha, \beta, p_s, q_c)$.
The experience acquired allows us to assign a realistic peak-to-peak noise level to each input/output NN signal. Even if the degradation of signals depends on the specific electronic

devices, the actual level of air turbulence and the type of aircraft, it has been decided to adopt a conservative approach based on the worst-case scenario, incorporating the combination of highest noise level detected within the Alenia data base.

Variable	Noise Level (peak to peak)
α	±0.1 deg
q_c	±0.1 mBar
n_x	± 0.1 g
n_y	± 0.1 g
n_z	± 0.1 g
ϑ	± 1E-2 deg
ϕ	± 1E-2 deg
p	± 0.15 deg/sec
q	± 0.15 deg/sec
r	± 0.15 deg/sec
δ_e	± 3E-3 deg
δ_r	± 3E-3 deg
δ_a	± 3E-3 deg

Tab. 1 Noise level derived from flight test data

A white noise with zero mean value have been added to each signal and the peak to peak values have been selected according to Tab. 1.

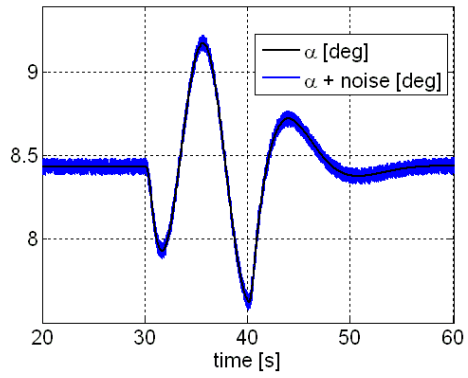


Fig. 4 Example of angle of attack without and with noise

Fig. 4 shows an example related to the angle-of-attack. It must be noted that for the angle of attack estimation no electronic filters have been considered.

Results using noise-free signals

To assess the feasibility of the method it has been decided to simulate the best-case scenario, implementing a NN trained and tested on noise-free signal: performance are evaluated on the bases of ±1 deg error specification for the angle of attack. The virtual sensor has been developed using a state-of–the-art tool, developed by third parts under the Matlab simulation environment [13]. The kind of neural network used is a multi layer perceptron (MLP) it has one hidden layer with 15 neurons characterized by hyperbolic tangent activation function and one output layer, as described in Fig. 5.
The system identification exploits the NNOE model trained with the Marquardt method [14]. The input variables have been chosen from Eq. 11 and they are directly processed by the neural network without any kind of buffer of memory, which would have been needed in the case that past inputs were also requested.

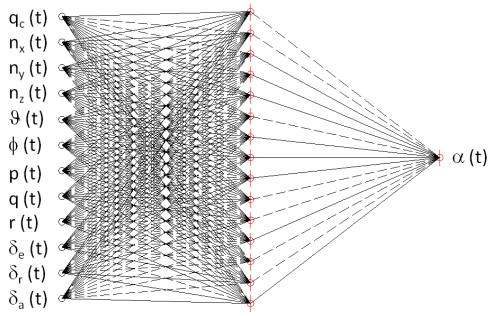


Fig. 5 Neural network scheme (NN1) used for AOA estimation: the solid lines stand for positive weights, the dashed lines stand for negative weights, the red circles are the network neurons

This kind of network has the enormous advantage to be light in terms of hardware resources: it only needs to be connected with the requested variables in input ports at current time t to give an estimation of $\alpha(t)$. This kind of neural network can be identified as NN1 and the NN1C is, in particular, the NN trained using noise-free 'clean' signals, otherwise if NN1 is trained with noisy signals it can be

indicated as NN1N. The results presented in this section are obtained using the net NN1C. The estimation error for the angle of attack is plotted in Fig. 6 and the maximum value is less than 0.1 deg (peak to peak).

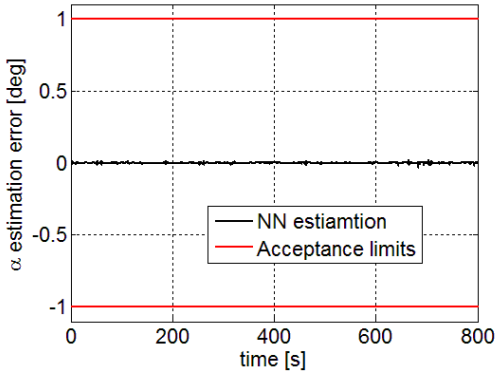


Fig. 6 Estimation error obtained using the training maneuver of Fig. 2 for NN1C

Fig. 7 shows the estimation error for the test maneuver using clean signals to feed the virtual sensor.

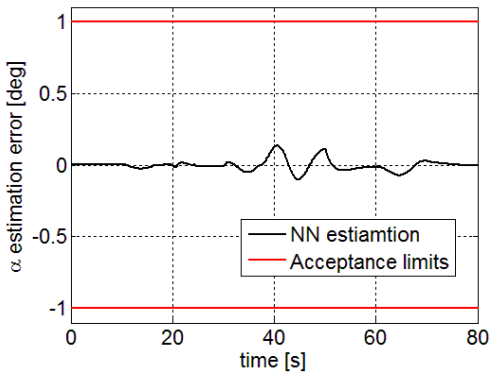


Fig. 7 Estimation error obtained using the test maneuver for NN1C

As could be expected, when noisy inputs are used with a net NN1C the errors are very large (Fig. 8). To reduce these errors a training using noisy signals can be performed as described in the following section.

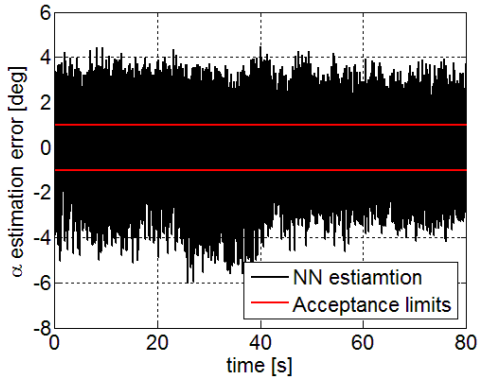


Fig. 8 Estimation error obtained using the test maneuver and noisy inputs for NN1C

Validation using noisy signals

Using the same scheme of neural network of Fig. 5 and described at the beginning of the previous section, a new training was performed using a combination of clean and noise-corrupted signals (net NN1N) [14].
It is worth noting that in literature other training algorithms have been developed to work better in presence of noisy signals [15] and they could have employed, but the main task of this phase of the activity is to investigate the worst-case performance losses rather than optimizing them.
The results in terms of estimation error of angle of attack are plotted in Fig. 9.

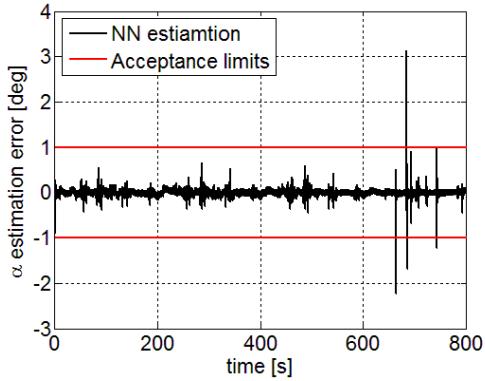


Fig. 9 Estimation error obtained using the training maneuver and noisy inputs for NN1N

It is quite easy to understand that the scheme NN1 has the advantage to be very light in terms of computational costs but does not give the net any information about the development of the

input signals. In order to give more information about the signal evolution to the net, the virtual sensor has been fed with current input variable values and also with previous input values. A more uniform error over the entire flight envelope can be obtained using 2 of past inputs, which account for *pi=2*. In other words the condition *pi=2* guarantees a better performance, in terms of root mean square error than pi=1. So, a new virtual sensor has implemented based on the neural network (NN3) with 45 neurons presented in Fig. 10.

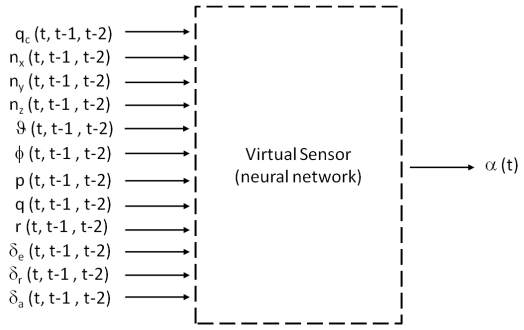


Fig. 10 Neural network scheme (NN3) used for AOA estimation with pi=2

Fig. 11 shows the error obtained during the training of NN3 using (the same) noisy signals, the maximum error is about 0.2 deg, ten times more than of maximum errors obtained with NN1C.

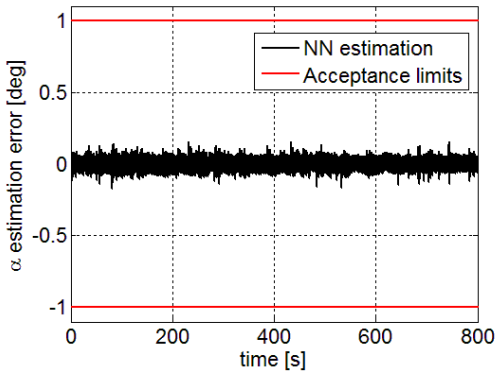


Fig. 11 Estimation error obtained using the training maneuver and noisy inputs for NN3N

Through the comparison of Fig. 11 and Fig. 9, however, it is possible to appreciate the performance improvements obtained using past input values of all the variables involved, which have clearly the drawback of the extra computing cost in terms of memory buffer.

The estimation error on the test maneuver is plotted in Fig. 12. If the last figure is compared with the Fig. 7 it can be noted that the maximum error has doubled. This is due to the particular noise level: the lower the noise level the lower the estimation error; the higher the *pi* value the lower the estimation error.

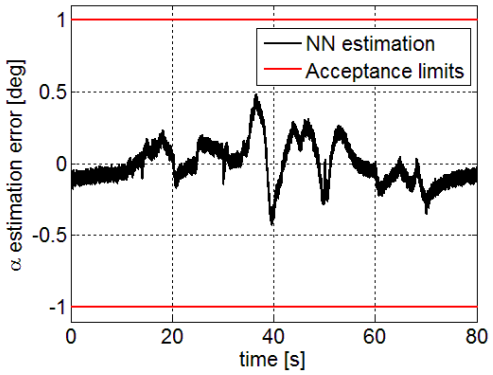


Fig. 12 Estimation error obtained using the training maneuver and noisy inputs for NN3N

Conclusions

The present work presented an analysis of the degradation which occurs on a neural virtual air-data sensor, when noise-corrupted data are processed. In particular this simple analysis has highlighted the potential errors that might be introduced using a scarcely realistic data source to estimate the virtual sensor parameters.

The first part of the paper deals with the development of a virtual sensor based on noise-free signals: the functional relationship between the angle-of-attack and the most relevant flight variables has been derived and assessed in an ideal environment. In particular, the virtual air-data sensor, implemented ad-hoc for the De Havilland DHC-2 “Beaver” analytical model, has been conceived as a function of data derived from gyros, accelerometers, pilot commands and a Pitot tube (dynamic pressure).

The second part of the paper deals with the training of the neural network using noisy signals, for which each input variable has been provided with an artificial noise level derived from available flight test data.

The work has highlighted that is not possible to train a neural network with clean signals and then operate the NN with noisy signals: in other words there are no many chances of good performance to develop and train a virtual sensor using the aircraft mathematical model only and then using this sensor with actual signals for real applications. Even if the NN were extended to receive past inputs and trained with noise-corrupted signals, the net performance would suffer a degradation of about five order of magnitude with respect to a neural network trained with noise-free signals.
As shown, the best practice in order to design a virtual sensor which satisfies such performance specifications is training the neural networks with noisy signals. As a matter of fact, a copious and meaningful set of flight data would be highly desirable, but it might be the case that this is not available. Can it be substituted with a set of data artificially corrupted with a noise model derived as the one used in the present paper? Giving an answer to this question means reformulating the technique comparing artificially noise-corrupted simulated data with flight data collected for the same aircraft, which is clearly a future developments of the present work.
Is also realistic to keep neural network dimensions within certain limits (in terms of hardware resources): the state-of-the-art in the neural network field proposes many techniques that involve more complex neural network architectures which can be expressed with a smaller amount of parameters. The optimization of the neural technique can be regarded as an option in this first phase, but it remains an issue that must be addressed, especially in consideration of future implementation of the virtual sensor as an analytical redundancy tool for the on-board sensor management system.

Acknowledgements

The present work has been performed through a close cooperation with Alenia Aeronautica S.p.A. and through the use of Confidential Information and Data, property of Alenia Aeronautica S.p.A., which remains the sole owner of all such relevant IP rights. The result of the present work shall be, therefore, property of Alenia Aeronautica S.p.A.

Bibliography

[1] Napolitano M.R., An Y., Seanor B.A, “A fault tolerant flight control system for sensor and actuator failures using neural networks”, Aircraft Design, 3, pp. 103-128, (2000).
[2] Oosterom M., Babuska R., “Virtual sensor for fault detection and isolation in flight control systems - fuzzy modeling approach”, Proceedings of the 39th IEEE Conference on Decision and Control, 3, pp. 2645-2650, (2000).
[3] Tomayk, J.E., Computers Take Flight: A History of NASA’s Pioneering Digital Fly-By-Wire Project, NASA SP-2000-4224, NASA, Washington, D.C., 2000.
[4] Norgaard M., Jorgensen C. C., Ross J. C., “Neural Network Prediction of New Aircraft Design Coefficients”, NASA Technical Memorandum 112197, 1997.
[5] Oosterom M., Babuska R., “Virtual Sensor for the Angle-of-Attack Signal in Small Commercial Aircraft”, 2006 IEEE International Conference on Fuzzy Systems. Vancouver, BC, Canada July 16-21, ISBN 0-7803-9488-7, 2006.
[6] Samara P.A., Fouskitakis G.N., Sakellariou J.S. and Fassois S.D., “Aircraft Angle-Of-Attack Virtual Sensor Design via a Functional Pooling Narx Methodology”, Proceedings of the European Control Conference (ECC), Cambridge, UK, 2003.
[7] Crowther W.J., Lamont P.J., “A prediction model for vehicle sideslip angle based on neural network” Information and Financial Engineering (ICIFE), 2nd IEEE International Conference, ISBN: 978-1-4244-6927-7, pp. 451 – 455, 2010.
[8] Rohloff T. J., Whitmore S. A., Catton I., “Air Data Sensing from Surface Pressure Measurements Using a Neural Network Method”, AIAA Journal, Vol. 36, No. 11, November 1998.
[9] Samy I., Postlethwaite I., Green J., “Neural-Network-Based Flush Air Data Sensing System Demonstrated on a Mini Air Vehicle”, Journal of Aircraft, Vol. 47, No. 1, January–February 2010.
[10] Calia A., Galatolo R., Poggi V., Schettini F., “Multi-hole probe and elaboration algorithms for the reconstruction of the air data parameters”, ISIE08 - International Symposium on Industrial Electronics - IEEE 2008.

[11] Rauw M. “FDC 1.2 – A Simulink Toolbox for Flight Dynamics and Control Analysis”, http://dutchroll.sourceforge.net/fdc.html, 2001.
[12] Stevens B. L., Lewis F. L., “Aircraft Control and Simulation”, John Wiley And Sons Inc, New York, ISBN 0471613975 / 0-471-61397-5, 1992.
[13] Norgaard M., “The nnsysid toolbox for use with Matlab”, http://www.iau.dtu.dk/research/control/nnsysid.html
[14] Réda A., Aoued B., “Artificial Neural Network-Based Face Recognition”, Second International Symposium on Communications, Control and Signal Processing, ISCCSP, 2006.
[15] Van Gorp J., Schoukens J., and Pintelon R., “Learning Neural Networks with Noisy Inputs Using the Errors-in-Variables Approach”, IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 11, NO. 2, MARCH 2000.

CEAS 2011

Space Robots and Manipulators

3rd CEAS Air&Space Conference
21st AIDAA Congress

Decentralized Walking Control for a Hexapod Robot for Space Exploration using Dynamical Neural Networks

M. Massari
Politecnico di Milano, Dipartimento di Ingegneria Aerospaziale
Via La Masa, 34 – 20156 – Milano (MI), Italy
mauro.massari@polimi.it

Keywords: *neural networks, hexapod robot, walking, advanced control, biomimicry.*

Abstract

In this work a biomimetic approach has been adopted to design a decentralized walking motion generator for a Hexapod robot. In particular, Continuous Time Recurrent Neural Networks (CTRNN) have been investigated for the realization of parts of the modules composing the controller. Starting from the design of efficient controllers for every phase of the step cycle, a selector regulating the transition among phases has been developed producing a complete single leg controller. At this level mutual influences among neighboring legs have been introduced and only in the last phase of the work global coordination parameters have been taken into account. The preliminary evaluation of the controller performances obtained using a multi-body numerical simulation tool are presented.

1 Introduction

The reproduction of what the nature offers us has been for long time one of the most exciting challenges to scientists and engineers. The thrust to copy the animals and their behavior is justified by the fact that biological systems tend to have a better efficiency than those found in their artificial counterparts. This concept can also be applied in space exploration, replacing the classical wheeled rovers with legged robots inspired by the animals. A wheeled system is almost always simpler, cheaper and more reliable than a legged one, and it also shows continuous stability, that can be achieved without any control strategy. However, there are conditions in which legs become preferable to wheels. In particular, legs don't require a continuous area of solid ground for moving and so the navigation is not constrained to specific kind of surfaces. A legged system is not stable per se, like a wheeled one, because its stability depends fundamentally on how many legs touch simultaneously the ground. However, If during the walk, three or more legs are always in contact, the motion is a mere juxtaposition of equilibrium state and the system is called statically stable. This means that the stability is not time dependent and a global stability controller is not required if a solution with four or more legs is considered.

The exploration of a space environment requires a very robust behavior, so a statically stable robot with an high number of legs is considered. The complexity of an high DoF system is balanced by its adaptability to an unknown environment and its flexibility to achieve lots of different tasks. It also makes it

more fault tolerant, since the robot can lose a leg without losing its capability for walking. For this reason it has been decided to design a controller for a hexapod robot with a total of 18 rotational DoF. The controller design is based on observed natural behavior and the main source of inspiration is the stick insect, because of its high level of mobility and of the ease with which it can be studied.

Biological findings about various animals show that the movements of individual legs are managed by fairly independent control systems [1]. On the other hand each leg has an influence on the other ones and it's the combination of all these influences that allows to generate an efficient walk. This means that the global motion problem can be solved by accomplishing two macro-tasks: control the motion of each single leg and coordination of all the legs. The control of the single leg it's fundamentally the problem of generating the step cycle.

During the walking, the individual legs typically move cyclically and, in order to facilitate the analysis, the motion of a leg is often partitioned into phases: the control of each phase and transition among phases are called subtasks of the walking problem.

The coordination among legs is essential to regulate global parameters such as advancing speed, stability margin or body attitude and position. The first task consists in maintaining a steady motion and it can be obtained introducing local weighted influences among neighboring legs. The main advantage of a such structured solution is the possibility to design each module independently because they are only weakly coupled. This allows the use of a simpler, decentralized control system, since it requires a less performing CPU than a global controller. The possibilities offered by using a decentralized control matches the requirements of the low-performances space-certified CPUs.

In this work a bottom-to-top approach has been adopted to create an incrementally realistic walking motion generator. In particular, Continuous Time Recurrent Neural Networks (CTRNN) [2] have been investigated for the realization of parts of the modules composing the decentralized controller. Starting from the design of efficient controllers for every phase of the step cycle, a selector regulating the transition among phases has been developed producing a complete single leg controller. At this level mutual influences among neighboring legs have been introduced and only in the last phase of the work global coordination parameters have been taken into account accomplishing the global task of walking control. However, the present work will concentrate on the controller of a single leg motion, which is responsible of the actual control of the motion. The leg-coordination and gait generation, that are needed to define target leg motions, can be implemented using different approaches. The most promising solution identified for those modules is just outlined in this work.

In section 2 an introduction to the problem of walking motion generation is introduced together with the proposed approach, while in section 3 and 4 the details on the implemented controller are outlined together with preliminary results obtained with numerical simulations.

2 The Problem of Walking

Walk can be defined, in the most general way, similarly to a method of terrestrial locomotion that uses limbs. The main differences with other kinds of movement are the presence of a non-continuous contact with a substrate and the usage of more than one multiple DoFs appendage.

Walking is one of many behaviors where machines still lag notably behind the performance of animals, so it is natural to examine walking in animals to look for hints to improve the performance of machines. The most important global aspects of the walking problem can be summarized as follows:

Redundancy. the number of degrees of freedom is normally larger than the one necessary to perform the task.

Autonomy. The redundancy requires the system to select among different configuration alternatives without external command;

Situatedness. A walking system is situated in a complex and unpredictable environments, which means that any movement may be influenced by the environment.

The best solution to fulfill all these tasks is a decentralized control structure that consists of a number of distinct modules each one solving particular subtasks. The division into modules and the choice of the subtasks have been accomplished assembling logical components to model the system properties in a biomimetic way. Although many animals have body appendages that can be used for walking, only a few species have been investigated in sufficient detail. Among them, the stick insect is the more suitable for the purposes of this work thanks to its capability to walk on very complex and irregular substrates.

The stick insect can be modeled as a rigid body supported by six legs: two legs are called *ipsilateral* if situated on the same side, *contralateral* if situated on the opposite side. Each leg has been modeled as a manipulator with three joints and divided into three movable sections: the *coxa*, closest to the body, the *trochanter-femur*, or just femur as they are fused and the *tibia*.

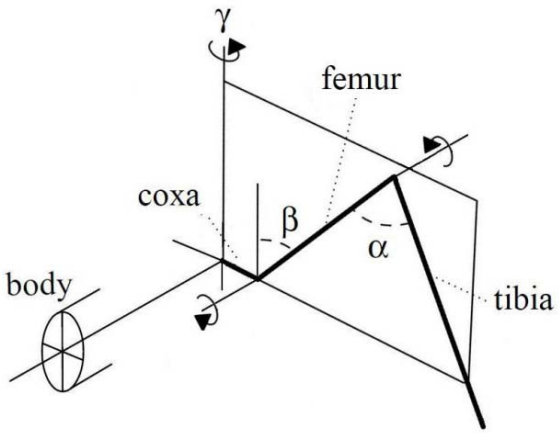


Fig. 1 Leg geometric parameters

In Fig. 1, the geometric parameter defining the leg are illustrated. Starting from the body it's possible to identify: the Body-Coxa (BC) joint (angle γ); the Coxa-Trochanter (CT) joint (angle β) and the Femur-Tibia (FT) joint (angle α). The γ angle determines mostly how much near to the front or to the back the leg plane is positioned, the β angle specifies how high or low the knee is and the α angle determines the lateral proximity of the tarsus to the body. Note that all the joints influence motion in all directions, i.e. joint variables are non-linear with respect to Cartesian coordinates.

2.1 Control Strategy

The modular approach adopted in this work allows the simplification of the problem of walking control in a large measure, but needs a good distinction among all the tasks.

The approach to the problem of walking control is not unique: it heavy depends on external parameters such as substrate properties and global required speed. Two different methods have been applied for the realization of walking controller: central pattern generator (CPG) and reflexes-based systems.

Typically a CPG controls each single leg and the coordination is obtained by a temporal synchronization of all the CPGs. This system allows to reach very high advancing speeds and a regular motion but require legs with particular reactive characteristics and powerful motors.

A reflex-based controller can be viewed as a closed loop control system with fixed input that produce an output only in relation with an environment depending input. The coordination is typically based on rules that mediate the reflexes of every single leg. The reflex driven approach seemed to provide a simple way to stabilize the walking patterns, also in very difficult conditions, by providing a set of fixed situation-reaction rules to external disturbances and as a way to regulate leg coordination among multiple independent legs.

For a space-oriented project, like this one, it's clear that, at this level of development, a very fast CPG-based robot has only little utility. In an unknown environment, like planetary surfaces, the first priority for a robot is to ensure a steady motion in any moment.

A walking robot, besides, must be able to reach locations out of the range of wheeled ones, balancing its complexity with its higher capabilities. Furthermore a slow locomotion allows to reduce energy consumption and to avoid shocks to the scientific payload. All these considerations favored the choice on a reflex-based controller.

The approach followed in this work allows to realize a decentralized reflex-based controller,

in which each modules is responsible of subtasks which are coordinated by a set of rules.

3 Single Leg Controller

The problem of controlling a single leg can be simplified partitioning the step cycle into the *stance* and *swing* phases. During the *stance* phase, the leg maintains ground contact and is retracted to propel the body forward, while supporting the weight of the robot. During the *swing* phase, the leg is lifted off the ground and moved in the direction of walking, to touch down at the location where the next stance should begin. Fig. 2 shows these two phases.

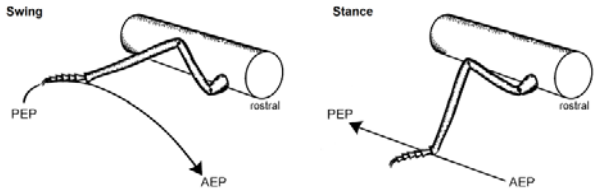


Fig. 2 Walking phases

The phases are mutually exclusive behaviors as a leg cannot be in swing and in stance at the same time. Therefore, the control structure must include a mechanism for deciding the transition between swing and stance phase, creating the complete step cycle. This criterion can be identified on the basis of spatial variable: at the AEP the transition from swing to stance occurs; at the PEP the transition from stance to swing occurs. This means that to produce a stereotypical step cycle, the leg controller can be composed by only three modules each of them fulfilling a subtask: a swing module that controls the leg during the transfer phase; a stance module that controls the leg during the support phase; a selector that works at an higher level and switches between using the outputs of either the stance or the swing module. To guarantee a robust and efficient motion on irregular surfaces a stereotypical step cycle is not sufficient. For this reason a limited number of standard reactions to typical unpredictable disturbances, called reflexes, has been added to the standard controller. The scheme of the complete single-leg controller is illustrated in Fig. 3 with particular attention to the input/output relation.

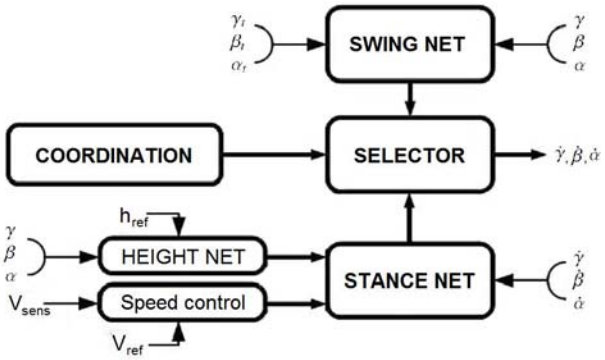


Fig. 3 Single leg controller

Controlling a swing movement is easier than controlling a stance movement, because a leg in swing is mechanically uncoupled from the environment and, owing to its small mass, essentially uncoupled from the movement of the other legs. Therefore, whatever a leg does during a swing movement, it has virtually no impact on the movements of the other legs.

Since the stick insect regulates swing movements and compensate for external perturbations, the control system must receive proprioreceptive sensory feedback to account for a closed-loop control. Moreover, because swing movements are known to be targeted towards a given location, the controller must also receive information about a desired posture.

A general swing can be completely defined by three points: the posterior extreme position (PEP), where the leg lifts off the ground to start swing, the anterior extreme position (AEP) where the leg ends swing by touching the ground and the swing extreme position (SEP), the dorsal extreme position of the swing movement. To fulfill the swing task, the controller must show two different features: the leg has to lift off from the PEP and the foot has to move to the target position (AEP).

The one permits to avoid all the obstacles below the trajectory so, with an high SEP, the walk will be possible also on rough terrain. The second one implicitly imposes that the AEP is an attractor for the controlled system, but only in the two dimensions of the ground plane. In the direction perpendicular to the support surface, the attraction point must be placed

below the ground reference position thus prohibiting the foot to move too much slowly immediately before the ground contact.

In a complex environment, the task of controlling the stance movements of all the legs on the ground poses several major problems. It is not sufficient to assign the motion of each leg on its own, because the mechanical coupling through the substrate implies the coordinated action of all the joints that connect the body to the ground. Thus, the coordination of up to 18 joints may be needed at any moment in time. In classic approach to walking machine, this problem can be solved only by computationally intensive algorithms that take into account the complete dynamics of all the legs in stance. The nature of the mechanical interactions and the search for a global optimum requires a single central control system. However, in technical solutions, such centralized system makes real-time control difficult, even in simple cases.

Although the task to coordinate many joints appears complex, insects master it with their simple nervous system. This means that they use, instead of a centralized controller, a decentralized approach able to exploit the dynamics of the body-environment interaction.

To solve this particular problem in a very easy way, Cruse et al. [3] propose to replace a central controller with distributed control in the form of local positive feedback. The positive feedback occurs at the level of the single joint, feeding back the position signal to control the motor output of the same joint, without direct relationship with the rest of the kinematic chain.

To understand how positive displacement feedback works, one must consider, a standing legged system that begins to move one joint, while keeping all the feet on the ground. Owing to the mechanical connections, all other joints of the moved leg, and even joints of the other legs, passively adjust to the active joint movement. Thus, the movement direction and speed of each joint do not have to be computed because this information is already provided by the physics of the whole system. This example is related to postural control, a problem ruled by resistance reflexes that maintain the position of each leg with local negative feedback circuits. In this work, the control of walking is realized replacing the resistance reflex with an assistance reflex: it can be modeled by a positive feedback that transforms this passive movement into an active one.

3.1 Controller Implementation

The control of the swing phase can be achieved with a single module, consisting in an artificial neural network (called *swing-net*), that generates the desired trajectory during this phase. This trajectory must show two different behaviors: the lift off from the ground and the targeting on the AEP.

The choice of a dynamical neural network is based on its superior capability to simulate a time-dependent behavior. In particular, Continuous Time Recurrent Neural Networks (CTRNN) will be investigated. In this work the current angles have been considered as states and the target ones as inputs. This choice allows to update the state with the measured angular displacements and to use directly the state derivative calculated by the network without any integration, reducing numerical problems in both cases.

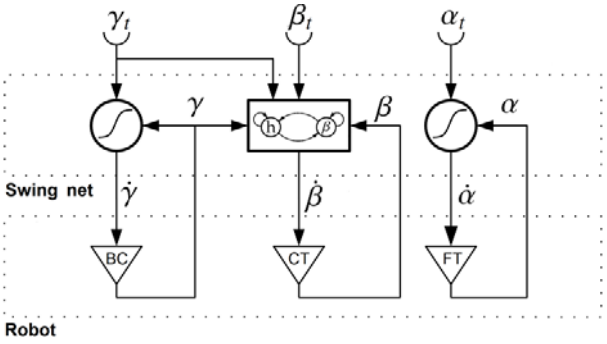


Fig. 4 Swing-net structure

The structure of the network is reported in Fig. 4. The network is composed by three units, one for each joint angle. The α and γ units are composed by a single neuron and the α unit is not interconnected with the other two. The β unit is composed by two interconnected neurons, one of which is completely interconnected also with the γ unit. The mathematical model of the dynamic recurrent ANN can be summarized by the equations:

$$
\begin{aligned}
\tau_\gamma \dot{\gamma} &= w_{\gamma\gamma} \tanh(S_\gamma) + w_{\gamma\gamma_t} \gamma_t \\
\tau_\beta \dot{\beta} &= w_{\beta\gamma} \tanh(S_\gamma) + w_{\beta\beta} \tanh(S_\beta) + \\
&+ w_{\beta h} \tanh(S_h) + w_{\beta\gamma_t} \gamma_t + w_{\beta\beta_t} \beta_t \\
\tau_h \dot{h} &= w_{h\gamma} \tanh(S_\gamma) + w_{h\beta} \tanh(S_\beta) + \\
&+ w_{hh} \tanh(S_h) + w_{\beta\gamma_t} \gamma_t + w_{\beta\beta_t} \beta_t \\
\tau_\alpha \dot{\alpha} &= w_{\alpha\alpha} \tanh(S_\alpha) + w_{\alpha\alpha_t} \alpha_t
\end{aligned}
\tag{1}
$$

remembering that the variable S_i are the weighted sum defined as follows:

$$S_i = y_i + \theta_i \tag{2}$$

where y_i is a general state. The state h, also called hidden state, has been added to correctly reconstruct the behavior of the CT joint angle. The recurrence of the network is guarantee by the leg that acts as an integrator on the three angular velocities produced by the swing net and fed back. The hidden state is integrated numerically.

It's possible to see in **Errore. L'origine riferimento non è stata trovata.** how this CTRNN is able to produce precise trajectories on a wide range of PEPs and AEPs. This network shows also a great robustness since it reacts in a very effective way to external disturbances as shown in Fig. 6.

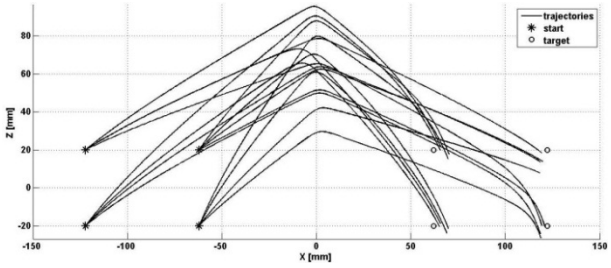


Fig. 5 Swing-net performances

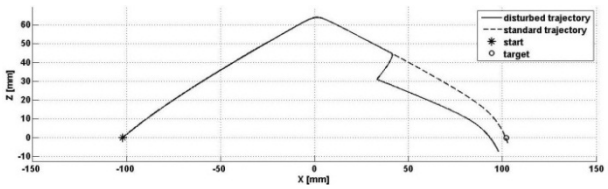


Fig. 6 Swing-net response to disturbance

The control of the stance phase is achieved with a dedicated module realized combining classical controllers with ANN. The module implements both the LPVF controller on the BC and CT joint angles and the other aspects that involve the power stroke, especially the height, velocity and direction control. The Local Positive Velocity Feedback (LPVF) is fundamentally a closed-loop system with an high-pass filter (HPF) on the feedback path and a unitary gain on the forward path. A typical first order high-pass filter has a transfer function that can be expressed as follows:

$$G(s) = \frac{\tau s}{1 + \tau s} \tag{3}$$

where τ is the time constant of the HPF, the inverse of the cutoff frequency. The closed-loop transfer function between input and output becomes:

$$y(s) = (1 + \tau s) x(s) \tag{4}$$

This means that the output of such a system corresponds to the sum of the input and the derivative of the input. In the local positive displacement feedback, feeding directly back the joint angle, the correspondent angular velocity could continuously increase in an unpredictable way. The introduction of the HPF avoids this behavior by maintaining the speed almost constant as required to a stance control system. In this case, for example, a short velocity impulse given at the input, after the pulse, leads to an almost continuously constant velocity output value.

In order to avoid backward walking induced by directional biases, e.g. when gravity opposes the walking direction on a steep slope, a supervisory system has been introduced. It specifies the correct walking direction, but allows also to stop or start the entire system. This is also necessary to avoid any influences of the previous swing phase on the successive stance phase. This formulation allows to give directly an angular velocity command to the actuator. However, the sensors mounted on the joints measure angular displacements only. To obtain an information about the angular velocity it is necessary to numerically calculate the derivative of the measure. In the present work a

second order backward finite difference formula has been used.

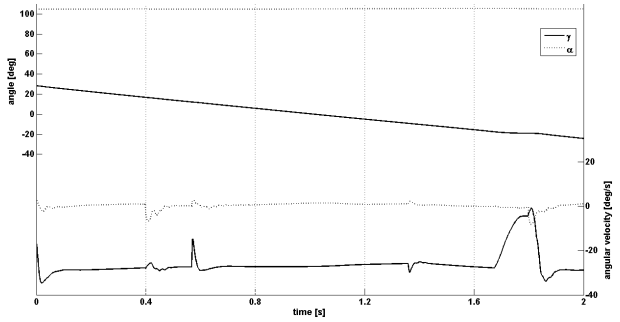


Fig. 7 Stance-net performances

In Fig. 7 the results obtained for the stance phase are shown. During a reference stance phase the most part of the motion depends on the γ angle that move the leg backward, propelling the body forward. The angular velocity shows the behavior previously predicted for a system with LPVF, remaining almost constant after a starting initial perturbation, given by the walking direction offset. The angle α shows small variations only, due to numerical error, but also to the motion of the other legs. This behavior can justify a different control strategy, e.g. a closed-loop negative feedback to increase lateral stability.

The height control of the body depends completely on the CT joint, the γ and α angles are responsible for ground parallel movements and thus they provide the thrust by using LPVF, whereas the β angle doesn't show any reversal reflex, but controls the movement in the direction perpendicular to the ground. This action is performed by the simplest control system possible. The required angular velocity is calculated by a classical closed-loop regulator that uses negative feedback. The actual body height, evaluated at the leg insertion point, is compared to a reference value and the resulting error is transformed into a desired angular velocity. The choice of using a simple proportional controller has its biological counterpart in the findings on height control in stick insects that show an individual proportional element acting as a servomechanism on each leg.

The control law can be summarized as follows:

$$\dot{\beta} = K_{\beta}(h_{sens} - h_{ref}) \tag{5}$$

The results obtained with this controller are illustrated in Fig. 8.

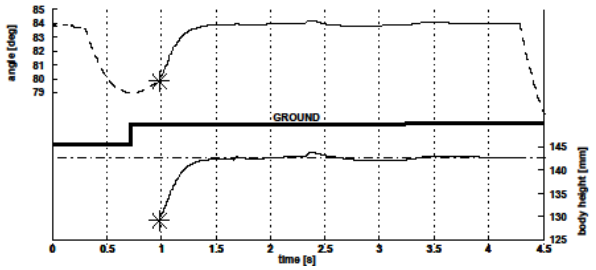


Fig. 8 Height control performances

If the ground has a step (bold line), when the leg passes from swing (dashed line) to stance (solid line), it sees a change in the local body height. The controller tries to maintain the reference body height (dash-dotted line), by feeding-back a command proportional to the error on height.

The value of the current body height isn't directly measured by any sensor but can be easily deduced by the measure of the joint angles. The complete analytical formulation uses twice a trigonometric function, increasing the computational cost. The best way to reduce the computational cost is to approximate the function by an ANN with two outputs, one input and only one neural unit on the hidden layer. The calculation time decrease of almost one order of magnitude.

To maintain a constant advancing speed it is necessary to correct the values of the angular velocities previously calculated. Since the movements on the plane parallel to the ground mainly depend on the γ and α angles, the correction has been applied only on them. The velocity correction works as a classic negative feedback supervisor that modifies the amplification of the positive feedback control signal as follows:

$$\begin{aligned} \dot{\gamma}(k+1) &= K_{\gamma}\, K_{v}\, \dot{\gamma}(k) \\ \dot{\alpha}(k+1) &= K_{\alpha}\, K_{v}\, \dot{\alpha}(k) \end{aligned} \tag{6}$$

where K_{γ} and K_{α} are the fixed static gains of the LPVF controller. To evaluate the gain K_{v} it is necessary to define the error e as the difference

between the required velocity and the measured velocity V_{sens}. The required speed descends from a higher control level, e.g. a human operator or a pattern optimizer, and can be varied over a normalized range from 0 to 1 where 0 means a null speed and 1 indicates the maximum allowed value. The current advancing speed of the single leg can be obtained indirectly from the angular variables as follows:

$$\begin{aligned} V_{sens} &= -((L_c + L_f \sin\beta + \\ &+ L_t \sin(\alpha-\beta))\cos\gamma\dot{\gamma} + \\ &+ (L_f \cos\beta\dot{\beta} + \\ &+ L_t \cos(\alpha-\beta)(\dot{\alpha}-\dot{\beta}))\sin\gamma) \end{aligned} \tag{7}$$

The correction depending on advancing speed can be calculated as follows:

$$K_v = \begin{cases} 1+e & : e \geq 0 \\ \dfrac{1}{1-e} & : e < 0 \end{cases} \tag{8}$$

These nonlinear characteristics, proposed by Schmitz et. al [4], ensures values either bigger or smaller than 1 at the output when walking too slowly or too fast respectively.

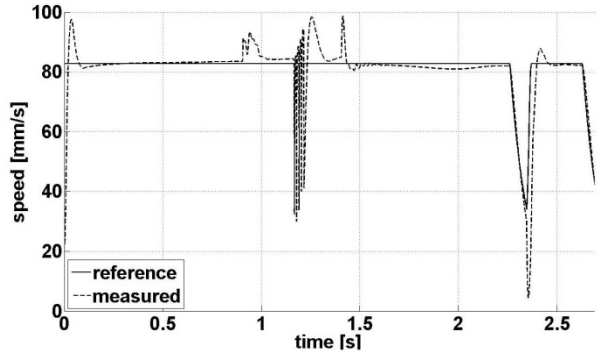


Fig. 9 Velocity control performances

The control of the global velocity accords with the biological findings on stick insects reported by Cruse, who found a velocity control rather than a position control during leg retraction. The results of the velocity control for a typical ripple gait are shown in Fig. 9. The actual trend fits very well the reference value except for some little variation and a sudden oscillation that happens at phase transitions of the leg or adjustments in the body attitude due to the motion of other legs.

The direction control, in the same way as height and velocity one, relies on a negative feedback strategy. Owing to its closed-loop nature, this system can compensate for unbalanced coupling factors or other inequalities between right and left legs. The yaw-turning velocity is assumed to be known. If the yaw reference is set to zero, the system moves straight with small, side-to-side oscillations in heading like all that can be observed in walking insects (see Kindermann [5]). To produce curve walking, a small positive or negative bias is added to the reference value in order to determine the curvature direction and magnitude. This model makes it possible robust course control and performs very good also with very small turning radii, up to 1.5 times the total body length.

The actual yaw angular velocity can be estimated on the basis of the current joint angular variable of the supporting legs, but the analytical formulation of this problem is quite complex and can amplify every error on the measures. To avoid this problem it is necessary to introduce an angular rate sensor.

The state selector is the module deputed to govern the transition between swing and stance phase and thereby to control the rhythm of a leg's motion. An oscillator based on external inputs has been used to control the rhythmic alternation of swing and stance phases. Cruse et al. suggests that this requirement can be fulfilled in the easiest way by a module driven by two sensory input and the information about the activity of both the states. This module can be realized with a simplified four-unit ANN.

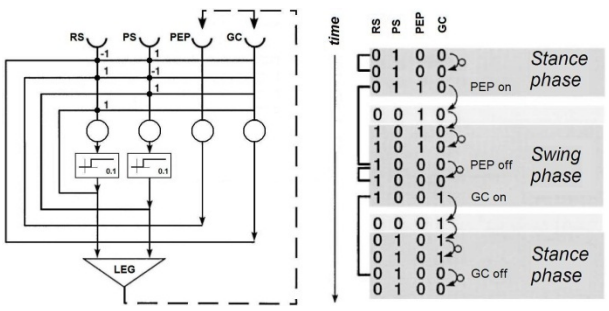


Fig. 10 State Selector structure

In this network, the motor outputs that are active during return stroke and during power stroke are each represented by a single artificial unit (RS and PS in Fig. 10). The activity of a state is indicated by a Boolean truth value that works as switch for the velocities calculation: if the RS is true and the PS is false the angular velocities are given by the swing net, in the opposite case by the stance net.

The sensory inputs are given by two distinct units: a ground contact unit GC, that signals the onset of foot contact with the substrate, and a PEP unit that observes whether the leg has reached a reference posterior extreme position PEP. The GC unit simply consists in the readings of the force measured by the sensor mounted on the foot.

The PEP unit produce its output comparing only the *x* Cartesian coordinate of the PEP and of the foot current position, because it is the most significant. Since the actual position of the foot is not directly measured, it needs to be calculated starting from the current joint angles. This analytical expression, as the body height one, involves trigonometric functions that produces an increase in the computational cost. The solution is the same: replace the exact formulation with an approximating ANN. Even in this case the calculation time becomes one order of magnitude smaller than the original case.

3.2 Reflexes

A reflex is a type of behavior that can be roughly defined as a rapid, automatic involuntary response triggered by external stimuli. The response persists for the duration of the stimulus and its intensity is correlated with the stimulus's strength. Reflexes are very useful during locomotion, making the animal able to react to any unpredictable situation rapidly.

The control system described in this work for the generation of a step cycle in each leg of the robot, is based on reflexes. as it is illustrated above. This approach allows an higher flexibility than for CPG-based systems, however, only a part of the total reactive behaviors shown by a stick insect has been shown so far: an assistance reflex to propel the body and a step reflex to lift-off the leg.

In real animals other reflexes exist, that allow them not only to generate a regular walking pattern, but also to produce the correct behavior in reaction to a large number of unpredictable changes in the environment. Three more different typologies of reflex are always present in insects and have been implemented:

stepping reflex: when a leg assumes a poorly stable posture, performs a limited aerial phase to return in a stable configuration;

avoidance reflex: when a leg touches an obstacle, moves away from it and changes its elevation to overcame it;

searching reflex: when a leg doesn't find the ground at the end of swing phase, starts a rhythmical motion in an increasing region of space until a supporting surface is found.

In Fig. 11 such behaviors are shown.

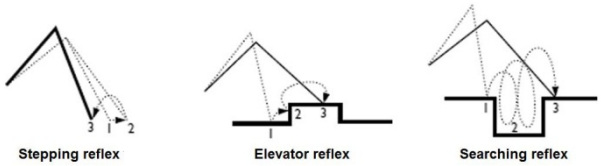


Fig. 11 Reflexes

4 Coordination Strategies

The problem of coordination must consider two things at least. First, the timing of lift-off and touchdown in each leg must not impair the stability of the whole body. Second, the number of legs in stance determines the upper bound of propulsive force. Therefore, efficient timing and coordination of power and stance direction of all legs in stance must be controlled to ensure their synergic action.

These two first requirements can be fulfilled without a central planning. To reach a satisfying solution it is sufficient to insert local influences that link each limb only with its neighbors. This method allows to simplify the problem considerably, reducing computational costs.

Some other problems need a global supervision on all the legs to be correctly solved: the main of them are the control of the body's height and attitude and the generation of curve trajectories. To manage these parameters,

it's necessary both to measure and control variables of all the legs and to perform global calculation.

Walking coordination is generally described using Gaits. The concept of gait is strictly related to the stride concept, that is a period of locomotion defined by the complete cycle of a reference limb. In six-legged walkers the gaits that ensure the best performances can be chosen from biological findings, show a periodic behavior and are called *wave-gaits*. The most important characteristic of these patterns is their capability to ensure a statically balanced motion on a wide speed range.

On really rough terrain, cyclic gaits are not suitable and the *free gait*, a gait that can produce both regular or irregular pattern depending on coordination influences, is used instead.

The coordination rules implemented that regulate the gait in a six-legged walker have been identified by Cruse starting from his experiment on stick insects and can be summarized as six influences among both contralateral and ipsilateral legs. No direct influences between the diagonal legs have been found.

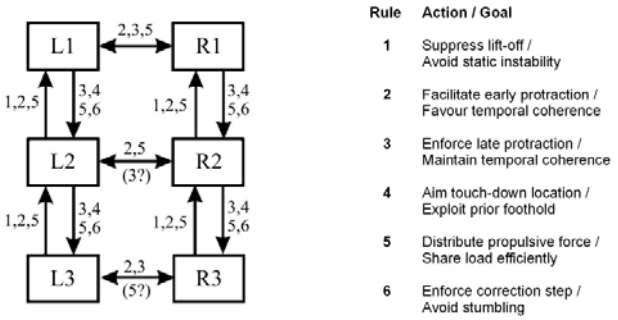


Fig. 12 Coordination rules

A scheme of how this coordinating system works is shown Fig. 12. Each of the six legs signals information about its current state to its neighbors (arrows). The known coordination rules are numbered according to the list to the right. Numbers next to the arrows denote the rules known to be present in a given signaling pathway. The actions/goals correspondence that can be associated with the experimental evidence, is listed aside. The preliminary evaluation of the controller performances have been obtained using a multi-body numerical simulation tool and combining all the developed modules. In Fig. 13 Stepping patterns - ripple the stepping patterns obtained for a ripple wave gait are shown.

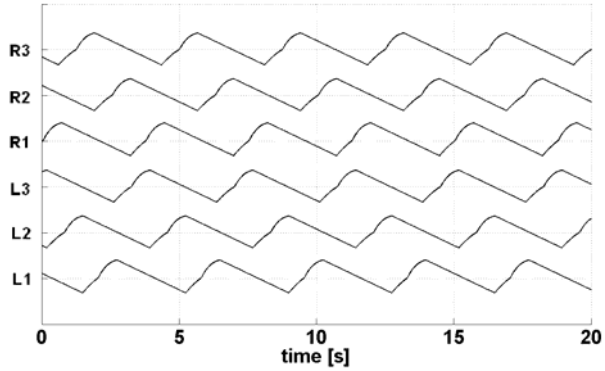


Fig. 13 Stepping patterns - ripple

5 Conclusion

The proposed biomimetic approach to the problem of walking generation and control for Hexapod robot shows promising results both in terms of controller performances and computational costs. The last aspect is very important in the possible application of such techniques to space exploration. The use of CTRNN increase the robustness to external disturbances which suggest a better adaptability to uneven terrain.

References

[1] H. Cruse, V. Dürr, and J. Schmitz. "Insect walking is based on a decentralized architecture revealing a simple and robust controller". *Philosophical Transation of the Royal Society A*, No.365, 2007, pp. 221–250.

[2] R.D. Beer. "On the dynamics of small continuous-time recurrent neural networks". *Adaptive Behavior*, Vol. 3, No. 4, 1995, pp. 469–509.

[3] H. Cruse, T. Kindermann, M. Schumm, J. Dean, and J. T. Schmitz. "Walknet - a biologically inspired network to control six-legged walking". *Neural Networks*, Vol. 11, 1998. pp. 1435–1447.

[4] J. Schmitz, T. Kindermann, M. Schumm, and H. Cruse. "Simplifying the control of a walking hexapod by exploiting physical properties". *European Mechatronics Colloquium, Euromech 375 – Biology and Technology of Walking*, Munich, Germany, March 1998, pp. 296–303.

[5] T. Kindermann. "Behavior and adaptability of a six-legged walking system with highly distributed control". *Adaptive Behavior*, Vol. 9, 2002, pp. 16–41.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Autonomous Capture of Non-Cooperative Spacecraft with a Space Free-Flyer

Cesare Guariniello and Luigi Ansalone and Fabio Curti
DIAEE - Sapienza University of Rome, Italy

Keywords: space, robotics, capture, virtual manipulator

Abstract

This work deals with the problem of performing rendezvous and capture of a non-cooperative spacecraft by means of a space free-flyer, i.e. a satellite base equipped with a robotic manipulator. Though this kind of manoeuvres addresses the solving of relevant existing problems such as debris removal, satellite servicing, orbit changing, only few spaceborne experiments have been conducted, all keeping strong working hypothesis. In this work, a few techniques and strategies to obtain more realistic algorithms, taking into account relative motion and computational load, are presented. An orbit and attitude dynamics simulator has been developed to experiment the proposed strategies.

1 Introduction

Since the launch of the first artificial satellite, Sputnik 1, in 1957, more than 6000 satellites have been put in orbit, 3000 still orbiting Earth together with 12000 space debris.

Many issues may occur due to this enormous population of space waste objects. Deorbiting manoeuvres are not always provided for, or cannot be executed due to failures, thus causing hazards or expensive manoeuvres for incoming spacecraft. Furthermore, satellites may need simple servicing operations, like refueling or deployment of entangled structures, or recovery from failed orbit insertions, or orbit changes that require external help. The main idea that has inspired the present work is the possibility of having a robotic space vehicle equipped with manipulators, a free-flyer, to approach and grab a spacecraft for executing planned operations. While autonomous rendezvous and docking manoeuvres have been developed and realized many times, the use of robotic manipulators in space has always been limited to human-controlled handling and berthing of structures. In this paper a study on the autonomous capture of non-cooperative satellites is presented. A chaser, equipped with a robotic arm, must approach a tumbling non-cooperative target endowed with a grapple fixture and capture it, as shown in Fig. 1.

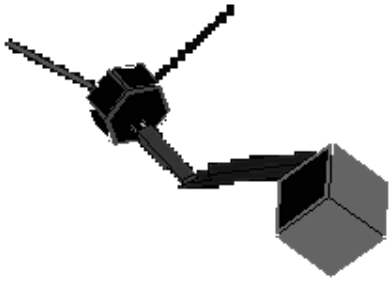


Figure 1: **The free-flyer captures the target (V-bar and R-bar shown)**

The dynamics of the manipulator is quite dif-

ferent from the dynamics of robotic arms with a fixed base; in fact in space the motion of a joint affects the attitude motion of the vehicle on which the manipulator is attached. According to the concept of the Virtual Manipulator and the Virtual Ground, the dynamics of the whole multi-body chaser has to be referred to the center of mass of the system, which moves when external forces are applied and when the capture is executed. Previous works, like [6] and [5], solved this problem using inertial wheels to keep the base attitude fixed with respect to an inertial reference frame, thus simulating a fixed base but wasting three degrees of freedom. The dynamics of a space free-flyer also involves heavy computational loads: an algorithm to reduce the load has been proposed and tested. The problem of the trajectory planning to guide a robotic manipulator to rendezvous and capture a non-cooperative target satellite is closely related to the application of the vision in the robotic capture of moving objects with known dynamics. In the study, the chaser is supposed to have a certain quantity of information about the target, i.e. an estimation of the rotational state and the geometry, the mass and the moment of inertia tensor. A fly-around strategy has been implemented, thus simulating the presence of a vision system apt to get and/or increase the quantity of information before starting the final capture manoeuvre. Capturing strategies for a free-floating space robot in grasping of a tumbling target with model uncertainty are presented. The proposed strategies take into account the relative orbital motion between the chaser and the target, thus dropping the flat-space assumption used in [5].

2 The free-flyer[4]

2.1 Historical Background

Although studied since the early nineties, space robotics has almost entirely been reduced to heavy robotic manipulators attached to much heavier spacecraft (e.g. the *RMS* onboard the Space Shuttle or the Canadarm2 on the *ISS*); this choice allows to avoid the difficulty concerning the kinematics and the dynamics of a space free flyer, because due the heaviness of the base, this can be regarded as fixed with reference to the center of mass of the system. Nevertheless this solution causes two problems: first of all, heavier payloads mean more expensive launches, while a free-flyer can be a medium sized satellite; in addition, large robotic manipulators in space are always human-operated, being the decisions left to astronauts and the computational duties put in the computers' charge.

During last twenty years, many articles have been published about different techniques to perform operations with a free-flyer, but such a wide field of study has never been deeply explored and tested: Japanese satellite ETS-VII ([5]) was the first and up to date the only space-borne experiment that performed rendezvous and capture operations with a space free-flyer. After that, many researchers suggested strategies based on a certain number of hypothesis to make the study easier, but sometimes losing the possible advantages of the free-flyer (e.g., the hypothesis of flat space i.e. lack of relative motion between the free-flyer and the target may lead not to use this relative motion to get a better position of the arm with respect to the grapple fixture).

2.2 Kinematics

To study the motion and the geometry of a space system, is advisable decoupling the equations describing the orbital motion, i.e. the motion of the Center of Mass (CM) of the system, and the equations describing the attitude motion, i.e. the motion around the CM. A space free-flyer, however, is a multi-body system with moving link: therefore, its CM is not a physical point of the system; nevertheless, the position of the CM can be determined as a function of the base attitude and of the joint variables, i.e. the angles of the revolute joints and the shift of the prismatic joints.

2.2.1 The Virtual Manipulator

The free-flyer robotic system is represented by a base whose center of mass CM_0 is identified by means of the vector ${}^c\vec{r}_{c_0}$ with reference to the CM of the whole system. Similarly, ${}^c\vec{r}_{c_i}$ is

the position vector of the center of mass CM_i of the i-th link of the manipulator. These vectors are defined with respect to an in-orbit inertial reference frame RF_c centered in CM.

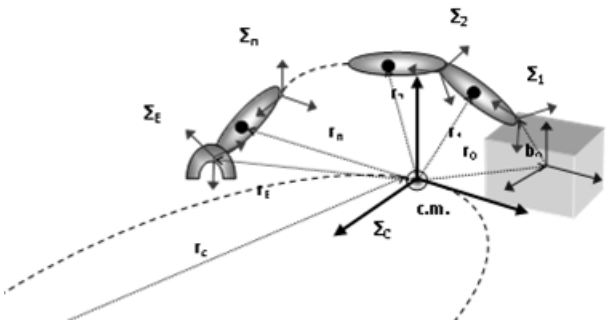


Figure 2: **The free-flyer kinematics**

For each link, in addition to the Denavit-Hartenberg parameters, two more vectors are introduced: ${}^c\vec{a}_i$ is the position of CM_i with respect to the origin of the i-th reference frame RF_i, expressed in RF_c; ${}^c\vec{b}_i$ is the position of the origin of the i+1-th reference frame RF_{i+1} with respect to CM_i, expressed in RF_c.

As for the platform, RF_0 is the reference frame centered in CM_0 and having its axes along the principal axes of inertia of the platform.

According to the given definitions of the position vectors, the following relations between the ${}^c\vec{r}_{c_i}$ can be written:

$$\begin{cases} {}^c\vec{r}_{c_1} = {}^c\vec{r}_{c_0} + {}^c\vec{b}_0 + {}^c\vec{a}_1 \\ \vdots \\ {}^c\vec{r}_{c_i} = {}^c\vec{r}_{c_{i-1}} + {}^c\vec{b}_{i-1} + {}^c\vec{a}_i \\ \vdots \\ {}^c\vec{r}_{c_n} = {}^c\vec{r}_{c_{n-1}} + {}^c\vec{b}_{n-1} + {}^c\vec{a}_n \end{cases} \quad (1)$$

Such equations are recursive, ${}^c\vec{r}_{c_0}$ being the only unknown quantity.

Let m_0 be the base mass and m_i the mass of the i-th link; being the ${}^c\vec{r}_{c_i}$ referred to the CM, the static moment of the whole system is equal to zero:

$$m_0\ {}^c\vec{r}_{c_0} + \ldots + m_i\ {}^c\vec{r}_{c_i} + \ldots + m_n\ {}^c\vec{r}_{c_n} = 0 \quad (2)$$

Given the Eq. (1),

$${}^c\vec{r}_{c_0} = \sum_{i=0}^{n} c_i^{(0)} \left({}^c\vec{b}_i + {}^c\vec{a}_{i+1} \right) \quad (3)$$

being

$$M = \sum_{i=0}^{n} m_i \quad (4)$$

$$c_i^{(0)} = \frac{\sum_{k=0}^{i} m_k}{M} - 1 \quad (5)$$

Calling

$$c_i = \frac{\sum_{k=0}^{i} m_k}{M} \quad (6)$$

the position of each CM can be calculated as

$${}^c\vec{r}_{c_k} = \sum_{i=0}^{n} c_i^{(k)} \left({}^c\vec{b}_i + {}^c\vec{a}_{i+1} \right) \quad (7)$$

being

$$c_i^{(k)} = \begin{cases} c_i & i < k \\ c_i^{(0)} = c_i - 1 & i \geq k \end{cases} \quad (8)$$

Since the position of the end effector (EE) is ${}^c\vec{r}_{EE} = {}^c\vec{r}_{c_n} + {}^c\vec{b}_n$, the expression of the *direct kinematics* of the space free-flyer is

$${}^c\vec{r}_{EE} = \sum_{i=0} nc_i \left({}^c\vec{b}_i + {}^c\vec{a}_{i+1} \right) \quad (9)$$

Comparing the expression of the direct kinematics with the geometric path given by ${}^c\vec{r}_{EE} = {}^c\vec{r}_0 + {}^c\vec{b}_0 + {}^c\vec{a}_1 + {}^c\vec{b}_1 + \ldots + {}^c\vec{a}_n + {}^c\vec{b}_n$, it can be noticed that the position of the EE with respect to the CM of the whole system is equivalent to that of a virtual manipulator with the base located in CM and links parallel to that of the real manipulator, but having the lengths ${}^c\vec{a}_i$ and ${}^c\vec{b}_i$ scaled by the coefficients c_i. In this virtual description, the base of the real manipulator is represented by means of a spherical joint.

2.3 Differential Kinematics

The linear velocity of the EE, ${}^c\vec{v}_E$ can be expressed as a function of the linear and angular velocities of the base, ${}^c\vec{v}_{c_0}$ and $\vec{\omega}_0$, and of the velocity of each joint $\dot{q}_i$ as follows:

$${}^c\vec{v}_E = {}^c\vec{v}_{c_0} + \vec{\omega}_0 \times ({}^c\vec{r}_{EE} - {}^c\vec{r}_{c_0}) + \sum_{i=1}^{n} J_{L_i} \dot{q}_i \quad (10)$$

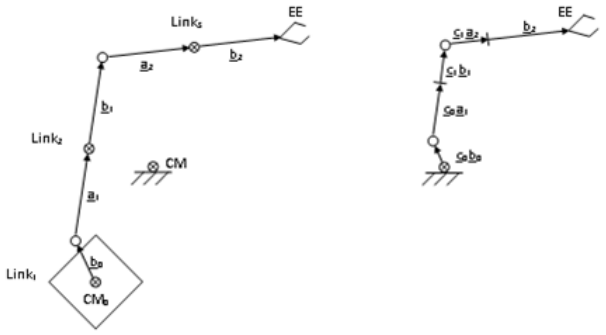


Figure 3: **The virtual manipulator**

being J_{L_i} column arrays based on the joint type:

$$J_{L_i} = \begin{cases} {}^c\hat{z}_i & Pris \\ {}^c\hat{z}_i \times ({}^c\vec{r}_{EE} - {}^c\vec{r}_i) & Rev \end{cases} \quad (11)$$

where ${}^c\vec{r}_i$ is the position of the origin or RF_i with respect to the origin of RF_c, i.e. ${}^c\vec{r}_i = {}^c\vec{r}_{c_{i-1}} + {}^c\vec{b}_{i-1}$.
Let be ${}^c\vec{r}_{OE} = {}^c\vec{r}_{EE} - {}^c\vec{r}_{c_0}$ and let also be

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \Rightarrow \tilde{v} = \vec{v}\times = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (12)$$

Then the linear velocity of the *EE* can be written as

$${}^c\vec{v}_E = {}^c\vec{v}_{c_0} - {}^c\tilde{r}_{OE}\vec{\omega}_0 + J_L\dot{\vec{q}} \quad (13)$$

being

$$J_L = \begin{bmatrix} J_{L_1} & \dots & J_{L_n} \end{bmatrix} \qquad \dot{\vec{q}} = \begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \vdots \\ \dot{q}_n \end{bmatrix} \quad (14)$$

To calculate ${}^c\vec{r}_{c_0}$, assuming that *no external force is acting on the free-flyer*, the conservation of momentum gives, starting with initial momentum equal to zero, $m_0\,{}^c\vec{v}_{c_0} + m_1\,{}^c\vec{v}_{c_1} + \dots + m_n\,{}^c\vec{v}_{c_n} = 0$.

Since for the i-th link

$${}^c\vec{v}_{c_i} = {}^c\vec{v}_{c_0} + \vec{\omega}_0 \times ({}^c\vec{r}_{c_i} - {}^c\vec{r}_{c_0}) + J^{(i)}_{L_c}\dot{\vec{q}} \quad (15)$$

where the columns of $J^{(i)}_{L_c}$ are

$$J^{(i)}_{L_{c_j}} = \begin{cases} {}^c\hat{z}_j & j \le i \quad and\ prism \\ {}^c\hat{z}_j \times ({}^c\vec{r}_{c_i} - {}^c\vec{r}_j) & j \le i \quad and\ rev \\ \vec{0} & j > i \end{cases} \quad (16)$$

Due to the conservation of momentum and setting ${}^c\vec{r}_m = \frac{1}{M}\sum_{i=1}^n m_i\vec{\omega}_0 \times ({}^c\vec{r}_{c_i} - {}^c\vec{r}_{c_0})$, the final result is

$${}^c\vec{v}_E = \left(J_L - \frac{1}{M}\sum_{i=1}^n m_i J^{(i)}_{L_c}\right)\dot{\vec{q}} + ({}^c\tilde{r}_m - {}^c\tilde{r}_{OE})\,\vec{\omega}_0 \quad (17)$$

As for the angular velocity of the *EE*,

$${}^c\vec{\omega}_E = {}^c\vec{\omega}_0 + \sum_{i=1}^n J_{A_i}\dot{q}_i = {}^c\vec{\omega}_0 + J_A\dot{\vec{q}} \quad (18)$$

where the columns of J_A are defined as follows, basing on the joint type:

$$J_{A_i} = \begin{cases} \vec{0} & Prism \\ {}^c\hat{z}_i & Rev \end{cases} \quad (19)$$

Thus, the final expression for the *differential kinematics* of the free-flyer is

$$\begin{bmatrix} {}^c\vec{v}_E \\ {}^c\vec{\omega}_E \end{bmatrix} = J_{FF}\begin{bmatrix} \dot{\vec{q}} \\ \vec{\omega}_0 \end{bmatrix} \quad (20)$$

being J_{FF} the Jacobian of the free-flyer:

$$J_{FF} = \left[\begin{array}{c:c} \underbrace{J_L - \frac{1}{M}\sum_{i=1}^n m_i J^{(i)}_{L_c}}_{J_A} & \underbrace{{}^c\tilde{r}_m - {}^c\tilde{r}_{OE}}_{I_{3x3}} \end{array}\right] \quad (21)$$

2.4 Dynamics

The equations of the free-flyer dynamics can be derived using an approach based on the lagrangian mechanics. Assuming that there is not any conservative forces acting on the spacecraft (i.e., neglecting the effects of gravity gradient), the lagrangian term is only the kinetic energy T of the free-flyer:

$$T = \frac{1}{2}\sum_{i=0}^{n}\left({}^{c}\vec{\omega}_i^T \; {}^{c}I_i \; {}^{c}\vec{\omega}_i + m_i \; {}^{c}\vec{v}_{c_i}^T \; {}^{c}\vec{v}_{c_i}\right) \quad (22)$$

The expression for ${}^{c}\vec{v}_{c_i}$ is described in Eq. (15). The angular velocity of the i-th link is ${}^{c}\vec{\omega}_i = {}^{c}\vec{\omega}_0 + J_A^{(i)}\dot{\vec{q}}$, where the columns of $J_A^{(i)}$ are defined as follows, basing on the joint type:

$$J_{A_j}^{(i)} = \begin{cases} \vec{0} & j \leq i \quad \textit{and prismatic} \\ {}^{c}\hat{z}_j & j \leq i \quad \textit{and revolute} \\ \vec{0} & j > i \end{cases} \quad (23)$$

Taking into account the equations of differential kinematics,

$$T = \frac{1}{2}\begin{bmatrix} \vec{\omega}_0^T & \dot{\vec{q}}^T \end{bmatrix}\begin{bmatrix} H_{\omega\omega} & H_{\omega q} \\ H_{q\omega} & H_{qq} \end{bmatrix}\begin{bmatrix} \vec{\omega}_0 \\ \dot{\vec{q}} \end{bmatrix} \quad (24)$$

being H_{FF} the *generalized inertia matrix* of the space free-flyer.

The blocks $H_{\omega\omega}$, $H_{\omega q}$, $H_{q\omega}$ and H_{qq} can be calculated using the equations of the kinematics and differential kinematics.

The classical lagrangian approach then allows to calculate the evolution of the dynamics, given the torques $\vec{\tau}$ acting on the spacecraft:

$$\frac{d}{dt}\left[\frac{\partial T}{\partial \dot{\vec{q}}}\right] - \frac{\partial T}{\partial \vec{q}} = \vec{\tau} \quad (25)$$

However, in the free-flyer dynamics equations the base angular velocity $\vec{\omega}_0$ constitutes a non-holonomic constraint (causing the possibility to get to a previous configuration without reaching the previous state), thus not integrable.

A different approach, based on the so-called *quasi-coordinates* ([3]) allows to solve the problem of the derivation of the lagrangian term.

2.4.1 Quasi-Coordinates Lagrangian Approach

When the equations of motion are not restricted to *true coordinates* (e.g., the derivatives of the joint variables), they use non integrable variables, called *quasi-coordinates*, that can be expressed as

$$\omega_s = \alpha_{1s}\dot{q}_1 + \alpha_{2s}\dot{q}_2 + \ldots + \alpha_{ns}\dot{q}_n = \sum_{i=1}^{n}\alpha_{is}\dot{q}_i \quad (26)$$

where the coefficients α_{rs} are function of the generalized coordinates q. In the matrix form, $\vec{\omega_{ql}} = \alpha^T\dot{\vec{q}}$. Assuming that the matrix α is not singular, $\dot{\vec{q}} = \beta\vec{\omega_{ql}}$ with $\beta\alpha^T = I$.

Calculating the relations between the kinetic energy as function of the true coordinates and the kinetic energy T as function of the quasi-coordinates allows to write the equations of dynamics in the integrable form

$$\begin{aligned}\alpha\frac{d}{dt}\left[\frac{\partial T}{\partial \vec{\omega}_{ql}}\right] + \left[\vec{\omega}_{ql}^T\beta^T\left\{\frac{\partial \alpha}{\partial \vec{q}}\right\}\right]\frac{\partial T}{\partial \vec{\omega}_{ql}} - \frac{\partial T}{\partial \vec{q}} + \\ - \left[\vec{\omega}_{ql}^T\beta^T\left[\frac{\partial \alpha}{\partial \vec{q}}\right]\right]\frac{\partial T}{\partial \vec{\omega}_{ql}} = \vec{\tau}\end{aligned} \quad (27)$$

where $\left\{\frac{\partial \alpha}{\partial \vec{q}}\right\}$ does not involve summation over the indices of α while $\left[\frac{\partial \alpha}{\partial \vec{q}}\right]$ does.

The computational load of such a calculation is huge, involving inversion of 6-by-6 matrices.

2.5 Inverse Kinematics

When designing a robotic manipulator control system, there is a fundamental choice between control in the joint space and control in the task space; nevertheless, even when choosing a centralized control based on the whole dynamics of the system, i.e. in the task space, the desired trajectory is converted in joint positions, velocities and accelerations.

The inverse kinematics allow to calculate the joint variables, given a desired pose of the *EE*. While the inverse differential kinematics is based on a simple inversion of a matrix, the inverse kinematics is not always easily derivable in an analytical form; thus, a numeric algorithm has to be implemented to realize the inverse kinematics.

The numeric algorithms are based on the analytical Jacobian matrix of the direct kinematics map $\vec{r} = f_r(\vec{q})$:

$$J_r(\vec{q}) = \frac{\partial f_r(\vec{q})}{\partial \vec{q}} \quad (28)$$

The easiest way to calculate the analytical Jacobian matrix, however, is through the geometric Jacobian J_G, which is the matrix mapping the derivatives of the joint variables into the linear and angular velocities of the *EE*:

$$\begin{bmatrix} \dot{\vec{r}}_{EE} \\ \vec{\omega}_E \end{bmatrix} = J_G \left[\dot{\vec{q}}\right] \tag{29}$$

The Jacobian Matrix J_{FF} in Eq. (21) maps the derivatives of the joints variables and the angular velocity of the base into the linear and angular velocities of the *EE*. If we choose a parametric representation of the attitude of the base, e.g. by means of a set of Euler angles $\vec{\phi}$, the relation between angular velocity and derivatives of the Euler angles can be expressed as $\vec{\omega}_0 = T(\vec{\phi})\dot{\vec{\phi}}$ and for the *EE* $\vec{\omega}_E = T(\vec{\phi}_{EE})\dot{\vec{\phi}}_{EE}$. Thus a relation between J_{FF} and J_G is given by

$$\begin{bmatrix} \dot{\vec{r}}_{EE} \\ \vec{\omega}_E \end{bmatrix} = J_{FF} \begin{bmatrix} \dot{\vec{q}} \\ \vec{\omega}_0 \end{bmatrix} = J_{FF} \begin{bmatrix} I & 0 \\ 0 & T(\vec{\phi}) \end{bmatrix} \begin{bmatrix} \dot{\vec{q}} \\ \dot{\vec{\phi}} \end{bmatrix} \tag{30}$$

Thus, $J_G = J_{FF} \begin{bmatrix} I & 0 \\ 0 & T(\vec{\phi}) \end{bmatrix}$.

Similarly, the relation between J_G and J_r is function of $T(\vec{\phi}_{EE})$. In conclusion

$$J_r = \begin{bmatrix} I & 0 \\ 0 & T^{-1}(\vec{\phi}_{EE}) \end{bmatrix} J_{FF} \begin{bmatrix} I & 0 \\ 0 & T(\vec{\phi}) \end{bmatrix} \tag{31}$$

2.5.1 Newton Method

The first numeric method to calculate the inverse kinematics is the Newton method, based on the linearization of the evolution of the direct kinematics $\vec{r} = f_r(\vec{q}) = f_r(\vec{q}^k) + J_r(\vec{q}^k)(\vec{q} - \vec{q}^k) + \circ(\|\vec{q} - \vec{q}^k\|^2)$. The step of the algorithm thus is

$$\vec{q}^{k+1} = \vec{q}^k + J_r^{-1}(\vec{q}^k)\left[r - f_r(\vec{q}^k)\right] \tag{32}$$

The convergence can be reached only if the initial guess is close enough to the real solution; in this case the algorithm has quadratic convergence rate.

2.5.2 Gradient Method

Another method is based on the minimization of the quadratic error and the step is

$$\vec{q}^{k+1} = \vec{q}^k + \alpha J_r^T(\vec{q}^k)\left[r - f_r(\vec{q}^k)\right] \tag{33}$$

The scalar parameter α has to be chosen so as to guarantee a decrease of the error function at each iteration: too large values may lead the algorithm to miss the minimum; on the other side, if α is too small, the convergence is extremely slow. This algorithm never diverges and is computational simpler than the other.

An efficient iterative scheme has been devised by combining initial iterations with Gradient method, sure but slow, then switch to Newton method (quadratic final convergence rate). Choices to be made concern the initial guess, the step size in gradient method and the stopping criteria (cartesian error lower than a certain limit or algorithm increment lower than a certain limit).

3 Proposed strategies and results

The proposed mission scenario involves a target and a chaser orbiting on circular LEO orbits on the same plane at different height; the state of the chaser free-flyer is supposed to be known. The free-flyer is equipped with a 3-links robotic arm.

3.1 Rendezvous Strategy

Faraway operations are similar to those of an automated rendezvous and docking mission. Therefore, standard phasing manoeuvres are used: orbital plane corrections, Hohmann transfers, R-bar transfers ([2]). Nevertheless, the proposed strategy includes two fundamental characteristics that make it quite different from a standard rendezvous manoeuvre:

- while the first phase of manoeuvres is executed using *AGPS*, standard homing manoeuvres are not suitable to get closer to a non-cooperative spacecraft, because there is no possibility to use relative positioning. Such manoeuvres need to be performed

by means of absolute positioning (knowing orbital parameters about the object to be chased) and, when getting closer to the target, by visual systems;

- the trajectory planner needs to know not only relative position and velocity (in order to calculate relative orbital motion), but also the state of motion of the target, in addition to the position of the grapple fixture, when present aboard the target. To achieve the goal of getting information about the state of motion, the chaser will perform a fly-around manoeuvre.

The proposed rendezvous strategy can therefore be describes as follows:

1. the chaser performs orbital correction manoeuvres to get the orbital plane of the target and drifts in order to phase with the target;
2. the chaser performs height change transfer manoeuvres in order to reach a station keeping point a few hundreds of meters behind the target;
3. the chaser performs an R-bar transfer or a V-bar transfer in order to reach a station keeping point a few meters behind the target, at a distance suitable to use the motion estimation visual system;
4. the chaser performs a complete fly-around manoeuvre in order to get starting information about the state of motion of the target. When come back to the starting point of the manoeuvre, capture strategies can be put into execution.

Performing a closing transfer while neglecting the perturbation effects, may lead to a quite large error, as shown in Fig. 4.

3.2 Capture Strategies

Due to the complexity of the many variables involved in an autonomous capture mission, a great number of strategies could be implemented and experimented; nevertheless, a few cases of interest have been identified and studied during the development of this thesis:

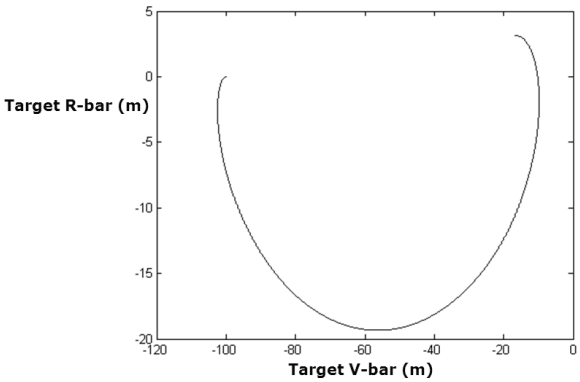


Figure 4: **Closing manoeuvre trajectory with perturbations**

3.2.1 Free-flyer *CM* and Grapple Fixture at Similar Heights

When the state of motion of target causes the grapple fixture to be in such a position that exists a point whence the robotic arm could reach the grapple fixture without moving the free-flyer *CM* out of the target orbit during the *final capture*, the best choice is to manoeuvre the free-flyer as to perform the final capture leaving unaltered the *CM* position with respect to the target.

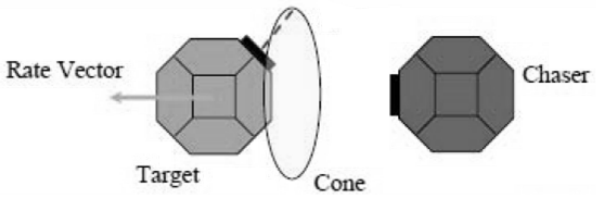


Figure 5: **Direct final capture**

3.2.2 Free-flyer *CM* and Grapple Fixture at a Different Height

When the state of motion of the target causes the grapple fixture to be in such a position that the robotic arm cannot reach it without taking the free-flyerCM out of the target orbit (e.g., the grapple fixture is on the rotation axis or coning near it and far from the H-bar - V-bar plane), the suggested strategy involves a transfer manoeuvre able to exploit the relative motion to

carry the free-flyer *CM* to a point whence the grapple fixture can be reached. The trajectory planner must take into account the relative motion to calculate the required pose and velocity of the *EE* with respect to the grapple fixture; during the execution of the relative motion, the manipulator is controlled in order to reach the desired goal.



Figure 6: **Final capture with relative motion**

3.3 Trajectory Planner

The trajectory planner is activated before the beginning of the final capture manoeuvre and whenever the trajectory needs to be recalculated. The planner needs three inputs:

- current state of the joint variables and of the attitude of the base: determinable by means of proprioceptive sensors;
- manoeuvre execution time: function of the manipulator physical limitations, it must be long enough to guarantee the possibility of replanning. It also can be calculated in function of optimization criteria, taking into account the grapple fixture motion and the relative motion between the chaser and the target;
- final state of the joint variables, of the attitude of the base and of the joint velocities required to execute the capture; they are determined as follows: the relative motion of the spacecraft and the manoeuvre execution time allows to calculate the position and velocity of the *EE* with respect to the free-flyer *CM* required to execute the capture; required velocity of the joints can be obtained by a simple inversion of the geometric Jacobian matrix, while the required joint variables can be calculated through kinematic inversion.

3.3.1 Kinematic Inversion

As stated in section 2.5, an important stress must be given to the choice of the numeric algorithm to use to calculate the inverse kinematics by means of the analytical jacobian: the experiments have in fact shown that the parameter for the gradient method has to be chosen little enough as to guarantee convergence of the algorithm, but this entails very slow convergence, sometimes more than 100 iteration steps and various minutes of computation, as shown in Fig. 7

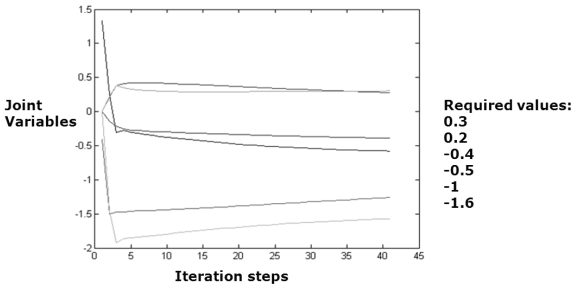


Figure 7: **Gradient Method**

on the other side, the Newton method can sometimes reach the solution with an error lower that 10^{-15} m in less than 10 steps, but can also lead to divergence, as shown in Fig. 8

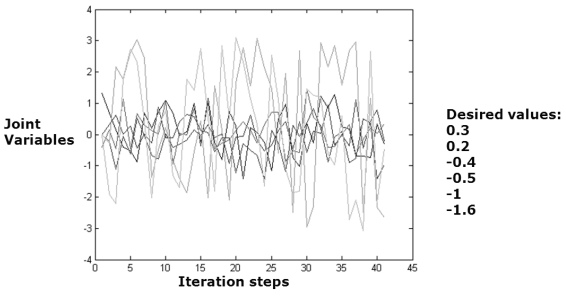


Figure 8: **Newton Method**

thus, an hybrid algorithm has been used: the gradient method, slow but assuring convergence,

is initially used to shift from the current variables, used as the first guess, to variables closer to the required solution; at this point, the Newton method starting from the last step of the gradient algorithm, guarantees convergence at quadratic rate in a few steps (Fig. 9).

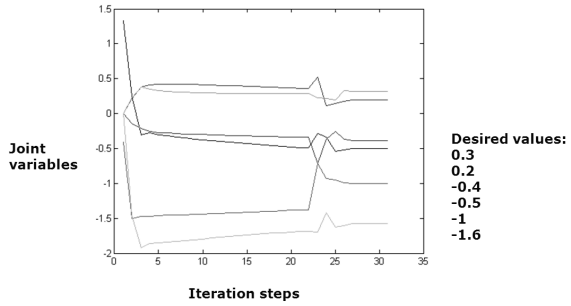


Figure 9: **Hybrid Method**

3.3.2 Planning

Knowing the initial position and velocity and the required final position and velocity for each joint variable, a cubic polynomial represents a solution for this problem.

3.4 Control Laws

After having calculated the desired trajectory for the joints, the free-flyer needs a controller to determine the torques to apply on the base and on the joints in order to execute the desired movement and to perform the capture. Two main families of controllers have been used in this work.

3.4.1 Joint Space Decentralized Control

The use of quasi-coordinates to calculate the dynamics of the manipulator leads to an expression having the form $H(q)\ddot{q} + C(q, \dot{q}, \omega)\dot{q} = \tau$. A controller in the joint space will use only the linear part of the dynamics equations, thus considering the matrix C, i.e. the contribution due to the configuration, as a disturbance. Each joint is controlled as a Single Input Single Output (*SISO*) system; the lack of the gravity term allows to use a simple proportional controller, with position and velocity feedback. The performances of such controllers decay when the desired trajectory is too fast; derivative of the desired motion to compensate the disturbance with a feedforward predictor may lead to improvement of this kind of controller. A delay in the expected trajectory is always present.

3.4.2 Centralized Control

This kind of controller takes into account the non-linear coupling between the joints, considering the free-flyer as a *MIMO* system. The so-called inverse dynamics control allows to decouple and linearize the original system $H(q)\ddot{q} + C(q, \dot{q}, \omega)\dot{q} = u$, with input $u = H(q)y + C(q, \dot{q}, \omega)\dot{q}$ depending on the manipulator state. Choosing a new input $y = -K_P q - K_D \dot{q} + r$, the resulting system is asymptotically stable and decoupled if the matrices K_P and K_D are symmetric and positive-definite. The desired trajectory following is guaranteed by the choice of $r = \ddot{q}_d + K_D \dot{q}_d + K_P q_d$.

Such kind of controllers are suitable to be made robust or adaptive.

3.5 Execution of Manoeuvres

Since the complete dynamics of the free-flyer manipulator is computationally hard to handle, in this work a decoupling between the base attitude motion and the joint motion has been proposed: when the trajectory planner calculates the desired trajectory, the joint motion is simulated in absence of platform control; this technique allows to calculate the base rotation due to the joint movement. If only the joints would rotate, the base would get an attitude R_x. The system then computes the *over-rotation* between the current attitude of the base and R_x and the base is actuated with such torques as to get to the desired final attitude minus the calculated over-rotation. Having the free-flyer moved as a whole body (the joints were not actuated during the base motion), the relative position of the joints with respect to the base has not changed: thus, realizing the previously simulated joint movement will add the over-rotation, taking the base to the desired attitude. The to-

tal manoeuvre time of the joints must be scaled down to take into account the separate base rotation and it has been experimented that this scaling does not affect the final attitude of the base.

3.6 Results

As an example of the obtained results, two tables containing the position error (centimeters) of the *EE* at the end of the capture manoeuvres are shown.

Table 1: Decentralized control

	Base case	Perturb.	Fast dyn
Fixed target	0.3 (delay)	1.2 (delay)	0.8 (delay)
V-bar rot	2.2 (delay)	2.5 (delay)	7.1 (delay)
H-bar rot	2.1 (delay)	3.8 (delay)	8.4 (delay)

Table 2: Centralized control with 2 steps

	Base case	Perturb. no feedback	Exstim. errs. no feedback
Fixed target	0.1	4.6	0.8
V-bar rot	1.2	5.0	2.0
H-bar rot	1.2	6.1	5.5

3.7 Comparison

Experimental comparison of various algorithms provides that:

- decentralized algorithms is computationally the lightest, but requires a sampling dynamics faster than the system dynamics and provides delay;
- centralized algorithms are stable due to feedback, but the inversion of the complete dynamics equations (involving inversion of 6-by-6 matrices) are impossible to be used in real-time.
- centralized algorithms with the proposed 2-step manoeuvre are computationally lighter than the complete dynamics inversion (involving two 3-by-3 matrices inversion) and achieve good results; such kind of manoeuvres are more sensible to rapid dynamics than the previous algorithm, since they reduce the capture manoeuvre time.

4 Conclusion

In conclusion, a rendezvous and capture mission with a free-flyer has been studied, designed and tested with a newly implemented orbit and attitude dynamics multisatellite simulator. A few problems of previous studies and mission have been addressed to and solved, such as kinematics inversion, use of relative motion, dynamics computational load. Future studies will involve optimization criteria for the strategies, robust and adaptive control and multiple-arms free-flyers, e.g. as proposed in [1] and [4].

References

[1] Dimitar Dimitrov. *Dynamics and Control of Space Manipulators During a Satellite Capturing Operation*. PhD Dissertation Thesis - Tohoku University, 2005.

[2] Wigbert Fehse. *Automated Rendezvous and Docking of Spacecraft*. Cambridge University Press, 2003.

[3] Leonard Meirovitch. *Methods of Analytical Dynamics*. McGraw-Hill, 1970.

[4] Evangelos Papadopoulos e S. Ali A. Moosavian. *Dynamics and Control of Space Free-Flyers with Multiple Manipulators*. Department of Mechanical Engineering and Centre for Intelligent Machines, McGill University.

[5] Kazuya Yoshida. *Space Robot Dynamics and Control: To Orbit, From Orbit, and Future*. Robotics Research: the 9th International Symposium, 2000.

[6] Kazuya Yoshida, Dimitar Dimitrov, e Hiroki Nakanishi. *On the Capture of Tumbling Satellite by a Space Robot*. Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Bidirectional Converter for Single-Cell Li-Ion Batteries in a Moon Rover

M. Macellari and L. Schirone
La Sapienza Università di Roma – DIAEE
Via Salaria 851, 00138 – Roma – Italy
luigi.schirone@uniroma1.it
Tel. 0644589762

Keywords: *Power conversion, Bidirectional, Li-ion, BCDR, Moon Rover*

Abstract

This work deals with a circuit intended to operate as a BCDR in small space vehicles, and especially in the cases when battery sizing is driven by the need to complement the power provided by primary power sources during high peaks of load, while a comparatively small energy is required during eclipses. In particular, a Bidirectional Three Level Neutral Point Clamped (TLNPC) converter was selected, designed and built in a prototype capable to operate as a BCDR to interface a 12V bus with a single-cell Li-ion battery.

After providing a rationale of the single-battery-cell approach, and of the criteria adopted to select the TLNPC converter circuit, the paper reports principles of operation of the converter and indicates design criteria for the main circuit parameters. Extensive characterization of the laboratory prototype completes the work.

1 Introduction

Energy storage systems normally consist of several low-voltage cells, connected in series strings in order to provide battery voltages suitable to be raised to suitable dc-link voltages by means of bidirectional converters [1-7], operating both in discharging and charging modes.

High voltage strings are beneficial for the system efficiency, but slight mismatches or temperature differences cause charge imbalance when the series string is charged as a unit [8]. This results in increased stress on the batteries and reduces the operating lifetime. Charge equalization cycles [9,10] or complicated single-cell management systems [11] can be used to mitigate the effects of imbalances, but a more effective approach to reduce imbalances is to diminish the number of cells in a string [12], in the framework of a modular storage system. Therefore, it is important to use a bidirectional dc–dc converter capable to manage an increased voltage diversity and to be easily connected in parallel to similar units.

The most used cell technology is lithium–ion (Li-Ion) batteries, providing an energy density larger than 150 Wh/kg, but highly demanding from the point of view of the control of charge and discharge conditions. Battery voltage varies more than 30%, depending on the state of charge.

A rover for Moon exploration will be continuously exposed to solar radiation throughout the 354 hours of lunar day. During daytime it only could experience short eclipses

just after deployment, when it could fall under the shadow projected from the lander, or when it approaches a boulder while Sun elevation is low. The secondary energy source is accordingly required to provide full power for a short time, of the order of a few tens of minutes. On the other hand, typical Li-ion cells for space applications are characterized by a limited power/energy ratio so that the required power can only be achieved when an excess energy has been taken on board.

Let us consider a rover with 100W power absorption and a 60 minutes maximum eclipse duration. The typical current/capacity ratio is around 1 A/Ah, so that 100W maximum power corresponds to nearly 30 Ah capacity. This could be either achieved by a 5S1P battery of SAFT MPS, or a 6S3P battery of Sony 18650, or a single (1S1P) SAFT VES 100 cell. Alternatively the single cell could be a GS LSE 50, slightly oversized. As can be observed in Tab. 1, by far the most lightweight solution is provided by a single VES 100 cell.

Table 1

	SAFT VES100	SAFT MPS	SONY 18650	GS LSE 50
Config.	1S1P	5S1P	6S3P	1S1P
Power (W)	100	100	100	180
Capacity (Ah)	27	5.6	1.5	50
Mass (g)	800	1500	860	1500
Volume (cm^3)	520	600	480	800
Voltage (V)	2.7 - 4.2	13.5-20.5	16.2-24.6	2.7 - 4.2

Additional benefits are provided as a single-cell battery would be very simple both from the fabrication and management points of view: it would remove any issue about cell balancing, while the harness reduction would save mass. As a counterpart, an high-voltage-diversity bidirectional converter is required as a bus interface.

This is normally obtained by means of transformer-based (insulated) switching DC-DC converters [3, 13-15]. Furthermore, in such converters, wide voltage variations complicate the transformer design, resulting in a reduction of conversion efficiency. Thus, best conversion efficiencies are obtained by using transformerless (uninsulated) configurations [16-24]. Indeed, some of this converters can operate in such a way to transfer part of the energy directly to the load (Direct Energy Transfer - DET), without the dissipation associated to a storage phase in the magnetic parts, always necessary in a transformer-based converter. Further dissipation reduction are associated to a higher flexibility of the design of Transformerless converters in the presence of large variation of the input voltage.

Among the proposed configurations, the requirement of high voltage diversity limits our interest to either the circuit proposed by [12] or the Three Levels Neutral Point Clamped (TLNPC) circuit [25]. The TLNPC circuit provides optimum efficiency when operating with a Voltage Ratio VR = 4, but it can comfortably operate in the range VR = 2.5÷6. This makes it a perfect candidate to interface a single-cell Li-Ion battery (V = 2.7÷4.2 V) to a 12 V bus. For this application the larger circuit complexity of the converter in [12] is not justified as it is associated with larger dissipation and lower reliability.

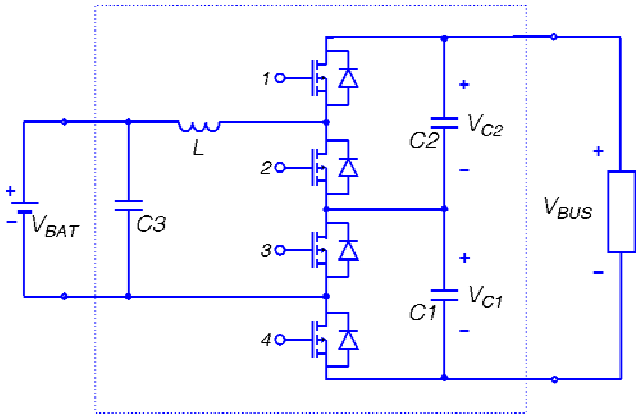


Fig. 1 The Bidirectional TLNPC Converter

Additional benefits of the TLNPC circuit are related to the multilevel structure, allowing to use components with halved voltage rating, and so with a reduced series resistance. It also reduces power losses for doubling the apparent

switching frequency applied to the core with respect to actual switch commutation frequency. The TLNPC converter was originally proposed for inverters [26-27]. Later, several versions have been proposed [28] and have been applied to dc-dc converters after [29-30]. To achieve bidirectionality its synchronous version has to be used, incorporating 4 switches according to the schematic reported in fig. 1.
The section 2 below will analyze circuit operation both in BDR-BOOST mode and BCR-BUCK mode. In section 3 the control issues of the converter are briefly described, while section 4 reports an extensive characterization of the laboratory prototype under several load/battery/bus conditions.

2 Circuit Operation

2.1 Buck Mode

The BCR-BUCK operating mode (energy flows from the right to the left in Fig. 1, from V_{BUS} to V_{BATT}) can be split into 4 phases (the simulated waveform are reported in Fig. 2):
Phase I – S_1, S_3 ON and S_2, S_4 OFF: energy is directly transferred from the capacitor C_2 to the battery and the inductor current linearly increases proportionally to the applied voltage $V_{C2} - V_{BAT}$;
Phase II – S_2, S_3 ON, and S_1, S_4 OFF: the current continues to flow through the inductor towards the battery and decreases proportionally to the applied voltage $-V_{BAT}$;

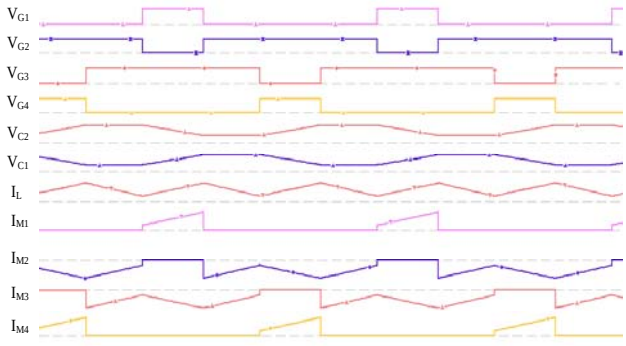


Fig. 2 – BCR-BUCK mode waveforms for the TLNPC converter

Phase III – S_2, S_4 ON and S_1, S_3 OFF: the energy is again directly transferred to the battery but from the capacitor C_1 and the inductor current increases proportionally to $V_{C1} - V_{BAT}$;
Phase IV – S_2, S_3 ON, and S_1, S_4 OFF: similarly to phase II, the inductor current flows towards the battery and decreases proportionally to $-V_{BAT}$.
The voltage reduction achieved for a given duty cycle D is exactly doubled with respect to the classical BUCK converter:

$$V_{OUT\_BUCK} = \frac{D}{2} V_{IN} \qquad (1)$$

where D = effective duty-cycle = $2D_{SW}$ (and D_{SW} is the duty-cycle of the switches S_1 and S_4, that necessarily must be smaller than 0.5).

With simple calculations, the inductor current ripple can be estimated:

$$\Delta I_L = \frac{V_{BUS} \cdot D_{SW} \cdot (1 - 2D_{SW})}{2 \cdot f_{SW} \cdot L} \qquad (2)$$

where f_{SW} is the switching frequency.

2.2 Boost Mode

The BDR-BOOST operating mode (energy flows from the left to the right in Fig. 1) can be similarly split into 4 phases (the simulated waveform are reported in Fig. 3):
Phase I – S_2, S_3 ON and S_1, S_4 OFF: energy from the battery is stored in the inductor while the inductor current linearly increases, proportionally to the applied voltage V_{BAT};
Phase II – S_1, S_3 ON, and S_2, S_4 OFF: energy is transferred from the inductor towards the bus to the capacitor C_2 and the inductor current decreases proportionally to the applied voltage $V_{BAT} - V_{C2}$;
Phase III – S_2, S_3 ON and S_1, S_4 OFF: as in phase 1, the energy is stored in the inductor and the inductor current increases proportionally to V_{BAT};
Phase IV – S_2, S_4 ON, and S_1, S_3 OFF: energy is again transferred from the inductor towards the bus, but now to the capacitor C_2 and the inductor current decreases proportionally to $V_{BAT} - V_{C1}$.

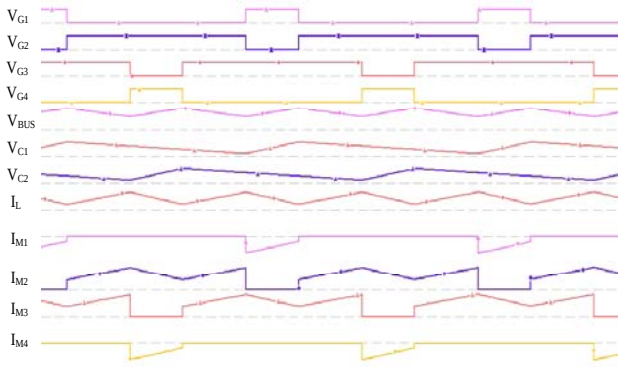


Fig. 3 – BDR-BOOST mode waveforms for the TLNPC converter

The voltage gain is again doubled with respect to the classical BOOST converter:

$$V_{OUT\_BOOST} = \frac{2}{1-D} V_{IN} \qquad (3)$$

where D = effective duty-cycle = $2D_{SW}-1$, (and D_{SW} is the duty-cycle of the switches S_2 and S_3, that must be necessarily greater than 0.5).

Inductor current and capacitor voltage ripples can be calculated with Eq. (4) and (5), helping in the design of the main converter components:

$$\Delta I_L = \frac{V_D \cdot (1-D_{SW}) \cdot (2D_{SW}-1)}{2 \cdot f_{SW} \cdot L} \qquad (4)$$

$$\Delta V_{C_{1,2}} = \frac{V_B}{f_{SW} \cdot C_{1,2} \cdot R_L \cdot (1-D_{SW})} \qquad (5)$$

where R_L is the load.

The voltage ripples on the bus capacitors C_1 and C_2 are in opposition and so the combined bus ripple will have a double frequency and about an halved amplitude (it will be exactly halved in case of D = 0.5 and then D_{SW} = 0.75), allowing to reduce the size of the storage devices (inductor and capacitors).

It is to be mentioned that in a three level converter, both in the Buck and Boost operating mode, the two high–side capacitors C_1 e C_2 will be operated to alternately supply energy to (or draw energy from) the battery. This operation will produce ideal waveforms such as reported in Fig. 2 and 3. However, unavoidable tolerances and variations of control circuits, switches, capacitors, etc…, will result in a small difference in the amount of energy supplied or drawn by each capacitor at each cycle. As a consequence, the bus capacitors will have different voltages, the presented waveform will be asymmetrical and the converters performance will not be optimal. As a consequence, in order to cope with ripple specifications the reactive components (inductor and capacitors) have to be slightly oversized with respect to ideal values.

3 Control Principle

In Fig. 4 and 5, block diagrams of the control schemes for the two operating modes are reported.

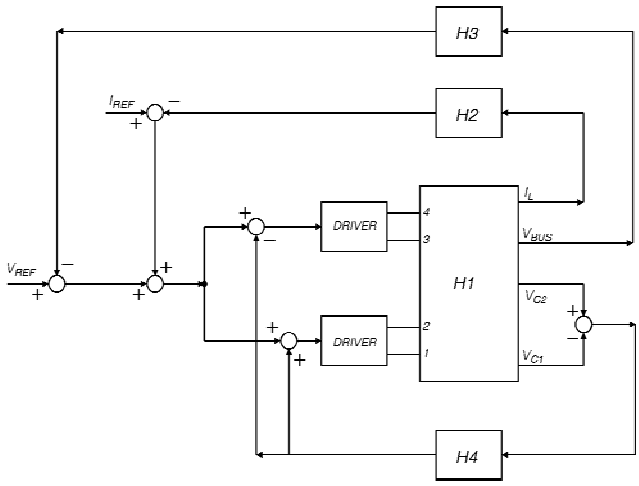


Fig. 4. Converter control scheme during BDR-BOOST operation

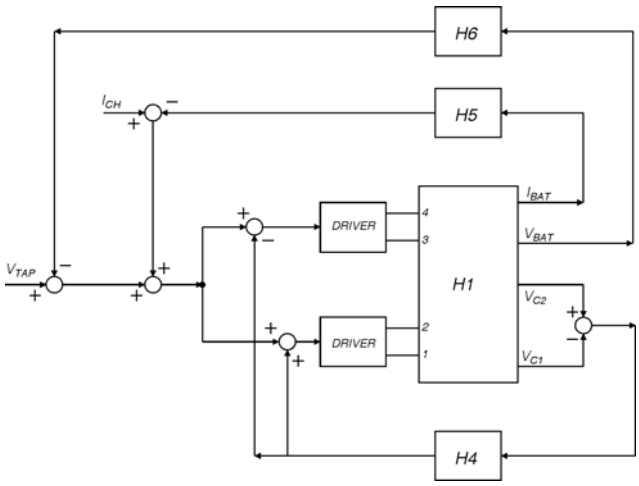


Fig. 5. Converter control scheme during BCR-BUCK operation

The two control schemes include various control loops, but they are based on classical current-control mode, with the inner current loops (H2 and H5 in the figures) that allow to limit the maximum current cycle-by-cycle.

To solve the above mentioned issue of the high – side capacitors unbalancing, a slower control loop (H4) is provided in both the operating mode: the voltages on C_1 and C_2 capacitors are compared and the resulting control signal is adequately merged into the control loop to compensate the duty-cycle of each single switch and to keep the neutral point at the center of the bus lines.

In order to reduce the part count, the controllers for buck and boost mode share many circuits, such as current and voltage sensors, drivers, ramp generators, comparators and the feedback (or feedforward) loop for the high–side capacitors balancing.

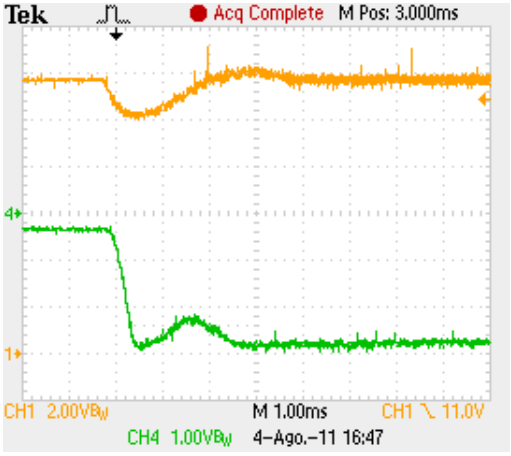
4 Experimental Results

We obtained a laboratory prototype capable to supply a 100 W / 12 V load from a single Li-ion cell discharged down to 3 V, corresponding to approx. 40 A absorption from the battery. When other voltage sources take control of the bus, the same bidirectional converter is able to recharge the same Li-ion cell while modulating the current both on the basis of battery voltage and of the excess power available from bus. The battery charging is initially carried out at constant current and is switched to constant-voltage charging when the battery voltage exceeds a specified threshold (tapering mode).

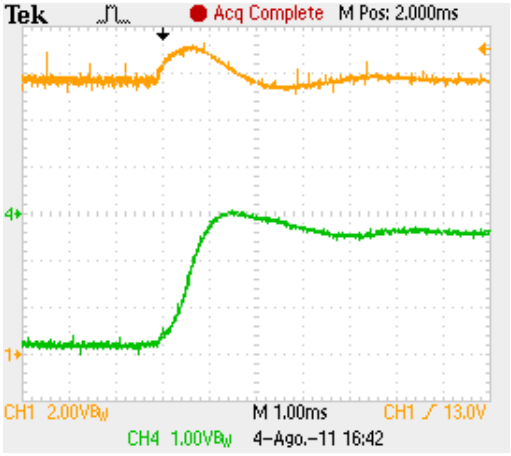
The prototype at first has been tested separately in the BDR-Boost mode and in the BCR-Buck mode. Then transitions between the two operating modes have been tested.

4.1 BDR Mode

In BDR-BOOST mode the converter has to supply the main bus during short eclipse or when an excess of power is required from loads. To test the control stability and effectiveness in boost mode, a step change in the power required from loads has been applied. The experimental results for two typical transients are shown in Fig. 6 and the bus voltage and the battery discharge current are reported. The results show a good stability and a fast response with a total transient duration of about 4 ms with a maximum bus voltage deviation of less than 15% in both directions of the load step change.



(a)



(b)

Fig. 6. Typical transients in BDR-boost mode: bus voltage (yellow curve) in response to a pulsed R_L transition with battery discharge current (green curve, -2A/V) from 4 A to 28 A: (a) Load increase; (b) Load decrease.

The efficiency of the TLNPC converter prototype has been measured in BDR mode as a function of the discharge power at different battery level, from 2.7 V to 4.1 V. In Fig. 7 the overall efficiency is shown. At full power (100W) and at a battery voltage of 4.1 V the

efficiency is about 91%, while it reaches nearly the 94% at half the full power.

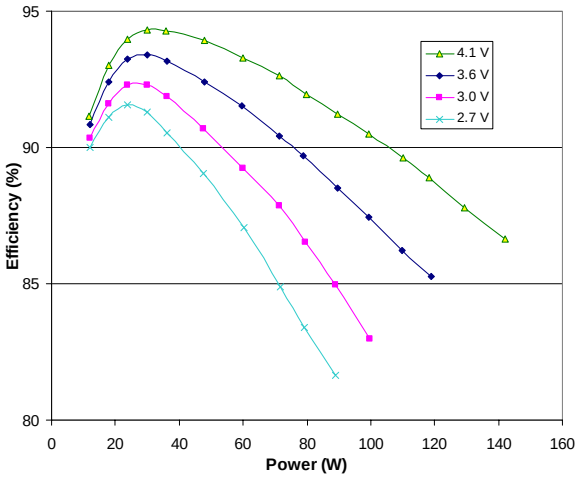
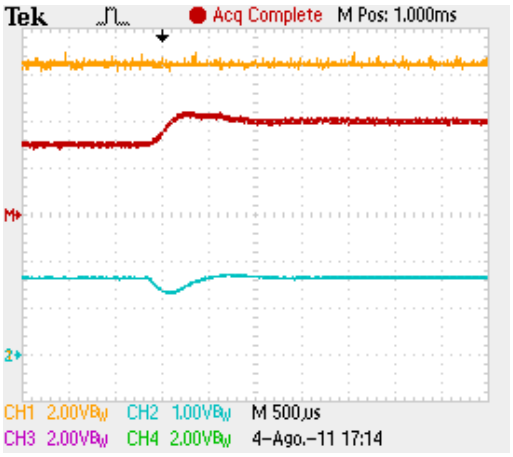


Fig. 7. Overall efficiency Vs. output power in BDR-BOOST mode

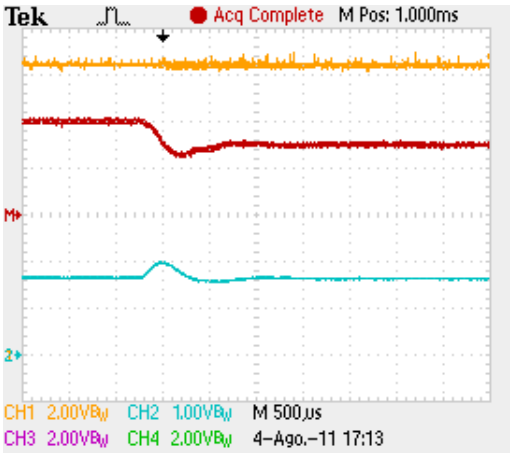
4.2 BCR Mode

The control is capable to manage the charging of a Li-ion battery delivering a constant current, independent from the battery voltage (variations of about 1% has been measured). When the battery voltage exceed a specified voltage threshold (4.1V), the control switches to a constant-voltage charging mode (tapering mode) to complete battery charging.

Just to test the converter control stability, the response to fast variations both on bus voltage and on battery voltage has been tested. In order to stress controller stability, transients amplitude and speed were by far larger than expected on actual operating conditions and were obtained by replacing the battery by an electronic load. As reported in Fig. 8 for battery voltage transients, the prototype shows very good behaviour both in terms of stability and response speed (duration of about 1 ms, with current variations of less than 20% in both directions and no ringing or overshoot).



(a)



(b)

Fig. 8. Transient in BCR-BUCK mode at a charging rate of 8 A: battery charging current (cyan curve, 1A/V) in response to a pulsed transition on the battery voltage (red curve) between 3 V and 4 V: (a) V_{BAT} increase; (b) V_{BAT} decrease

4.3 BCR-BDR mode transitions

Step transitions between the two main operating modes have been forced by applying large variations of the power required from loads. The experimental set-up is shown in the diagram block reported in Fig. 9.

In Fig. 10 a BCR to BDR transition is reported. Before the transition, the bus is controlled and regulated by an Array Power Regulator supplied by a photovoltaic source, and the power available from the solar array is enough to supply the bus and to recharge the battery. Then the TLNPC converter operates in BCR-BUCK mode recharging the battery with a constant current of about 8 A (green curve).

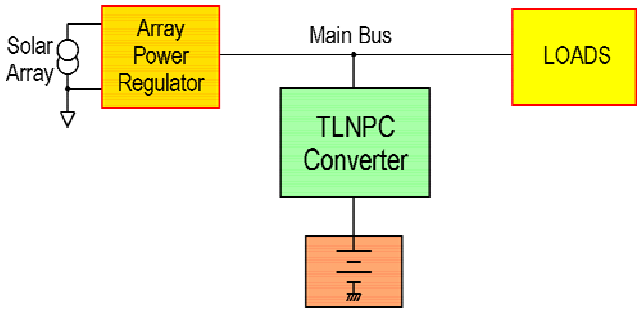


Figure 9. Experimental set-up for the TLNPC operating mode transitions

As the load undergoes a large step increase, the power from solar array is no more sufficient to supply the bus, and the converter changes its operating mode to BDR and regulate the bus with a battery discharge current of 24 A. The main bus has a voltage drop of about 2 V amplitude (less than 20%) for a duration of about 15 ms. The entire duration of the transition is about 30 ms and no instability and ringing has been detected on the main bus.

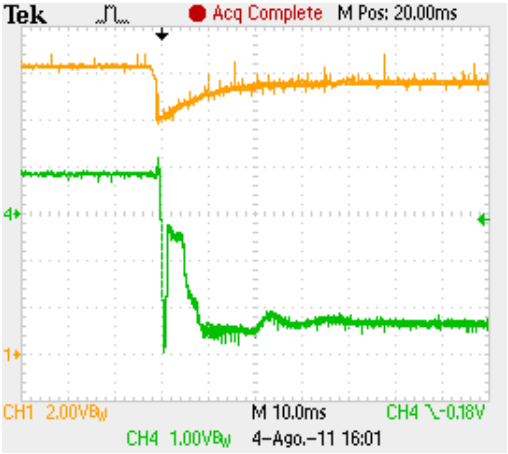


Figure 10. Typical transition from BCR to BDR operating mode, corresponding to a pulsed load transition during battery charge (from 8A in charging mode, positive current, to 24A in discharging mode, negative current).

Conclusions

A Battery Charge Discharge Regulator, capable to interface a single Li-Ion cell with a 12 V bus. This will simplify the design of the energy storage subsystem for small space vehicles, intended to undergo high peaks of load power with comparatively small requirements about stored energy.

References

[1] Venkatesan K., "Current mode controlled bidirectional flyback converter", *Proc. IEEE PESC*, pp. 835-842, June–July 1989.

[2] Ray B., "Bidirectional dc–dc power conversion using quasi-resonant topology", *Proc. IEEE PESC*, pp.617-624, June 1992.

[3] Jain M., Daniele M. and Jain P. K., "A Bidirectional DC-DC Converter Topology for Low Power Application", *IEEE Transactions on Power Electronics*, Vol.15, No.4, pp.595-606 July 2000.

[4] Chung H., Wai-Leung Cheung and Tang K. S., "A ZCS Bidirectional Flyback DC/DC Converter", *IEEE Transactions on Power Electronics*, Vol.19, pp1426-1434, Nov 2004.

[5] Kilders E. S., Ejea J. B., Ferreres A., Maset E., Esteve V., Jordan J., Garrigos A. and Calvente J., "Bidirectional Coupled Inductors Step-up Converter for Battery Discharging-Charging", *Proc. IEEE PESC*, pp. 64-68, 2005.

[6] Yamamoto K., Hiraki E., Tanaka T., Nakaoka M. and Mishima T., "Bidirectional DC-DC converter with full-bridge / push-pull circuit for automobile electric power systems", *Proc. IEEE PESC*, pp. 1-5, Jun. 2006.

[7] Lon-Kou Chang, Ru-Shiuan Yang, "An isolated full-bridge DC/DC converter with bidirectional communication capability," *IEEE Power Electronics and Motion Control Conference*, pp. 320-325, 2008.

[8] Pascual C. and Krein P. T., "Switched capacitor system for automatic series battery equalization", *Proc. IEEE APEC*, 1997, vol. 2, pp. 848–852.

[9] Giess H., "The operation of VRLA lead acid batteries in parallel strings of dissimilar capacity or Can we now sin?", *Proc. IEEE Telecommun. Energy Conf.*, 1999, pp. 18.1.1–18.1.5.

[10] Bouhours G., Klein E., Rebuffel C., Wallstein C., Asplanato R. and Borthomieu Y., "Lithium-Ion Cell Passivation in Flight Conditions", *Proc. 9th European Space Power Conference*, St.Raphael, France, 6-10 June 2011.

[11] Loche D., Barde H., "Lithium-ion battery cell balancing in LEO", *Proc. 6th European Space Power Conference*, Porto, Portugal, 6-10 may 2002, -ESA SP-502-, may 2002, pp.507-512.

[12] Wai R.J., Duan R.Y., "High-Efficiency Bidirectional Converter for Power Sources With Great Voltage Diversity", Transactions on Power Electronics, IEEE Volume: 22, Issue: 5;2007, pp. 1986-1996.

[13] Li H., Liu D., Peng F.Z. and Su G.J., "A Small Signal Analysis of A Dual Half Bridge Isolated ZVS Bi-directional dc-dc converter for Electrical Vehicle Applications", *Proc. PESC '05. IEEE 2005*.
[14] Peng F.Z., Li H., Su G.J. and Lawler J.S., "A New ZVS Bidirectional DC–DC Converter for Fuel Cell and Battery Application", *IEEE Transactions on Power Electronics*, Volume: 19 , Issue: 1
[15] Chu C.L. and Chen Y., "ZVS-ZCS Bidirectional Full-Bridge DC-DC Converter", *International Conference on Power Electronics and Drive Systems*, PEDS 2009.
[16] Caricchi F., Crescimbini F. and Di Napoli A., "20 kW Water-cooled Prototype of a Buck-Boost Bidirectional DC-DC Converter Topology for Electrical Vehicle Motor Drives", *Proc. APEC 1995*, Tenth Annual.
[17] Lei D., Xueping W., Zhen L. and Xiaozhong L., "A New Soft Switching Bidirectional Buck or Boost DC-DC Converter", *International Conference on Electrical Machines and Systems*, 2008. ICEMS 2008.
[18] Waffler S. and Kolar J.W., "A Novel Low-Loss Modulation Strategy for High-Power Bidirectional Buck+Boost Converters", *Transactions on Power Electronics*, IEEE Volume 24, Issue 6.
[19] Sable D.M., Lee F.C. and Cho B.H., "A Zero-Voltage-Switching Bidirectional Battery Charger - Discharger for the NASA EOS Satellite", *Proc. 7th APEC '92*.
[20] Northcott D.R., Filizadeh S. and Chevrefils A.R., "Design of a Bidirectional Buck-Boost DC-DC Converter for a Series Hybrid Electric Vehicle", *Proc.Vehicle Power and Propulsion Conference*, VPPC '09. IEEE 2009.
[21] Tonicello F., "The Control Problem of Maximum Point Power Tracking in Power Systems", *Proc. Seventh European Space Power Conference*, Stresa, Italy, 9-13 May 2005.
[22] Calvente J., Martinez-Salamero L., Garces P., Leyva R. and Capel A., "Dynamic Optimization of Bidirectional Topologies for Battery ChargeDischarge in Satellites", *Proc. 32nd Power Electronics Specialists Conference*, PESC. 2001.
[23] Pavlovsky M., Tsuruta Y. and Kawamura A., "Bi-directional Buck/Boost Dc-Dc Converter with Ultra High Efficiency Based on Improved SAZZ Topology", *Proc. Energy Conversion Congress and Exposition*, ECCE 2009. IEEE, 2009.
[24] Su N., Xu D., Chen M. and Tao J., "Study of Bi-Directional Buck-Boost Converter with Different Control Methods", *Proc. Vehicle Power and Propulsion Conference*, VPPC '08.
[25] Du Y., Zhou X., Bai S., Lukic S. and Huang A., "Review of Non-isolated Bi-directional DC-DC Converters for Plug-in Hybrid Electric Vehicle", *Proc. 25th IEEE Applied Power Electronics Conference*, APEC 2010.
[26] Nabae A., Takahashi I. and Akagi H., "A new neutral-point-clamped PWM inverter", *IEEE Transactions on Industry Applications*, Vol. IA-17, 1981, pp.518-523.
[27] Baker R. H., "Bridge converter circuit", *U.S. Patent* 4 270 163, May 26, 1981.
[28] Floricau D., Gateau G., Leredde A. and Teodorescu R., "The Efficiency of Three-level Active NPC Converter for Different PWM Strategies", *Proc. Compatibility and power electronics 2009*, CPE 2009, Badajoz, Spain, maj 20-22, 2009, pages 331 – 336.
[29] Pinheiro J. R. and Barbi I., "The three-level ZVS PWM converter—A new concept in high-voltage DC-to-DC conversion", *Proc. IEEE IECON*, 1992, pp. 173–178.
[30] Pinheiro J. R. and Barbi I., "Wide load range three-level ZVS-PWM DC-to-DC converter", *Proc. PESC*, IEEE 1993, pp. 171–177.

CONTACT FORCES OPTIMIZATION OF A BIO-INSPIRED CLIMBING ROBOT

Isacco Pretto
CISAS "G. Colombo", University of Padua, Italy

Yasong Li
School of Engineering Science, Simon Fraser University, Canada

Silvio Cocuzza
CISAS "G. Colombo", University of Padua, Italy

Stefano Debei
Department of Mechanical Engineering, University of Padua, Italy

Carlo Menon
School of Engineering Science, Simon Fraser University, Canada

24-28 October 2011, Venezia, Italy

Keywords: climbing robot, minimax, least squares, contact, force optimization

Abstract

Control torques applied at the climbing robot joints can be optimized in order to minimize the required strength of the synthetic dry adhesives that are used to keep the robot attached to an inclined surface. Optimal torques are computed by using MiniMax and Least Squares optimization approaches. This paper focuses on a preliminary investigation intended to identify the convenient control torques that the joint actuators of a Legged Climbing Robot (LCR) should exert, in order to improve its stability and take full advantage of the adhesive contact properties. The results obtained are compared to a natural reference condition in which all the joints are unloaded, by placing the robot in an over-constrained static configuration. An experimental setup is designed in order to test the control concepts and validate the simulation campaign. Results show that the use of optimal torques could be highly desirable as the maximum pulling force exerted at the contact tips is drastically reduced, enhancing the stability margin of the climbing robot configuration.

1 Introduction

The utilization of legged robots for space exploration represents a promising area of research, and several design architectures have been proposed by taking into account those available in nature [1], according to a biologically-inspired approach. Recent works have proven the feasibility of developing robots capable of climbing vertical walls by taking advantage of synthetic dry adhesive pads [2, 3, 4], and suggested for use on orbital space applications and space exploration [1].

The utilization of micro-machined dry adhesive materials as substrate of a performing contact surface is being extensively studied [5, 6] and this paper presents a work developed in the framework of the design of a climbing robot prototype relying on dry adhesive pads for space application purposes [8, 9].

In the context of joint torque control for this class of climbing robots [7], the concept introduced here relies in the analogy existing between grasping forces in manipulation robotics [10] and climbing forces in climbing robots, while introducing MiniMax and Least Squares optimization criteria [11].

The mechanical model of the climbing robot consists of a main body or payload to which several legs are mounted in order to form a structure with grasping capacities. Consider an inertial reference frame solid to the climbing surface $\vec{\Sigma}_c$ and express all the quantities in this frame. The external forces applied to the robot are its local-gravity weight $\boldsymbol{F}_b = m\,\boldsymbol{g}_{\otimes}$ and the resultant contact forces and torques at each contact pad $\boldsymbol{F}_{ci}$ and $\boldsymbol{T}_{ci}$, that from a closer insight at each contact pad, are the resultant action of a distributed surface stress field $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{r})$ applied at the points of contact $\boldsymbol{r}$ between the adhesion pads and the climbing surface. In this work we make the common assumption that the resultant torques at the contact tips are negligible in determining the contact stresses, while a complete contact wrench model can be easily extended.

The robot is conceived in order to take advantage of dry adhesive attachment pads, that in the developed prototype consist of Poly-

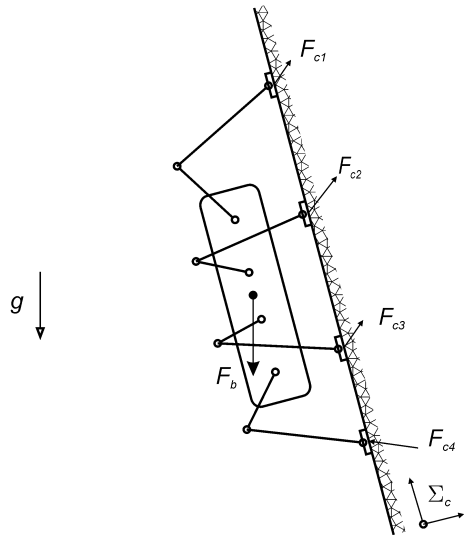


Figure 1: Climbing robot weight and contact forces.

dimethylsiloxane (PDMS) micro-machined layers. In order to take full advantage of the adhesion properties of the robot, and be able to perform a stable climbing of steep, vertical, and negative gadient surfaces, the control problem have been conceived by decomposition in two independent parts: 1) a model of adhesion that relates the resultant actions on the pads to the local surface stresses, that, together with a criticality measure of the local stress, give a prediction of the detachment conditions of the pads, and allow to formulate a measure of distance from the overall robot detachment conditions, to be minimized through active joint control; 2) a joint torque active control logic, that together with a structural model of the robot relating the joint torques to the surface contact forces, allows to take the best advantage of the system capacity in terms of grasping and adhesion. This paper concentrate on the control part, while for details on adhesion model principles see for example [11].

2 Experimental setup

The problem of controlling the surface contact forces is presented and resolved according to optimization criteria for an experimental proto-

type, that have been developed for the purpose of testing the proposed concepts [9]. A simulation counterpart that reproduces the geometrical and inertial properties of the experimental prototype have been developed in *Matlab*^{TM}, and validated in the *MSC Adams*^{TM} environment. The experimental setup, illustrated in Fig. 2, is composed of a planar robot in a fixed configuration and a set of load cells that retrieve the contact reaction forces normal to the climbing surface. The robot joints are powered by

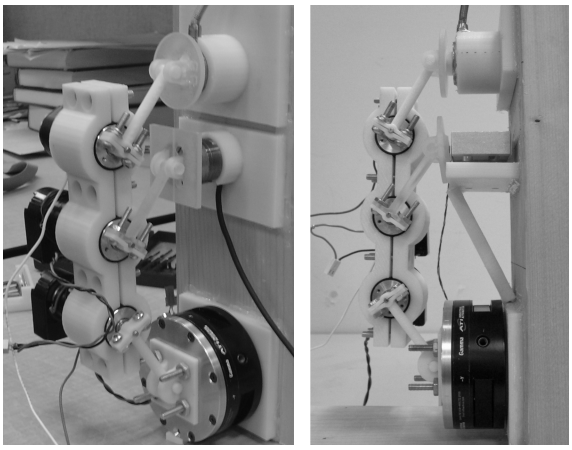


Figure 2: Experimental setup of the planar 3 DOF prototype, with three joint motors and three force sensors at the contact hinges. Side and lateral views.

3 servomotors that provide the desired torque. Joints and links are aligned in a vertical axis and numbered from top to bottom, such that joint 1 and link 1 are at the top of the robot, joint 2 and link 2 are at the center, and joint 3 and link 3 are at the bottom of the robot. The analyzed configuration simulates a vertical climb. The objective of the optimization is to define a set of desired control torques that keep the robot adhesion state far from the detachment conditions. The robot is fixed in an over-constrained configuration, such that its balance is achieved also with zero torque at the joints; in this way it is possible to compare the results of an active joint control with the *natural* unloaded joint conditions.

The geometrical and inertial properties of the robot that have been implemented also in the simulation models, and are reported in Tab. 1. The instrumentation part is composed of 3 load cells arranged along a vertical line, which are capable of retrieving the reaction forces in the direction normal to the climbing surface [9].

Property	Body	Link 1	Link 2	Link 3
ℓ_i [mm]	100	50	50	50
θ_i [deg]	90	50	50	−50
m_i [kg]	0.56	0	0	0

Table 1: Experimental climbing robot geometrical and inertial properties.

3 Statics model

In this section it is shown how it is possible to optimize the surface contact forces $\boldsymbol{F}_{c\,i}$ by means of applying a set of optimal control torques $\boldsymbol{\tau}$ at the robot joints. Consider the test robot with 3 powered 1-DOF joints and 3 legs and contact tips fixed to the climbing surface, and an external gravity acceleration of $\boldsymbol{g}$, generating a weight force $m\,\boldsymbol{g}$ about the robot CM.

In order to maintain the static balance of the robot, the resultant external forces and torques applied to the robot need to compensate. The static balance equation of the whole system, represents the statics *equality constraint* of the optimization problem. In this framework, be defined a pole for evaluation of the moments as the origin of the reference frame solid with the climbing surface $\vec{\Sigma}_c$.

Simplified adhesion model In order to simplify the problem and given the experimental setup, consider the problem of controlling the normal components of the contact forces instead of a complete function of contact forces and torques as proposed in the previous section. In this condition, the equivalent measure of surface stress to be controlled be:

$$\sigma_e = \bar{\sigma}_n = \frac{F_{cn}}{A_c} = \frac{\boldsymbol{F}_c^{\mathrm{T}}\boldsymbol{n}}{A_c} \qquad (1)$$

Where $\boldsymbol{n}$ is the unit vector normal to the flat climbing surface, oriented towards the robot

such that it corresponds to a detaching pulling force, and A_c is the adhesion pad contact area. For this experimental setup, the x-axis is horizontal, oriented towards the vertical surface in the direction opposite to the normal vector, and the y-axis, is in vertical direction, opposite to the gravity vector.

Observe that the detachment of the robot from the wall would eventually initiate when the equivalent stress in a contact point overcomes the threshold value. In general, the robot static state can be said to be as far as possible from the undesired detaching conditions when we are able to minimize the maximum value of the stress measure $\sigma_e(\boldsymbol{c})$. This problem, for the simplified model considered here, is equivalent to the minimization of the maximum value of the *pulling* contact force normal to the climbing surface:

$$\text{minimize} \max_i \{F_{cn\,i}\} \tag{2}$$

If we are able, by means of a set of joint torques, to lead the maximum contact force to a lower level, it can be said that the robot joints are working for stability. The optimization problem in this form, in the contact forces variables, becomes a *minimax* problem, as explained in the following paragraphs.

Static balance The overall balance of the robot in static configuration is described by the Newton-Euler equations for the static condition, and the resultant forces and torques applied to the robot do balance. The robot balance condition represents the equality constraint of the problem. In order to express this constraint, define the contact forces vector variable as the collection of all the contact forces applied at each tip to the climbing surface, equal to the opposite of the reaction forces that the surface applies to the robot:

$$\boldsymbol{c} \triangleq \begin{bmatrix} \boldsymbol{F}_1 \\ \boldsymbol{F}_2 \\ \boldsymbol{F}_3 \end{bmatrix} = - \begin{bmatrix} \boldsymbol{F}_{c1} \\ \boldsymbol{F}_{c2} \\ \boldsymbol{F}_{c3} \end{bmatrix} \tag{3}$$

Be G the robot center of mass of position $\boldsymbol{r}_G = \boldsymbol{r}_{OG}$. The external wrench due to the weight $\boldsymbol{F}_b$, composed of the resultant force and resultant torque about the system origin due to the weight is:

$$\boldsymbol{b} \triangleq \begin{bmatrix} \mathbf{1} \\ \boldsymbol{r}_G^{\times} \end{bmatrix} \boldsymbol{F}_b = \begin{bmatrix} \boldsymbol{F}_b \\ \boldsymbol{r}_G^{\times} \boldsymbol{F}_b \end{bmatrix} \tag{4}$$

By means of defining a *grasp matrix* $\boldsymbol{G}$, suggested from the analogies with problems of manipulation grasping, of the form:

$$\boldsymbol{G} \triangleq \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \boldsymbol{r}_{c1}^{\times} & \boldsymbol{r}_{c2}^{\times} & \boldsymbol{r}_{c3}^{\times} \end{bmatrix} \tag{5}$$

The system balance constraint can be expressed as:

$$\boldsymbol{G}\,\boldsymbol{c} = \boldsymbol{b} \tag{6}$$

In former Eq. (5), the positions vectors $\boldsymbol{r}_{ci}$ are the positions of the centers of the contact tips, while for a non-fixed robot configuration the grasp matrix depends on the configuration itself $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{\theta})$. Since the balance equation represents an under-constrained system in the contact forces variables $\boldsymbol{c}$, an optimization process can be carried out. The general solution of the contact forces can be expressed by means of an arbitrary vector $\boldsymbol{\xi}$ as:

$$\boldsymbol{c}_\xi = \boldsymbol{G}^{\dagger}\boldsymbol{b} + \boldsymbol{Z}_G\,\boldsymbol{\xi} \tag{7}$$

Where $\boldsymbol{Z}_G$ is the projector on the null space of $\boldsymbol{G}$, see [11]. The optimization of the contact forces $\boldsymbol{c}$ is the subject of the following paragraph, and can be carried out by observing that different configurations of contact forces can be achieved by means of different internal joint torques $\boldsymbol{\tau}$.

Once a desired set of contact forces $\boldsymbol{c}$ have been evaluated, the necessary joint torques that are able to produce those interaction forces are given by means of statics relations of the structure.

The forces Jacobian matrix is here denoted by $\boldsymbol{J}^F$, and the statics relation for the robot can be expressed in the form:

$$\boldsymbol{\tau} = \boldsymbol{J}^F\,\boldsymbol{c} \tag{8}$$

With $\boldsymbol{J}^F$ of the experimental robot as in Eq. (9).

For the robot structure of our tests, be α the inclination of the climbing surface, ℓ the length

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ -d/2 & 0 & d/2 & 0 & 0 & 0 \\ \ell \sin\theta_1 & 0 & 0 & -\ell\cos\theta_1 & 0 & 0 \\ 0 & \ell\sin\theta_2 & 0 & 0 & -\ell\cos\theta_2 & 0 \\ 0 & 0 & \ell\sin\theta_3 & 0 & 0 & -\ell\cos\theta_3 \end{bmatrix} \begin{bmatrix} c_{1x} \\ c_{2x} \\ c_{3x} \\ c_{1y} \\ c_{2y} \\ c_{3y} \end{bmatrix} = \begin{bmatrix} mg\cos\alpha \\ mg\sin\alpha \\ \tau_\alpha \\ -\tau_1 \\ -\tau_2 \\ -\tau_3 \end{bmatrix} \tag{9}$$

$$\text{with} \quad \tau_\alpha = -\tau_1 - \tau_2 - \tau_3 - mg \cdot (e_y \cos\alpha + e_x \sin\alpha)$$

of the links, and $\boldsymbol{e}$ the eccentricity of the center of mass G with respect to the geometrical center of the main body, the statics equations of the robot can be organized as in Eq. (9).

4 Contact forces optimization

In this section two optimization solutions will be considered to the contact forces distribution. Their performances have been tested on a Matlab simulator, and validated by means of the experimental robot prototype previously described. Furthermore, given the over-constrained static configuration of the robot, it is possible to compare the results with the natural solution without any torque load at the robot joints. The two optimization procedures examined are 1) the least squares (LS) optimization, and 2) the minimax (MMX) optimization, according to common numerical routines.

Minimax optimization The normal contact forces of all the robot tips can be extracted by means of a selection matrix $\boldsymbol{c}_n = -\boldsymbol{c}_x = \boldsymbol{S}_x\,\boldsymbol{c}$, and the MMX optimization problem of Eq. 2 reformulated by taking into account the balance constraint as:

$$\begin{array}{ll} \text{minimize} & \max_i (\boldsymbol{S}_x\,\boldsymbol{c})_i \\ \text{subject to} & \boldsymbol{G}\boldsymbol{c} = \boldsymbol{b} \end{array} \tag{10}$$

The solution is provided by means of common numerical routines, on which no further details are of interest for the purposes of these tests.

Least squares optimization On the other side, the least squares optimization of the pulling forces, can be expressed as:

$$\begin{array}{ll} \text{minimize} & \|\boldsymbol{S}_x\,\boldsymbol{c}\|^2 \\ \text{subject to} & \boldsymbol{G}\boldsymbol{c} = \boldsymbol{b} \end{array} \tag{11}$$

Note that for a general expression of the matrix $\boldsymbol{S}_x$ that could characterize a different cost function, the problem of Eq. (11), results to be a Least Squares Problem with Equality constraints problem, LSE. Anyhow, given the special formulation of the cost function given in this test-case, with the cost-function matrix composed only of elements in $\mathbb{B} = \{0, 1\}$, the problem, and its solution, have been found to be equivalent to a simple under-constrained least squares problem:

$$\begin{cases} \boldsymbol{G}\,\boldsymbol{c} = \boldsymbol{b} \\ \boldsymbol{c} \cong \boldsymbol{0} \end{cases} \tag{12}$$

and its solution is:

$$\boldsymbol{c} = \boldsymbol{G}^\dagger\,\boldsymbol{b} \tag{13}$$

For this reason in this context we use the notation LS to identify the solution to the problem in Eq. (11). The optimal contact forces obtained with these procedures, are then passed to the statics model of the robot of Eq. (8), that provides the corresponding joint control torques $\boldsymbol{\tau}$.

Simulations and experimental tests The force control concepts proposed in this paper, have been tested on the experimental robot prototype of Fig. 2, and simulations have been carried out by means of a validated simulator, in which the robot statics and the control paradigms have been implemented. The results of the optimization have been compared to the *natural* unloaded-joints configuration with $\boldsymbol{\tau} = \boldsymbol{0}$. In Fig. 3 are displayed the values of the normal contact forces obtained by means of the natural solution, and the minimax solution, displaying a good correspondence between the simulation results and the experimental data. For what concerns the maximum value of the pulling contact forces, from the analysis of Fig.s

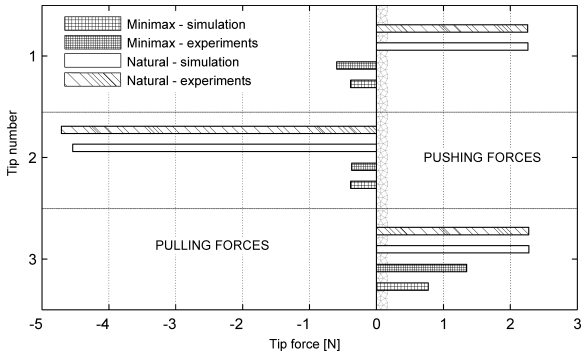


Figure 3: Experimental and simulation values of the normal contact forces at the tips, for the unloaded joints and the Minimax optimization.

4 and 3 it can be noticed that: 1) the MMX solution provides slightly better performances than the LS solution, resulting in a reduction of the maximum force of 31.0% for the simulation results, and showing also a slightly reduced control torque, whose maximum values is reduced of 19.4% with respect to the LS solution; 2) the LS and the MMX solutions highly improve the performances with respect to the unloaded natural configuration.

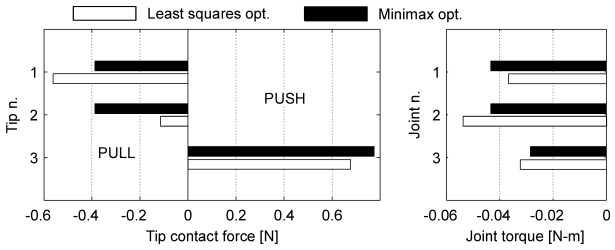


Figure 4: Normal contact forces and required joint torques comparison between Least Squares and Minimax optimizations.

The simulation comparisons result in a peak force reduction of 88.1% for the LS solution, and of 91.4% for the MMX solution, demonstrating the significant improvements on the stability of the climbing robot that can be attained by means of an active joint torque control. For what concerns the experimental data, the MMX solution still provides a reduction of 87.4% of the peak pulling force, confirming very good performances also in experimental conditions. The difference in terms of pulling forces results between the LS and MMX solutions are not significant for what concerns the experimental results and are below the accuracy of the experimental setup. In order to observe the performances of the optimization in different climbing conditions, a simulation campaign have been carried out by means of changing the angle α of the climbing surface (and its normal vector) with respect to the gravity vector. The results are reported in Fig. 5 for the unloaded, LS, and MMX solutions, and the maximum value of the pulling forces together with the maximum value of the required joint torques is provided. In this plot, positive values are aligned with the surface-normal vector, such that they identify the pulling forces; negative values are displayed if none of the contact forces is pulling, and in that condition the minimum (closer to zero) pushing force is displayed. The simulation tests confirm very good performances of the optimization laws for what concerns the reduction in pulling forces of LS and MMX solutions with respect to the natural solution, with a peak in the optimization performances for the vertical climbing conditions discussed in the experimental tests. When the climbing surface is about horizontal, the LS and MMX solutions give the same results, while the solution begins to differ when some pulling detaching components arise. In this case the MMX solution provides always slightly better performances than the LS solution. The experimental condition achieved for an angle of 90 deg, results favorable in terms of joint torques to the MMX solution, although for the inverted configuration achieved for an angle of 270 deg, the convenience is inverted as well and the LS solution provides lower torques. The natural configuration results convenient only for an angle of about 160 deg, where all the three solution provide very similar results thanks to a favorable distribution of the forces. It results also interesting to observe that both the optimal solutions always exhibit a pushing force at the bottom tip, suggesting a similar behaviour to what happens in nature for some reptiles, whose climbing apparatus relies on similar ad-

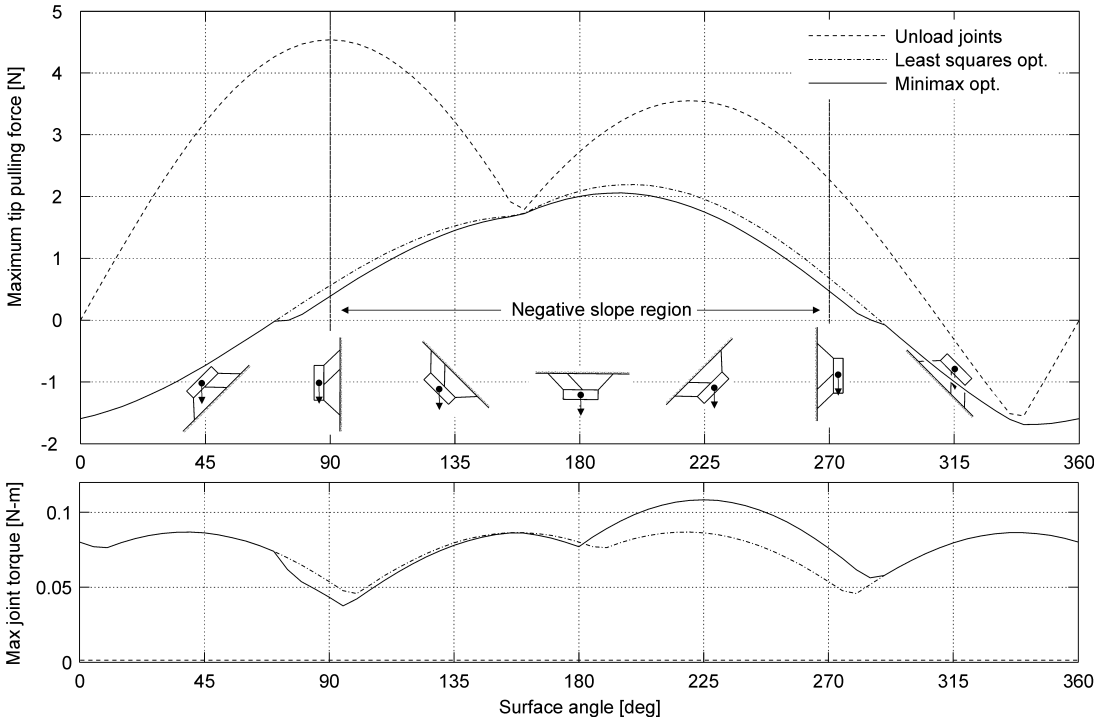


Figure 5: Contact forces and required joint torques for the natural, least squares, and minimax joint loading conditions, at different surface gradients.

hesive principles, and for which the bottom tip can figure as a tail, pushing against the climbing surface in order to preserve the static balance.

5 Conclusions

The problem of controlling the adhesion stability of a climbing robot standing in a static configuration along a generic gradient surface have been addressed and two optimization criteria have been proposed according to minimax and least squares methods. From the test campaign it can be asserted that 1) contact pulling forces can be greatly reduced by means of an active joint torque control, at the expense of higher values for the required joint torques; 2) the LS and MMX solutions do not significantly differ in performances with respect to the natural condition for the test configurations, suggesting further investigation in order to assess if a simple LS control, suitable for real-time implementation with a lower computational effort, can be generally effective for different operative conditions. The study of an improved inequality constrained problem that considers joint-torque limits to the solution may be interesting, in order to reduce the required joint torques, and especially if the robot presents a non null number of joint redundancies such that more freedom would be available in the solution domain. The good results of the simulation and test campaign motivate for the development of a climbing control system based on the proposed concepts, in the perspective of testing the robot behaviour in a 3D robot full-scale prototype.

References

[1] Menon, C., Murphy, M., Sitti, M., Lan, N., "Space exploration - towards new advanced bioinspired climbing robots", Bioinspiration and Robotics - Walking and Climbing Robots, M.K. Habib Ed., *Advanced Robotic Systems International and I-Tech*, pp. 544, 2007.

[2] Daltorio, K. A., Gorb, S., Peressadko, A., Horchler, A. D., Ritzmann, R. E., and Quinn, R. D., "A robot that climbs walls using microstruc-

tured polymer feet", *International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*, pp.131-138, 2005.

[3] Murphy, M., Sitti, M., "Waalbot: an agile small-scale wall climbing robot utilizing dry elastomer adhesives", *IEEE/ASME transactions on Mechatronics*, Vol. 12 , No. 3, pp. 330-338, 2007.

[4] Unver, O., Uneri, A., Aydemir, A., and Sitti, M., "Geckobot: a gecko inspired climbing robot using elastomer adhesives," *International Conference on Robotics and Automation*, pp. 2329-2335, 2006.

[5] Sameoto, D., and Menon, C. "Recent advances in the fabrication and adhesion performance of biomimetic dry adhesives", *Smart materials and structures*, Vol. 19, No. 10, 2010.

[6] Autumn, K., Dittmore, A., Santos, D., Spenko, M., and Cutkosky, M., "Frictional adhesion: a new angle on gecko attachment," *Journal of Experimental Biology*, Vol. 209, No. 18, p. 3569, 2006.

[7] Kim, S., Spenko, M., Trujillo, S., Heyneman, B., and Cutkosky, M., "Whole body adhesion: hierarchical, directional and distributed control of adhesion forces for a climbing robot", *IEEE Conference on Robotics and Automation*, pp. 1268-1273, 2007.

[8] Menon, C., Li, Y., Sameoto, D., and Martens, C., "Abigaille-I: towards the development of a spider-inspired climbing robot for space use", *IEEE RAS / EMBS International Conference on Biomedical Robotics and Biomechatronics*, Scottsdale, Arizona, USA, 2008.

[9] Li, Y., "Preliminary development of a bio-inspired hexapod climbing robot relying on dry adhesives", Masters thesis, *Simon Fraser University*, Burnaby, BC, Canada, 2008.

[10] Li, J., and Kao, I., "Grasp Stiffness Matrix - Fundamental Properties in Analysis of Grasping and Manipulation", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Pittsburgh, PA, USA, 1995.

[11] Pretto, I., "Base Reaction Control of Space Manipulators", Ph.D. Thesis, *University of Padua*, Padova, Italy, 2010.

CEAS 2011

Materials 2

Design and optimization of Radar Absorbing Structure using carbon nanostructured composite materials

D.Micheli, C.Apollo, R.Pastore, S. Laurenzi, R. Bueno Morles , D. Barbera , M.Marchetti.
"Sapienza" University of Rome- Astronautic, Electrical and Energetic Engineering Department (DIAEE),

Via Salaria 851, Roma, (phone: +39/0649919758; fax: +39/0649919757;

e-mail: carmelo.apollo@uniroma1.it

web site: www.saslab.eu

Keywords: *Modelling, Radar Absorbing Material, Winning Particle Optimization, Carbon*

1 Abstract

This paper presents the design of a radar absorbing structures that are based on a multilayer systems where each layer need to be optimized in terms of thickness and type of material. Available materials are accessible in a data base and are carbon-based nanostructured materials that have been characterized in terms of electric permittivity and conductivity in a previous work and here are assumed to be ready for optimization algorithms. The main goal is a multiple task as it tries to reduce the structure thickness while contemporary minimizing the electromagnetic reflected field. With respect to the most part of the current literature, the optimization takes place for a finite set of incidence angle of the electromagnetic field in a wide range of frequency. The algorithm employed is winning particle optimization. In order to validate the design procedure a simple multilayer structure has been built and tested using NRL arch method.

2 Introduction

A branch of scientific research about composite materials is focused on electromagnetic characterization and subsequent application of electric conductive polymers. The use of such structures is relevant in aerospace/aeronautics [1], for electromagnetic (EM) protection from natural phenomena (lightning)[2], and intentional interference with radar absorbing materials (RAM) [3], in electromagnetic compatibility (EMC)[4], for equipment-level shielding, high-intensity radiated fields (HIRF) protection, anechoic chambers (for the realizations of wedges and pyramidal arrays), and human exposure mitigation. In this work, composite reinforced by carbon nanostructured materials are considered, mainly because of their interesting electromagnetic characteristics, such as high electrical conductivity and excellent microwave absorption and shielding behavior[5]-[6]-[7]. Composite materials as well their absorption capability are analyzed and numerical design of wide frequency band microwave absorbing structures is presented and discussed in details. It is crucial to highlight the need of interdisciplinary research fields to go

through nanomaterials: besides nanotechnology, also electromagnetic wave propagation theory, composite materials manufacturing techniques, evolutionary computation algorithms, and use those to design the "quasi perfect absorber" are strongly required.

Composite materials considered in this work for microwave absorbing structure are based on epoxy matrix reinforced with several species of carbon nanomaterials. These latter have been chosen taking into account the lowest market prices: the economic aspects, normally neglected in small laboratory applications, are on the contrary important in real applications where the amount of carbon nanopowders fillers could be relatively high. In such scenario a good compromise in terms of cost/performances has been obtained using industrial grade multiwall carbon nanotubes (MWCNTs, about 300 $/kg). Materials used are MWCNTs (Multi-Wall Carbon Nanotubes type NANOCYLTM NC7000), bought at NANOCYL (diameter around 9.5 nm, length 1.5 μm, purity 90%); Epoxy-resin is PrimeTM 20LV(density 1.123 g/cm^3), Hardner (density 0.936 g/cm^3). CNFs (Carbon NanoFibers), bought at SigmaAldrich (diameter around 75 nm, length 50-100 μm).

As far as optimization algorithms, winning particle optimization (WPO) has been adopted [8][9].

WPO has been applied in order to design and optimize multilayer materials able to effective absorbing/shielding microwaves. Microwave absorption concepts are quite often exploited in RAM [10]-[11]-[12] while microwave shielding concepts are more often adopted in electromagnetic Interference (EMI) applications [13]-[14]. In the first part of the paper the electromagnetic model of multilayer absorber is in detail described. In the second part, the experimental validation of mathematical model of absorber is discussed, showing comparisons between simulations of microwave absorbing structures and measurements of some manufactured composite materials tiles.

3 WPO Algorithm

WPO algorithm is applied to solve a quite complex problem consisting in the design and optimization of microwave absorbing multilayer structure. Two kinds of structures are taken into account the first is RAM, the second is developed for shielding applications where requirements are simultaneously both: electromagnetic interference (EMI) shielding and absorbing capability of materials, i.e., interference suppression by usage of absorbing materials. This last has been called microwave-shielding structure (MSS).

Optimized multilayer structure is based on nanomaterials in particular carbon nanostructure-based composites materials. Conductive fillers have been uniformly dispersed in an epoxy resin at different weight percentages (0.5, 1, 2, 2.5, 3, %wt), and the resulting composites have been dielectrically characterized from electromagnetic point of view [15]. Waveguide measurements have been used to recover the dielectric properties of MWCNTs and carbon nanofibers CNFs, based composite materials. After that, numerical simulations of multilayer absorbing structures have been carried out using data base of dielectric properties of materials. In particular, electromagnetic analysis has been performed integrating the forward/backward propagation formalism to the in-house built WPO, thus able to carry out optimization upon oblique incidence over a finite angular range 0 to 80°. The design/optimization is basically a minimization procedure which seeks the best trade-off between structure thickness (to be minimized) and absorbed EM power (to be maximized). The absorbing power has been quantified through the so-called loss factor (LF) adopted to build the objective function [16]-[17].

$$LF = 1 - |RC|^2 - |TC|^2 \qquad (1)$$

In RAM systems the transmission coefficient TC=0 (because has been assumed perfect electric conductor (PEC) layer at the and of the multilayer), and the reflection coefficient (RC)

is the WPO-optimized reflection coefficient. Both are expressed in the linear form. LF physical meaning is related to the fraction of the incident power vanishing inside the materials because of localization and dissipation phenomena. Both normal and oblique incidences have been evaluated during the automatic optimization [18]. In Fig.1, a general scheme of the electromagnetic absorbing multilayer structure is presented, each layer is generally denoted by the index x.

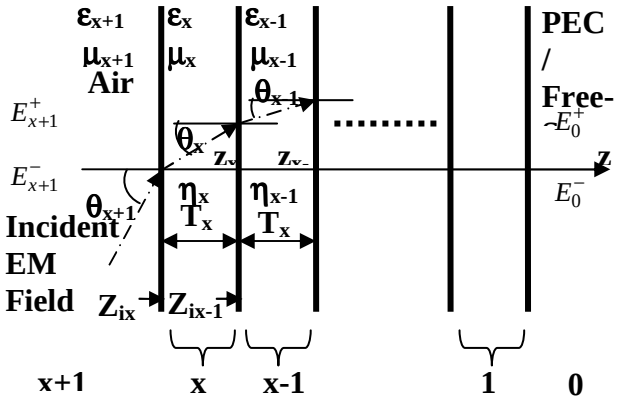


Fig. 1. General Multilayer scheme of electromagnetic absorbing structure; in RAM systems the back layer is a perfect electric conductor (PEC), whereas in MSS systems the structure is backed with free space

Each layer is made of one composite materials and TE / TM characteristic wave impedances η is computed knowing dielectric properties of composite materials and incidence angles of electromagnetic waves acting on each single layer:

$$\eta_x^{TM} = \sqrt{\frac{\mu_0}{\varepsilon_0}}\sqrt{\frac{\mu_{rx}}{\varepsilon_{rx}}}\cos\vartheta_x \qquad \eta_x^{TE} = \sqrt{\frac{\mu_0}{\varepsilon_0}}\sqrt{\frac{\mu_{rx}}{\varepsilon_{rx}}}\frac{1}{\cos\vartheta_x} \qquad (2)$$

where for Snell's law

$$\vartheta_x = \arcsin\left(\frac{\sqrt{\varepsilon_{r\,x+1}}}{\sqrt{\varepsilon_{r\,x}}}\sin\vartheta_{x+1}\right).$$

As far as geometric properties of multilayer structures are concerned, thickness of each layer can be defined by WPO in the range 0 to 6 mm, while the number of layers is upper-bounded to *maximum* 10 layers. The effective number of layers is established by the WPO optimization procedure which is able to remove one or more useless layers from the multilayer structure. The order of layer materials is not *a priori* fixed; instead WPO algorithm access the materials coded in the DB reported in Table I and decides the most appropriate material for each layer.

Layers Material	material code
Epoxy- Resin	1
MWCNT, 0.5 wt%	2
MWCNT, 1.0 wt%	3
MWCNT, 2.0 wt%	4
MWCNT, 2.5 wt%	5
MWCNT, 3.0 wt%	6
CNF, 1.0 wt%	7
CNF, 2.0 wt%	8
CNF, 3.0 wt%	9

TABLE I: Coding of Materials in the Data Base

In RAM applications, RC at the air-absorber interface can be evaluated by the following equation relating the free space impedance to the input impedance seen at the air-multilayer structure interface

$$RC^{TM/TE} = \frac{Z_i^{TM/TE} - Z_0^{TM/TE}}{Z_i^{TM/TE} + Z_0^{TM/TE}}$$

$$RC_{dB}^{TM/TE} = 20Log_{10}\left|\frac{Z_i^{TM/TE} - Z_0^{TM/TE}}{Z_i^{TM/TE} + Z_0^{TM/TE}}\right|, \qquad (3)$$

where Z0 ≅ 377 Ω is the free space impedance and Zi is the input impedance at the first air-absorber interface. The input wave impedance of the multilayer, backed by PEC, can be expressed iterating the following equation for x-th-layer [15]-[16]-[18],

$$Z_{ix}^{TM/TE} = \eta_x^{TM/TE} \frac{Z_{ix-1}^{TM/TE}\cos(k_x t_x) + j\eta_x^{TM/TE}\sin(k_x t_x)}{\eta_x^{TM/TE}\cos(k_x t_x) + jZ_{ix-1}^{TM/TE}\sin(k_x t_x)} \quad (4)$$

In (4), tx=zx -1 -zx is the x-th layer thickness in m, whereas the wave number kx is

$$k_x^2 = (2\pi f)^2 \mu_0 \varepsilon_0 \mu_{rx} \varepsilon_{rx} \cos^2 \vartheta_x = \\ = (2\pi f)^2 \mu_0 \varepsilon_0 \mu_{rx} (\varepsilon_{rx}^{'} - j\varepsilon_{rx}^{''}) \cos^2 \vartheta_x \quad (5)$$

where f is frequency of the incident electromagnetic wave in Hz, ε' is the real part of x-th-layer permittivity and ε" is the imaginary part of layer permittivity.

Complex waves appear in oblique incidence and lossy dielectrics problems [19]-[20]. Because the wave number become complex-valued, e.g., $\vec{k} = \vec{\beta} - j\vec{\alpha}$, the angle of refraction and possibly the angle of incidence may become complex-valued too. In calculating kx by taking square root of (5), is required, in complex-wave problems, to get the correct signs of their imaginary parts, such that evanescent waves are described correctly [21]. This leads to define an "evanescent" square root as follows. Let $\varepsilon_{rx} = (\varepsilon_{rx}^{'} - j\varepsilon_{rx}^{''})$ with $\varepsilon_{rx}^{''} > 0$ for an absorbing medium, then

$$k_x^2 = \begin{cases} \sqrt{(2\pi f)^2 \mu_0 \varepsilon_0 \mu_{rx} (\varepsilon_{rx}^{'} - j\varepsilon_{rx}^{''}) \cos^2 \vartheta_x}, \text{if } \varepsilon_{rx}^{''} > 0 \\ -j\sqrt{-(2\pi f)^2 \mu_0 \varepsilon_0 \mu_{rx} (\varepsilon_{rx}^{'}) \cos^2 \vartheta_x}, \text{if } \varepsilon_{rx}^{''} = 0 \end{cases} \quad (6)$$

Since all nanostrucutred composite materials available in the DB are with $\varepsilon_{rx}^{''} > 0$ then kx are usually computed. Evaluation of the material absorption and shielding performance also requires for the knowledge of transmission properties of MSS multilayer system. In calculating TC at the last interface of the multilayer – i.e., at the absorber-air interface – a more general formalism, based on the application of boundary conditions for tangential fields, has been used in place of the transmission lines approach, the mathematical formulation is reported in the appendix.

WPO try to minimize a global objective function (GOF) which in turn are built using elementary objective function (EOFs). The EOF for TM and TE modes are named as $CostRC_{TM}$, $CostRC_{TE}$, for reflection coefficient, $CostTC_{TM}$, $CostTC_{TE}$, for the transmission coefficient, CostT, for thickness. The formal definition of the EOF for TM and TE modes are shown in (7)-(8)-(9)-(10). We can observe that for each particle (Pa) of WPO, the corresponding EOF is evaluated over the entire frequency band and over the entire incidence range angle

$$CostRC_{TM}(Pa) = \left[\sum_{freq=f\,\min}^{freq=f\,\max} \left(\sum_{\vartheta=\vartheta\min}^{\vartheta=\vartheta\max} RC_{TM}(Pa, freq, \vartheta) \right) \right]$$

$$CostRC_{TE}(Pa) = \left[\sum_{freq=f\,\min}^{freq=f\,\max} \left(\sum_{\vartheta=\vartheta\min}^{\vartheta=\vartheta\max} RC_{TE}(Pa, freq, \vartheta) \right) \right] \quad (7)$$

where Pa is the current particle of WPO particles family, freq is the frequency step, fmin and fmax is the start and stop frequency band, θ, is the current step angle, θmin and θmax represent the angular range bounds.

$$CostTC_{TM}(Pa) = \left[\sum_{freq=f\,\min}^{freq=f\,\max} \left(\sum_{\vartheta=\vartheta\min}^{\vartheta=\vartheta\max} TC_{TM}(Pa, freq, \vartheta) \right) \right]$$

$$CostTC_{TE}(Pa) = \left[\sum_{freq=f\,\min}^{freq=f\,\max} \left(\sum_{\vartheta=\vartheta\min}^{\vartheta=\vartheta\max} TC_{TE}(Pa, freq, \vartheta) \right) \right] \quad (8)$$

The definition of the *EOF* for thickness is

$$CostT(Ch) = [thick_{10}(Ch) + thick_9(Ch) + ... + thick_1(Ch)] \quad (9)$$

where $thick_x$ is the thickness of he *x-th* layer.

Two different weighting factors are introduced, one is called α, (0≤α≤1), weighting CostT w.r.t. CostRC and CostTC, the other one is called γ, (0≤γ≤1), weighting CostRC w.r.t. CostTC. Such weighting factors are chosen by the user and their meaning have to be intended as the capability of the tool to design the multilayer structure making privilege to the electromagnetic performances w.r.t. the thickness when α tend to 1, or making privilege in lowering the transmission coefficient (MSS, mainly in shielding applications) rather than lowering the reflection

coefficient (RAM mainly in absorbing applications) when γ tend to 0. Final GOF is a linear combination of the described EOFs

$$GOF(Pa) = \\ = \alpha \cdot [\gamma \cdot OF1(Pa) + (1-\gamma) \cdot OF2(Pa)] + (1-\alpha) \cdot OF3(Pa) \quad (27)$$

where

$$OF1(Pa) = \left(\frac{CostRC_{TM}(Pa)}{A} + \frac{CostRC_{TE}(Pa)}{A} \right)$$

$$OF2(Pa) = \left(\frac{CostTC_{TM}(Pa)}{A} + \frac{CostTC_{TE}(Pa)}{A} \right),$$

$$OF3(Pa) = \left[\frac{CostT(Pa)}{C} \right]$$

A and C are normalization factors defined as $(FreqStepNumber) \times (AngleStepNumber)$ *and C defined as* $[\max(LayerNumbers) \times \max(LayerThickness)]$.

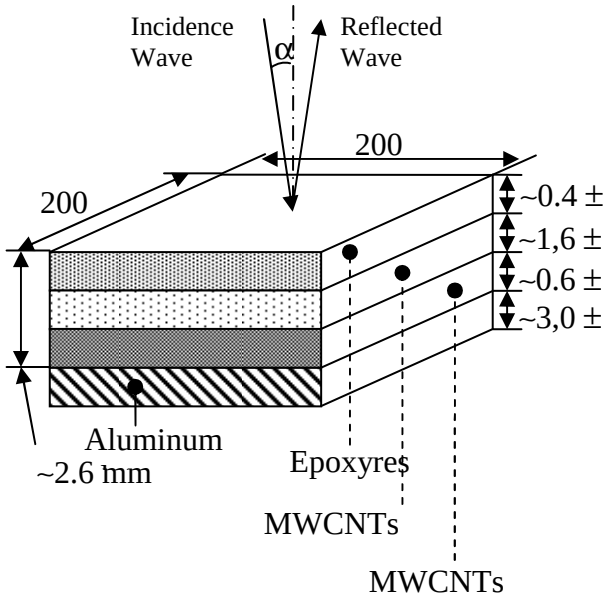


Fig. 2. RAM multilayer tile 250x2050 mm side

4 Experimental validation of mathematical model and design method

In order to experimentally validate the mathematical model and the design method described in previous sections, we manufactured and tested some tiles of composite materials.

RAM was simulated using only MWCNTs and a 250 x 250 mm related multilayer tile was manufactured. Choice of MWCNTs is mainly due to high absorption properties and to lower cost with respect to CNFs. RAM structure is schematically shown in Fig. 2.

Manufacturing process takes long time since about 20 hours is required for each layer to be accurately prepared and cured in oven. Ultrasonication 20 KHz has been exploited in order to obtain homogeneous disaggregation and dispersion of carbon nanopowders within the Epoxy-resin. Each layer has been made over the previous by injecting composite materials while was still in semi-liquid status in the sample holder. An in-house assembled press defined the thickness of each layer.

In Fig. 3 some pictures of tile manufacturing phases are shown.

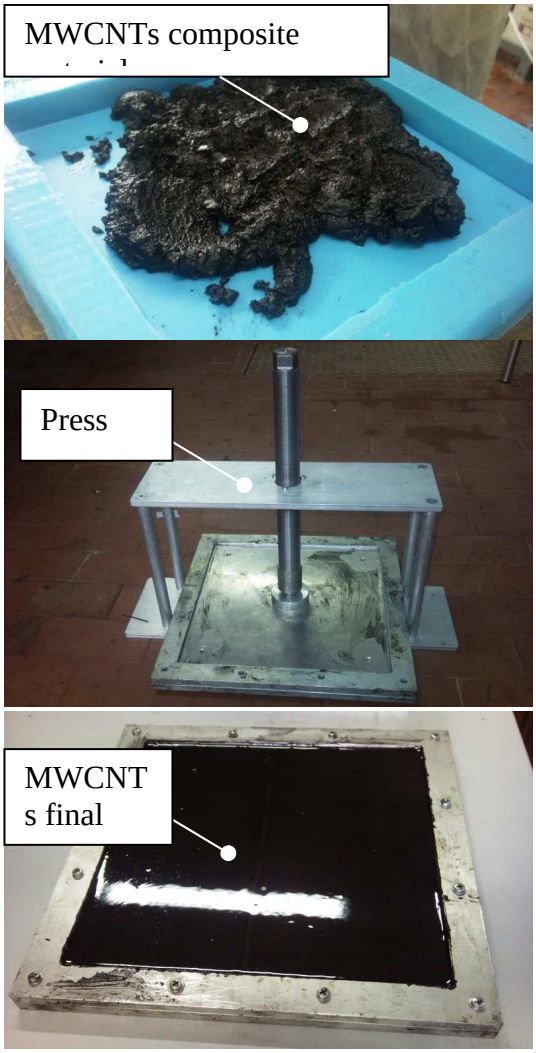
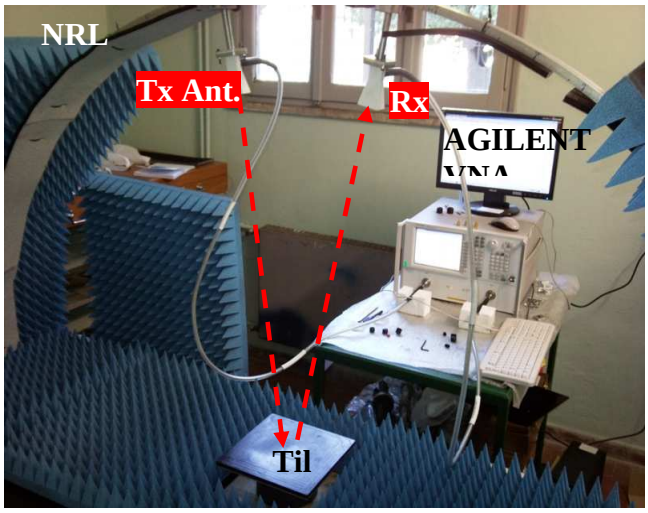
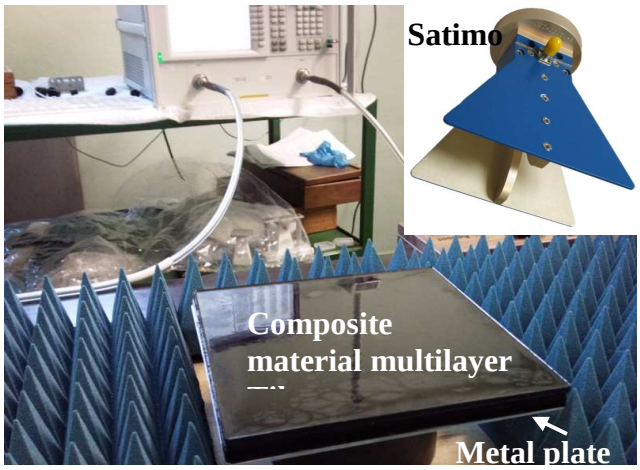


Fig. 3. a) MWCNTs composite; b) in-house constructed press,
c) final multilayer tile construction

In order to test reflection loss of tile we built an in-house NRL arch measurement, also called bi-static reflection method [22]-[23]. In this configuration, two antennas are used for transmitting and receiving signals respectively, and the microwave reflectivity at different incident angles can be measured. The NRL Arch is the industry standard for testing the reflectivity of materials. Reflectivity is defined as the reduction in reflected power caused by the introduction of an absorbent material. This reduction in reflected power is compared to a 'perfect' reflection which is approximated very well by the reflection of a flat metallic plate. A transmitting and receiving antenna are oriented towards the plate and can be located anywhere on the arch to allow measurements of performance at off normal angles of incidence. A network analyzer has been used for measurements on NRL Arch to provide both the stimulus and the measurement. It should be noted that in bi-static reflection measurements, the reflection is dependent on the polarization of the incident wave. Incident waves with parallel and perpendicular polarization usually result in different reflection coefficients. In Fig. 4, bistatic measurement system and the composite material tile are shown. In this configuration, bistatic measurement system is based on Agilent software 8571E (material measurement), and Agilent
PNA-L vector network analyzer. Antennas are SATIMO Dual Ridge Horn SH2000 type and are in the range 2-32 GHz. After calibration of the NRL bistatic system, measurement of a known sample consisting in ECCOSORB AN73 absorber has been performed in order to be aware about errors in the NRL measurement setup. Confidence of measurements was within 2 dB of interval with respect to reflection properties declared in ECCOSORB data sheet. At the end measurements of the manufactured RAM tile using respectively incidence angles of 5°, 20°, have been performed for TE and TM mode. In Fig.5, comparison between measured and simulated RAM is shown. It can be noticed that simulated curve shows similar behaviour with respect to measured one.

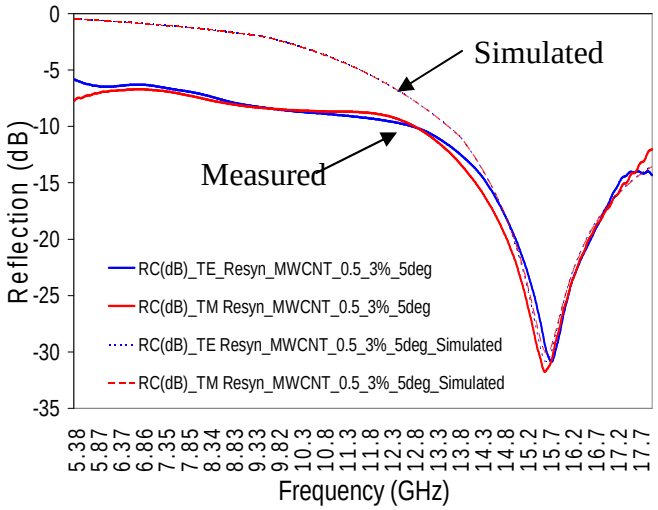


(b)

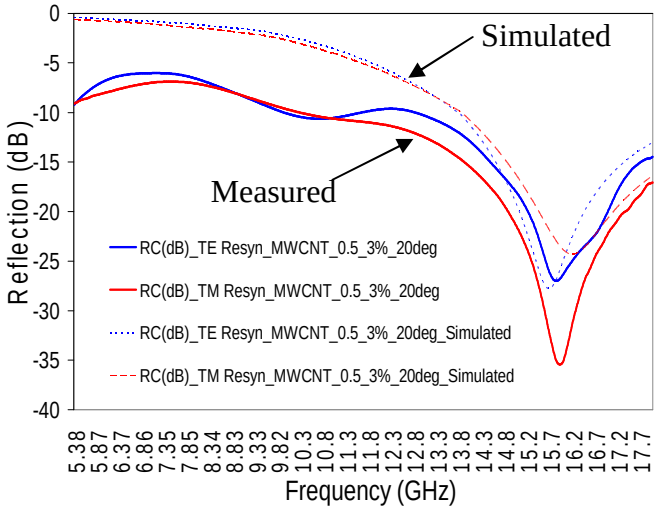


(c)

Fig. 4. a) bistatic measurement system, c) zoom of tile placed above the metal plate of bistatic measurement system and picture of SATIMO antenna.



(a)



(b)

Fig. 5. Comparison between simulated and measured 2.6 thickness RAM.

Reflection coefficient TE/TM for 5° Incidence angle

Reflection coefficient TE/TM for 20° Incidence angle

Differences in simulated and measured reflection loss are not greater then few dBs which are acceptable in such kind of experiment. Such differences are mainly due to:

- Difficulties in manufacturing tile multilayer structure. Here thickness of each layer has an accuracy of 0.1mm .
- Accuracy in determination of permittivity of composite materials using waveguide method.

• Intrinsic errors proper of bistatic measurement system mainly due to spurious reflections.

In terms of peak frequencies and corresponding reflection loss values, comparison between simulations and measurements confirm that mathematical model of absorber can be quite effectively used to design RAM structure.

two columns and should be 80 *mm* or 170 *mm* wide respectively.

5 Conclusion

The WPO based optimization techniques are an effective design methods having several degrees of freedom and so leaving the user free to decide of giving more importance to absorption properties with respect to the thickness or vice-versa. With the aid of optimization, we demonstrated the possibility of achieving very interesting microwave absorbing or shielding structures through the usage of carbon nanomaterials as filler in Resin-epoxy. Comparison between measured and simulated absorbers confirms validity of developed theoretical electromagnetic model.

References

[1] De Rosa, I.M.; Sarasini, F.; Sarto, M.S.; Tamburrano, A.” EMC Impact of Advanced Carbon Fiber/Carbon Nanotube Reinforced Composites for Next-Generation Aerospace Applications”, IEEE Transactions on Electromagnetic Compatibility, Digital Object Identifier: 10.1109/TEMC.2008.926818, Aug. 2008.

[2] Jihua Goua, Yong Tanga, Fei Lianga, Zhongfu Zhaoa, David Firsichb and Jennifer Fieldingc. “Carbon nanofiber paper for lightning strike protection of composite materials”, ScienceDirect Elsevier, doi:10.1016/j.compositesb.2009.06.009. July 2009.

[3] Zhuangjun Fana, Guohua Luo, a, , Zengfu Zhanga, Li Zhoua and Fei Wei, “Electromagnetic and microwave absorbing properties of multi-walled carbon nanotubes/polymer composites”, ScienceDirect, Elsevier, doi:10.1016/j.mseb.2006.02.045, 2006,

[4] Tianchun Zoua, , , Haipeng Lib, Naiqin Zhaoc and Chunsheng Shic, “Electromagnetic and microwave absorbing properties of multi-walled carbon nanotubes filled with Ni nanowire”, Journal of Alloys and Compounds, ScienceDirect, Elsevier, doi:10.1016/j.jallcom.2010.02.103., 2010.

[5] Yuzun Fana, Haibin Yangb, Xizhe Liua, Hongyang Zhub, Guangtian Zoub, “Preparation and study on radar absorbing materials of nickel-coated carbon fiber and flake graphite”, Journal of Alloys and Compounds, ScienceDirect, Elsevier, doi:10.1016/j.jallcom.2007.07.034, 2008.

[6] I.M. De Rosa, R. Mancinelli, F. Sarasini, M. S. Sarto, Senior, A. Tamburrano, " Electromagnetic Design and Realization of Innovative Fiber-Reinforced Broad-Band Absorbing Screens", IEEE Trans. on Electromagnetic Compatibility, Volume 51, Issue 3, Aug. 2009 Page(s):700 – 707.

[7] 3M, “New EMI material has rework potential”, ScienceDirect, Elsevier,doi:10.1016/S1350-4789(09)70250-8, 2009.

[8] Davide Micheli, Carmelo Apollo, Roberto Pastore, Mario Marchetti, “Modeling of Microwave Absorbing Structure using Winning Particle Optimization applied on Electrically Conductive Nanostructured composite material” , ICEM 2010 Roma- XIX International Conference on Electrical Machines, TPC-6: Theory, Modeling and Design, IEEE Industrial Electronics Society (IEE), September 6-8-2010 Rome, simposium, DOI: 10.1109/ICELMACH.2010.5607881

[9] Davide Micheli, Gabriele Gradoni. Carmelo Apollo, Roberto Pastore, Mario Marchetti. “Nanocomposite Materials, Theory and Application”, INECH, ISBN 978-953-7619-X-X. 2011

[10] Ignace Bogaert, Femke Olyslager, Yoeri Ari¨en, Davy Pissoort, “Modelling and Optimization of Advanced Multilayered Absorbers”, Proceedings of the 37th European Microwave Conference, Department of Information Technology, Ghent University, Emerson & Cuming Microwave Products, 2007.

[11] D. D. L. Chung, “Electromagnetic interference shielding effectiveness of carbon materials”, Carbon 39(2) 279–85, 2001.

[12] Yi Huang a, Ning Li a, Yanfeng Ma a, Feng Du a, Feifei Li a, Xiaobo He b, Xiao Lin b, Hongjun Gao b, Yongsheng Chen, “The influence of single-walled carbon nanotube structure on the electromagnetic interference shielding efficiency of its epoxy composites”, Carbon, 2007.

[13] Sui-Lin Shi and Ji Liang, “The effect of multi-wall carbon nanotubes on electromagnetic interference shielding of ceramic composites”,

Nanotechnology 19, 2008.
[14] Pitkethly, M.J., Radar absorbing materials and their potential use in aircraftstructures, IEEE Colloquium on Volume, Issue , 26 May 1992 Page(s):7/1 - 7/3.
[15] D. Micheli *, C. Apollo, R. Pastore, M. Marchetti. "X-Band microwave characterization of carbon-based nanocomposite material, absorption capability comparison and RAS design simulation", ScienceDirect, Composites Science and Technology, gen 2010.
[16] Davide Micheli, Roberto Pastore, Carmelo Apollo, Mario Marchetti, Gabriele Gradoni, Franco Moglie, Valter Mariani Primiani, "Carbon Based Nanomaterial Composites in RAM and Microwave Shielding Applications", 9th Nanotechnology Conference: IEEE NANO 2009, 2009.
[17] Davide M. Pozar, Microwave Engineering third edition, John Wiley, pp 174-180, 2005.
[18] Ramo, Whinnery, Van Duzer, Fields and Waves in Communications Electronics, JohnWiley and Son, 1994.
[19] Jin Au Kong, Electromagnetic Wave Theory, John Wiley & Sons, 1986.
[20] Sophocles J.Orfanidis, Electromagnetic Waves and Antennas, Rutgers University, pp 258 - 261. 2008.
[21] Juslius Adam Stratton. “Electromagnetic wave Theory”, IEEE Press Series on Electromagentic Wave Theory. 2007.
[22] Umari, M. H. Ghodgaonkar, D. K. Varadan, V. V. and Varadan, V. K. (1991). “A free-space bistatic calibration technique for the measurement of parallel and perpendicular reflection coefficients of planar samples”, IEEE Transactions on Instrumentation and Measurement, 40 (1), 19–24.
[23] Emerson & Cuming Microwave Products, NRL Arch Reflectivity Testing basic notes. Available: *http://www.eccosorb.com/*

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Analytical solution for the dynamic analysis of magneto-electro-elastic laminated beams

A. Milazzo
Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale, Università di Palermo, Viale delle Scienze, 90128 Palermo - Italy

Keywords: *Magneto-Electro-Elastic Composites, Laminated Beam, Free Vibrations, Smart Structures*

Abstract

A model is presented for the dynamic problem of magneto-electro-elastic generally laminated beams. A first-order mechanical beam theory is used and the electric and magnetic fields are assumed to be quasi-static. Different magneto-electric boundary conditions on the beam top and bottom surfaces are considered. The problem resolving equations, that are the equations of motions, are written. They involve the same terms of the elastic dynamic problem entering with effective stiffness coefficients, which take the magneto-electro-mechanical couplings into account. Additional terms, which involve the third spatial derivative of the transverse displacement, occur as a result of the piezoelectric and/or piezomagnetic behavior. It is also shown that the magneto-electric inputs can be treated as equivalent external axial forces and bending moments per unit length. Free vibrations analyses are presented to validate the model and show its effectiveness.

1 Introduction

Composites containing piezoelectric phases and piezomagnetic phases show a magneto-electric coupling effect which results in the ability to transform magnetic, electric and mechanical energies from one type to the others. For this reason, which makes magneto-electro-elastic composites suitable for applications in sensor and actuator devices and smart structure technologies, they are receiving increasing attention from the research and industrial communities. The magneto-electric effect, which refers to the capability of passively coupling the electric and magnetic fields [1, 2], is associated with the coexistence of piezoelectric and piezomagnetic phases that provide the coupling through the elastic field [3, 4]. This consideration addresses towards the employement of multilayer structures where piezoelectric and piezomagnetic layers are used to achieve the desired coupling effect. It is worth noting that the multilayered configuration has proven to be more effective for magneto-electric coupling with respect to particulate composites [5].

The promising application of magneto-electro-elastic composites in the form of multilayered structures makes mandatory the development of new methods and analysis tools with the aim to better understand their behaviour when subjected to mechanical, electrical and magnetic loads. Recently, researches have been focused on the magneto-electro-elastic laminate dynamics which models

the widely employed structural conditions for practical applications. Exact benchmark solutions for the static behaviour of magneto-electro-elastic laminates have been presented [6] and the approach extended to free vibration analysis [7]. State space approach has been proposed [8, 9] and finite elements have been extensively used to deal with multifield structural problems [10] and magneto-electro-elastic laminates [11, 12]. Boundary elements [13, 14] and meshless [15] approaches have been also proposed.

In the present paper a model for the dynamic analysis of general magneto-electro-elastic laminated beams is presented. Actually, the dynamic behavior of magneto-electro-elastic beams is highly influenced by the shear deformation even for low frequency vibrations [12, 16]. Thus, a first-order beam theory is employed to model the beam mechanical behaviour. Quasi-static behaviour is assumed for the electro-magnetic quantities with no electric charge and current density present. The equations of motion for the laminated beam are written after the magneto-electric problem is at first partially solved in terms of the beam generalized kinematical variables, that is axial and transverse displacement and cross sectional rotation. This leads to a system of coupled differential equations whose solution provides the beam generalized displacements. Once the mechanical problem is solved the quantities associated with the electro-magnetic behaviour are straightforwardly recovered. To validate the model the free vibrations problem is solved and results for different boundary conditions and loads are compared with literature solutions.

2 Problem statement and governing equations

Let us consider a multilayered beam having length L and cross section A, referred to the cartesian coordinate system $x-z$ as shown in Fig. 1. The beam consists of N magneto-electro-elastic layers with poling direction parallel to the z-axis. The k-th layer has faces located at $z = h_{k-1}$ and $z = h_k$, with $k = 1, \dots, N$. The layers are perfectly bonded.

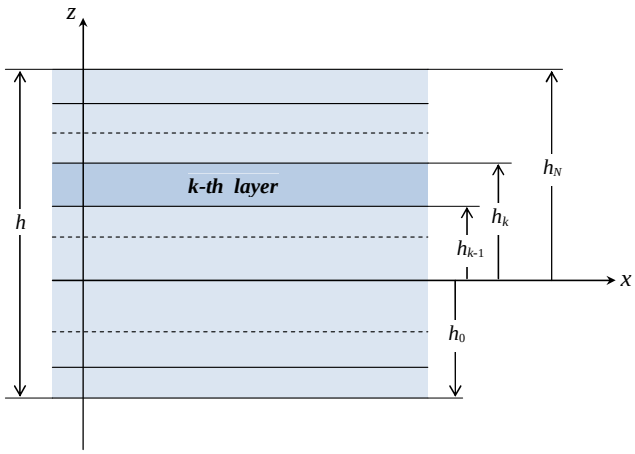


Fig. 1 Beam lamination scheme

According to the first-order beam theory [17] the kinematical model for the beam deformation is given by

$$u(x,\mathrm{z},\tau) = u_0(x,\tau) - z\,\vartheta(x,\tau) \qquad (1.a)$$

$$w(x,\mathrm{z},\tau) = w_0(x,\tau) \qquad (1.b)$$

where u and w are the displacement components of the generic beam point, u_0 and w_0 are the axial and transverse diplacement of the beam axis points whereas ϑ is the beam section rotation. τ denotes the time. Ignoring the effect of free current and time variation of electric diplacement and magnetic induction, the electric and magnetic states are described in terms of the scalar electric potential $\Phi = \Phi(x,z,\tau)$ and magnetic potential $\Psi = \Psi(x,z,\tau)$. The strain-displacement relationships read as

$$\gamma_{xx} = \frac{\partial u_0}{\partial x} - z\,\frac{\partial \vartheta}{\partial x} \qquad (2.a)$$

$$\gamma_{xz} = -\vartheta + \frac{\partial w_0}{\partial x} \qquad (2.b)$$

whereas the electric field components E_x and E_z and the magnetic field components H_x and H_z are obtained through the gradient laws. One has

$$E_x = -\frac{\partial \Phi}{\partial x}; \quad E_z = -\frac{\partial \Phi}{\partial z} \qquad (3.a,b)$$

$$H_x = -\frac{\partial \Psi}{\partial x}; \quad H_z = -\frac{\partial \Psi}{\partial z} \qquad (3.c,d)$$

To describe the magneto-electro-elastic state of the beam the stresses σ_{xx} and σ_{xz}, the electric displacements D_x and D_z and the magnetic

induction components B_x and B_z are introduced. Following the assumpion made in Ref. [10] and extending it also to the the magnetic induction [18], the effect of the axial components of the electric field and magnetic induction is neglected with respect to that of the corresponding components along the poling direction. Consequently, the constitutive law for each of the magneto-electro-elastic layers of the beam writes as

$$\sigma_{xx}^{\langle k\rangle} = c^{\langle k\rangle}\gamma_{xx} - e^{\langle k\rangle}E_z - d^{\langle k\rangle}H_z \qquad (4.a)$$

$$\sigma_{xz}^{\langle k\rangle} = c_{44}^{\langle k\rangle}\gamma_{xz} \qquad (4.b)$$

$$D_x^{\langle k\rangle} = e_{14}^{\langle k\rangle}\gamma_{xz} \qquad (4.c)$$

$$D_z^{\langle k\rangle} = e^{\langle k\rangle}\gamma_{xx} + \varepsilon^{\langle k\rangle}E_z^{\langle k\rangle} + \eta^{\langle k\rangle}H_z \qquad (4.d)$$

$$B_x^{\langle k\rangle} = d_{14}^{\langle k\rangle}\gamma_{xz} \qquad (4.e)$$

$$B_z^{\langle k\rangle} = d^{\langle k\rangle}\gamma_{xx} + \eta^{\langle k\rangle}E_z^{\langle k\rangle} + \mu^{\langle k\rangle}H_z \qquad (4.f)$$

where $c^{\langle k\rangle}$ and $c_{44}^{\langle k\rangle}$ are the elastic stiffness constants, $\varepsilon^{\langle k\rangle}$ and $\mu^{\langle k\rangle}$ are the dielectric constant and magnetic permeability, $e^{\langle k\rangle}$ and $e_{14}^{\langle k\rangle}$ are the piezoelectric costants, $d^{\langle k\rangle}$ and $d_{14}^{\langle k\rangle}$ are the piezomagnetic coupling coefficients and $\eta^{\langle k\rangle}$ denotes magneto-electric coupling coefficients. The notation $\bullet^{\langle k\rangle}$ indicates quantities associated with the k-th layer.

Finally, assuming that the electro-magnetic state of the beam can be modeled as quasi-static because the electro-magnetic waves propagation velocity is several order of magnitude higher than that of elastic waves, the governig equation for the problem are given by the beam equilibrium equations and Gauss laws for electrostatics and magnetostatics

$$\frac{\partial P}{\partial x} + p_x = J_0\frac{\partial^2 u_0}{\partial \tau^2} \qquad (5.a)$$

$$\frac{\partial V}{\partial x} + p_z = J_0\frac{\partial^2 w_0}{\partial \tau^2} \qquad (5.b)$$

$$\frac{\partial M}{\partial x} + V + m = J_1\frac{\partial^2 \vartheta}{\partial \tau^2} \qquad (5.c)$$

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_z}{\partial z} = 0 \qquad (5.d)$$

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_z}{\partial z} = 0 \qquad (5.e)$$

where P is the axial force, V is the shear force, M is the bending moment, p_x and p_z are the external axial and transverse force per unit length, m is the applied external moment per unit length. $J_0 = \int_A \rho\, dA$ and $J_1 = \int_A \rho z^2\, dA$ denote the inertia properties of the beam section.

The governing equations previously introduced have to be completed by the interface continuity conditions and the appropriate boundary conditions. The mechanical interface continuity conditions are trivially satisfied because of the kinematical model chosen. On the other hand, as regards the electro-magnetic problem the interface continuity conditions need to be satisfied and for $h = 1, \dots, N-1$ they read as

$$\Phi^{\langle k\rangle} = \Phi^{\langle k+1\rangle} \quad @\ z = h_k \qquad (6.a)$$

$$\Psi^{\langle k\rangle} = \Psi^{\langle k+1\rangle} \quad @\ z = h_k \qquad (6.b)$$

$$D_z^{\langle k\rangle} = D_z^{\langle k+1\rangle} \quad @\ z = h_k \qquad (6.c)$$

$$B_z^{\langle k\rangle} = B_z^{\langle k+1\rangle} \quad @\ z = h_k \qquad (6.d)$$

The mechanical boundary conditions are those of the first-order beam theory and are not discussed here for the sake of brevity. The electro-magnetic boundary conditions can be of different type and can be assigned at the laminate bottom and top surfaces or at the interfaces. Without loss of generality for the proposed approach, we focus on a class of common problems for which at least one surface of the beam, namely the bottom surface, is electroded and specified with the magnetic and electric potential whereas a second surface, namely the top surface, is specified with the electric and magnetic potentials or the electric displacements and magnetic induction. Thus electro-magnetic boundary conditions read as

$$\Phi(x, h_0, \tau) = \tilde{\Phi}_0 = \xi_1 \qquad (7.a)$$

$$\Psi(x, h_0, \tau) = \tilde{\Psi}_0 = \xi_2 \qquad (7.b)$$

$$\Phi(x, h_N, \tau) = \tilde{\Phi} = \xi_3 \quad \text{or} \quad D_z(x, h_N, \tau) = \tilde{D}_z = \xi_3 \qquad (7.c)$$

$$\Psi(x, h_N, \tau) = \tilde{\Psi} = \xi_4 \quad \text{or} \quad B_z(x, h_N, \tau) = \tilde{B}_z = \xi_4 \qquad (7.d)$$

In the preceding relationships the tilde denotes prescribed quantities. It is worth noting that other combinations of the electro-magnetic boundary conditions can be imposed. They are straightforwardly treated analogously to those presented in the paper.

3 Magneto-electro-elastic multilayer beam model

The strategy employed to derive the problem solution is based on a preliminary integration of the electro-magnetic governing equations. This allows to express the electro-magnetic variables Φ and Ψ in terms of the mechanical variables u_0, w_0 and ϑ. Finally, the problem is solved after substitution of the expressions obtained for the electro-magnetic quantities into the mechanical governing equations. In order to simplify some involved manipulations, matrix notation is employed where useful.

3.1 Magneto-electric state

Let us introduce the vector $\boldsymbol{\Phi} = \{\Phi \quad \Psi\}^T$, containing the electric and magnetic potentials, and the matrices

$$\boldsymbol{R}^{\langle k\rangle} = \begin{bmatrix} \varepsilon^{\langle k\rangle} & \eta^{\langle k\rangle} \\ \eta^{\langle k\rangle} & \mu^{\langle k\rangle} \end{bmatrix} \tag{8.a}$$

$$\boldsymbol{W}^{\langle k\rangle} = \begin{bmatrix} e^{\langle k\rangle} \\ d^{\langle k\rangle} \end{bmatrix} \tag{8.b}$$

$$\boldsymbol{W}_{14}^{\langle k\rangle} = \begin{bmatrix} e_{14}^{\langle k\rangle} \\ d_{14}^{\langle k\rangle} \end{bmatrix} \tag{8.c}$$

Taking the constitutive relationships Eqs. (4), the strain-displacement relationships Eqs. (2) and the gradient relationships Eqs. (3) into account, the Gauss's laws for electrostatic and magnetostatic, *i.e.* Eqs. (5.d) and (5.e), are written for the k-th layer as

$$\boldsymbol{R}^{\langle k\rangle}\frac{\partial^2\boldsymbol{\Phi}^{\langle k\rangle}}{\partial z^2} = \boldsymbol{W}_{14}^{\langle k\rangle}\frac{\partial^2 w_0}{\partial x^2} - [\boldsymbol{W}_{14} + \boldsymbol{W}]^{\langle k\rangle}\frac{\partial\vartheta}{\partial x} \tag{9}$$

Integration of Eq. (9) provides

$$\boldsymbol{\Phi}^{\langle k\rangle} = \left(\boldsymbol{A}^{\langle k\rangle}\frac{\partial\vartheta}{\partial x} + \boldsymbol{B}^{\langle k\rangle}\frac{\partial^2 w_0}{\partial x^2}\right)\frac{z^2}{2} + \boldsymbol{a}^{\langle k\rangle}z + \boldsymbol{b}^{\langle k\rangle} \tag{10}$$

where $\boldsymbol{a}^{\langle k\rangle}$ and $\boldsymbol{b}^{\langle k\rangle}$ are vectors of integration constants, $\boldsymbol{S}^{\langle k\rangle} = \boldsymbol{R}^{\langle k\rangle^{-1}}$ and

$$\boldsymbol{A}^{\langle k\rangle} = -[\boldsymbol{S}(\boldsymbol{W} + \boldsymbol{W}_{14})]^{\langle k\rangle} \tag{11.a}$$

$$\boldsymbol{B}^{\langle k\rangle} = [\boldsymbol{S}\boldsymbol{W}_{14}]^{\langle k\rangle} \tag{11.b}$$

Accordingly, from the constitutive law one has

$$\boldsymbol{\Delta}_x^{\langle k\rangle} = \boldsymbol{W}_{14}^{\langle k\rangle}\frac{\partial v}{\partial x} - \boldsymbol{W}_{14}^{\langle k\rangle}\vartheta \tag{12.a}$$

$$\boldsymbol{\Delta}_z^{\langle k\rangle} = \boldsymbol{W}^{\langle k\rangle}\frac{\partial u_0}{\partial x} - \boldsymbol{W}_{14}^{\langle k\rangle}z\frac{\partial\vartheta}{\partial x} - \boldsymbol{W}_{14}^{\langle k\rangle}z\frac{\partial^2 w_0}{\partial x^2} - [\boldsymbol{R}\boldsymbol{a}]^{\langle k\rangle} \tag{12.b}$$

where $\boldsymbol{\Delta}_x = \{D_x \quad B_x\}^T$ and $\boldsymbol{\Delta}_z = \{D_z \quad B_z\}^T$ are the vectors containing the electric displacements and magnetic induction x and z components.

The integration constants $\boldsymbol{a}^{\langle k\rangle}$ and $\boldsymbol{b}^{\langle k\rangle}$ have to satisfy $N-1$ relationships, determined by imposing the electro-magnetic continuity conditions at the laminate interfaces. Actually, from the interface continuity conditions of $\boldsymbol{\Delta}_z$, namely Eqs (6.c) and (6.d), for $k = 1, \dots, N-1$ one deduces that

$$\boldsymbol{a}^{\langle k\rangle} = \left[\boldsymbol{S}^{\langle k\rangle}\boldsymbol{R}^{\langle k+1\rangle}\right]\boldsymbol{a}^{\langle k+1\rangle} - \left[\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}|_{\langle k\rangle}^{\langle k+1\rangle}\right]\frac{\partial u_0}{\partial x} + \left[h_k\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}_{14}|_{\langle k\rangle}^{\langle k+1\rangle}\right]\frac{\partial\vartheta}{\partial x} + \left[h_k\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}_{14}|_{\langle k\rangle}^{\langle k+1\rangle}\right]\frac{\partial^2 w_0}{\partial x^2} \tag{13}$$

where $|\bullet|_{\langle k\rangle}^{\langle k+1\rangle}$ denotes the difference of quantities pertaining to the $(k+1)$-th and k-th layers. Again, from the interface continuity conditions of $\boldsymbol{\Phi}$, namely Eqs. (6.a) and (6.b), it follows that for $k = 1, \dots, N-1$ one has

$$\boldsymbol{b}^{\langle k+1\rangle} = \boldsymbol{b}^{\langle k\rangle} + \left[\left(\boldsymbol{S}^{\langle k\rangle}\boldsymbol{R}^{\langle k+1\rangle} - \boldsymbol{I}\right)h_k\right]\boldsymbol{a}^{\langle k+1\rangle} - \left[h_k\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}|_{\langle k\rangle}^{\langle k+1\rangle}\right]\frac{\partial u_0}{\partial x} + \left[\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}_{14}|_{\langle k\rangle}^{\langle k+1\rangle}h_k^2 - -\frac{1}{2}|\boldsymbol{S}(\boldsymbol{W} + \boldsymbol{W}_{14})|_{\langle k\rangle}^{\langle k+1\rangle}h_k^2\right]\frac{\partial\vartheta}{\partial x} + \left[\left(\boldsymbol{S}^{\langle k\rangle}|\boldsymbol{W}_{14}|_{\langle k\rangle}^{\langle k+1\rangle} - \frac{1}{2}|\boldsymbol{S}\boldsymbol{W}_{14}|_{\langle k\rangle}^{\langle k+1\rangle}\right)h_k^2\right]\frac{\partial^2 w_0}{\partial x^2} \tag{14}$$

where $\boldsymbol{I}$ is the identity matrix. The integration constants $\boldsymbol{a}^{\langle k\rangle}$ and $\boldsymbol{b}^{\langle k\rangle}$ are determined from Eqs (13) and (14) together with the electro-magnetic boundary conditions at the bottom and top surfaces (or at the interfaces). One writes

$$\boldsymbol{a}^{\langle k\rangle} = \boldsymbol{\alpha}_u^{\langle k\rangle}\frac{\partial u_0}{\partial x} + \boldsymbol{\alpha}_\vartheta^{\langle k\rangle}\frac{\partial \vartheta}{\partial x} + \boldsymbol{\alpha}_w^{\langle k\rangle}\frac{\partial^2 w_0}{\partial x^2} + \sum_{i=1}^{4}\boldsymbol{\alpha}_{\xi i}^{\langle k\rangle}\xi_i \qquad (15.a)$$

$$\boldsymbol{b}^{\langle k\rangle} = \boldsymbol{\beta}_u^{\langle k\rangle}\frac{\partial u_0}{\partial x} + \boldsymbol{\beta}_\vartheta^{\langle k\rangle}\frac{\partial \vartheta}{\partial x} + \boldsymbol{\beta}_v^{\langle k\rangle}\frac{\partial^2 w_0}{\partial x^2} + \sum_{i=1}^{4}\boldsymbol{\beta}_{\xi i}^{\langle k\rangle}\xi_i \qquad (15.b)$$

The expressions of the coefficient vectors appearing in Eqs. (15.a) and (15.b) can be determined in closed form. For the considered magneto-electric boundary conditions, they are given in the Appendix.

3.2 Mechanical state

Once the electric and magnetic potentials are resumed in terms of the kinematical generalized variables the expression of the stresses is inferred through the constitutive laws

$$\sigma_{xx}^{\langle k\rangle} = \left[c + \widehat{\boldsymbol{W}}\boldsymbol{\alpha}_u\right]^{\langle k\rangle}\frac{\partial u_0}{\partial x} + \left[(\widehat{\boldsymbol{W}}\boldsymbol{A} - c)z + \widehat{\boldsymbol{W}}\boldsymbol{\alpha}_\vartheta\right]^{\langle k\rangle}\frac{\partial \vartheta}{\partial x} + \left[\widehat{\boldsymbol{W}}\boldsymbol{B}z + \widehat{\boldsymbol{W}}\boldsymbol{\alpha}_w\right]^{\langle k\rangle}\frac{\partial^2 w_0}{\partial x^2} + \sum^{4}\left[\widehat{\boldsymbol{W}}\boldsymbol{\alpha}_{\xi i}\right]^{\langle k\rangle}\xi_i \qquad (16.a)$$

$$\sigma_{xz}^{\langle k\rangle} = c_{44}^{\langle k\rangle}\left(\frac{\partial w_0}{\partial x} - \vartheta\right) \qquad (16.b)$$

where $\widehat{\boldsymbol{W}}^{\langle k\rangle}$ denotes the transpose of $\boldsymbol{W}^{\langle k\rangle}$. Consequently the beam axial and shear forces and the bending moment are expressed as

$$P = \int_A \sigma_{xx}dA = \mathrm{K}_{Pu}\frac{\partial u_0}{\partial x} + \mathrm{K}_{P\vartheta}\frac{\partial \vartheta}{\partial x} + \mathrm{K}_{Pw}\frac{\partial^2 w_0}{\partial x^2} + \sum_{i=1}^{4}F_{Pi}\xi_i \qquad (17.a)$$

$$V = \int_A \sigma_{xz}dA = \mathrm{K}_{Vw}\frac{\partial w_0}{\partial x} - \mathrm{K}_{V\vartheta}\vartheta \qquad (17.b)$$

$$M = \int_A \sigma_{xx}zdA = \mathrm{K}_{Mu}\frac{\partial u_0}{\partial x} + \mathrm{K}_{M\vartheta}\frac{\partial \vartheta}{\partial x} + \mathrm{K}_{Mw}\frac{\partial^2 w_0}{\partial x^2} + \sum_{i=1}^{4}F_{Mi}\xi_i \qquad (17.c)$$

where the effective stiffness coefficients $\mathrm{K}_{\alpha\beta}$ are defined by

$$\mathrm{K}_{Pu} = \sum_{k=1}^{N}\left[(c + \widehat{\boldsymbol{W}}\boldsymbol{\alpha}_u)\mathfrak{I}_0\right]^{\langle k\rangle} \qquad (18.a)$$

$$\mathrm{K}_{P\vartheta} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}(\boldsymbol{A}\mathfrak{I}_1 + \boldsymbol{\alpha}_\vartheta\mathfrak{I}_0) - c\mathfrak{I}_1\right]^{\langle k\rangle} \qquad (18.b)$$

$$\mathrm{K}_{Pw} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}(\boldsymbol{B}\mathfrak{I}_1 + \boldsymbol{\alpha}_w\mathfrak{I}_0)\right]^{\langle k\rangle} \qquad (18.c)$$

$$\mathrm{K}_{Vw} = \mathrm{K}_{V\vartheta} = \Lambda\sum_{k=1}^{N}\left[c_{44}\,\mathfrak{I}_0\right]^{\langle k\rangle} \qquad (18.d)$$

$$\mathrm{K}_{Mu} = \sum_{k=1}^{N}\left[(c + \widehat{\boldsymbol{W}}\boldsymbol{\alpha}_u)\mathfrak{I}_1\right]^{\langle k\rangle} \qquad (18.e)$$

$$\mathrm{K}_{M\vartheta} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}(\boldsymbol{A}\mathfrak{I}_2 + \boldsymbol{\alpha}_\vartheta\mathfrak{I}_1) - c\mathfrak{I}_2\right]^{\langle k\rangle} \qquad (18.f)$$

$$\mathrm{K}_{Mw} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}(\boldsymbol{B}\mathfrak{I}_2 + \boldsymbol{\alpha}_w\mathfrak{I}_1)\right]^{\langle k\rangle} \qquad (18.g)$$

and

$$F_{Pi} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}\boldsymbol{\alpha}_{\xi i}\mathfrak{I}_0\right]^{\langle k\rangle} \qquad (19.a)$$

$$F_{Mi} = \sum_{k=1}^{N}\left[\widehat{\boldsymbol{W}}\boldsymbol{\alpha}_{\xi i}\mathfrak{I}_1\right]^{\langle k\rangle} \qquad (19.b)$$

In the former equations one sets $\mathfrak{I}_0^{\langle k\rangle} = \int_{A_k} dA$, $\mathfrak{I}_1^{\langle k\rangle} = \int_{A_k} z\,dA$ and $\mathfrak{I}_2^{\langle k\rangle} = \int_{A_k} z^2\,dA$, being A_k the cross section of the k-th layer. Λ is the beam section shear factor.

The expressions of the effective stiffness coefficients show that the classical stiffness coefficients of the pure elasticity case are modified by the effects of the magneto-electro-mechanical couplings. Moreover, the multifield

couplings give rise to further stiffness coefficients, namely K_{Pw} and K_{Mw}, which depend only on the piezoelectric and piezomagnetic material properties. These coefficients represent a contribution to the axial force and bending moment associated with the second derivative of the transverse displacement which activate only when piezoelectric and piezomagnetic behavior is present. Furthermore, both the axial force and the bending moment directly depend not only on the mechanical variables but also on the applied magneto-electric loads, entering the model through the boundary conditions, *i.e.* Eqs. (7). These loads give rise to contributions to the axial force and bending moment which, as a consequence, affect also the beam mechanical boundary conditions.

3.3 Equations of motion

Substituting Eqs. (17) into the equilibrium equations, Eqs (5.a), (5.b) and (5.c), one obtains the following equations of motion which represent the magneto-electro-elastic beam model

$$K_{Pu}\frac{\partial^2 u_0}{\partial x^2}+K_{P\vartheta}\frac{\partial^2 \vartheta}{\partial x^2}+K_{Pw}\frac{\partial^3 w_0}{\partial x^3} -J_0\frac{\partial^2 u_0}{\partial \tau^2}=-p_x-\sum_{i=1}^{4}F_{Pi}\xi_i \qquad (20.a)$$

$$K_{Vw}\frac{\partial^2 w_0}{\partial x^2}-K_{V\vartheta}\frac{\partial \vartheta}{\partial x}-J_0\frac{\partial^2 w_0}{\partial \tau^2}=-p_z \qquad (20.b)$$

$$K_{Mu}\frac{\partial^2 u_0}{\partial x^2}+K_{M\vartheta}\frac{\partial^2 \vartheta}{\partial x^2}+K_{Mw}\frac{\partial^3 w_0}{\partial x^3}+ K_{Vw}\frac{\partial w_0}{\partial x}-K_{V\vartheta}\vartheta-J_2\frac{\partial^2 \vartheta}{\partial \tau^2}=-m-\sum_{i=1}^{4}F_{Mi}\xi_i \qquad (20.c)$$

It is observed that the magneto-electro-elastic beam governing equations involve the same terms of the pure elastic beam whose coefficients are replaced by the magneto-electro-elastic effective stiffnesses defined by Eqs. (18.a), (18.b), (18.d), (18.e) and (18.f). These effective stiffnesses take the effects of the electro-magneto-elastic couplings into account and reduces to classical mechanical expressions if piezoelectric and piezomagnetic coupling coefficients are zero. In addition new terms appear in force of the piezoelectric and piezomagnetic behavior, which involve the third x-derivative of the transverse displacements. The effect of the applied magneto-electric loads, imposed through the boundary conditions on the beam surfaces and/or interfaces, enters the model as equivalent axial and bending loads given by Eqs. (18.h) and (18.i). It should be remarked that the magneto-electro-elastic coupling in the presence of magneto-electric loads give rise to contributions to the axial force and bending moment which actually modify the mechanical boundary conditions of the beam, which generally results into non-homogeneous and time-dependent expressions. This requires suitable techniques for the solution [19].

The solution of the system of Eqs. (20) provides the time-variation of the beam generalized displacements u_0, w_0 and ϑ. Once these quantities are determined the electro-magnetic state is straightforwardly recovered by using Eqs. (10) and (12).

4 Numerical results and model validation

In this section some applications of the proposed method are presented with the aim to validate it. To this aim the free vibrations problem is dealt with. According to the observations reported in the preceding section, to define the free vibrations problem for the magneto-electro-elastic beam homogeneous magneto-electric boundary conditions have to be considered. Otherwise, the magneto-electric boundary conditions generate mechanical forces which lead the problem to the case of forced vibrations. Therefore, the model for free vibrations analysis reduces to

$$K_{Pu}\frac{\partial^2 u_0}{\partial x^2}+K_{P\vartheta}\frac{\partial^2 \vartheta}{\partial x^2}+K_{Pw}\frac{\partial^3 w_0}{\partial x^3} +J_0\omega^2 u_0=0 \qquad (21.a)$$

$$K_{Vw}\frac{\partial^2 w_0}{\partial x^2}-K_{V\vartheta}\frac{\partial \vartheta}{\partial x}+J_0\omega^2 w_0=0 \qquad (21.b)$$

$$K_{Mu}\frac{\partial^2 u_0}{\partial x^2}+K_{M\vartheta}\frac{\partial^2 \vartheta}{\partial x^2}+K_{Mw}\frac{\partial^3 w_0}{\partial x^3}+ K_{Vw}\frac{\partial w_0}{\partial x}-(K_{V\vartheta}-J_2\omega^2)\vartheta=0 \qquad (21.c)$$

where ω is the circular frequency. The eigenvalues of the system, that is the mode frequencies, are determined by following the procedure reported in Ref. [18]

The free vibrations of a magneto-electro-elastic beam consisting of three layers of $BaTiO_3$ followed by three layers of $CoFe_3O_4$ are analyzed. The employed material properties are given in Ref. [12] to which the reader is referred for more details on the analyzed problem. The beam, with rectangular cross section, has length $L = 0.3\,m$ and unit width. The layers have equal thickness and the laminate is $0.02\,m$ thick. The cross section shear factor has been set equal to $8/9$. Different mechanical boundary conditions have been considered (clamped-free, clamped-clamped and simply-supported).

The results obtained are given in Table 1, 2 and 3 where the comparison with solutions found in the literature and obtained by finite elements (FEM) [12] is presented. Different magneto-electric boundary conditions are considered highlighting that they do not noticeably influence the natural frequencies. The results obtained are in very good agreement with those present in the literature. Some discrepancies are observed for the simply-supported case for which the boundary conditions at the beam ends of the present one-dimensional model and the two-dimensional model of Ref. [12] slightly differ.

The presented results validate the approach and show its effectiveness.

Mode	$\varphi(x,h_N)=\psi(x,h_N)=0$		$D_z(x,h_N)=B_z(x,h_N)=0$		
	Present	% Error	Present	% Error	FEM
1	187.74	0.51	187.74	0.51	188.70
2	1148.19	0.56	1148.25	0.55	1154.65
3	3100.12	0.66	3100.48	0.65	3120.80
4	4420.85	1.98	4421.01	1.98	4335.11
5	5791.14	0.81	5792.31	0.79	5838.36
6	9061.56	1.02	9064.30	0.99	9154.66
7	12761.49	1.26	12766.79	1.22	12924.90
8	13262.07	2.00	13262.59	2.01	13001.50
9	16779.19	1.51	16787.91	1.46	17036.50
10	21010.42	1.84	21023.12	1.78	21405.10

Table 1. Natural frequency for a six layers $BaTiO_3/CoFe_3O_4$ clamped-free beam with different magneto-electric boundary conditions

Mode	$\varphi(x,h_N)=\psi(x,h_N)=0$		$D_z(x,h_N)=B_z(x,h_N)=0$		
	Present	% Error	Present	% Error	FEM
1	1156.61	0.73	1156.66	0.72	1165.07
2	3053.21	0.82	3053.54	0.81	3078.60
3	5683.77	0.96	5684.87	0.94	5738.67
4	8834.89	1.74	8834.84	1.74	8683.41
5	8874.39	1.08	8877.37	1.05	8971.57
6	12468.40	1.36	12473.28	1.32	12640.40
7	16373.96	1.62	16382.07	1.57	16643.20
8	17674.39	1.82	17675.26	1.83	17358.00
9	20504.93	1.89	20517.25	1.83	20900.40
10	24798.53	2.21	24815.44	2.15	25360.00

Table 2. Natural frequency for a six layers $BaTiO_3/CoFe_3O_4$ clamped-clamped beam with different magneto-electric boundary conditions

Mode	$\varphi(x,h_N)=\psi(x,h_N)=0$		$D_z(x,h_N)=B_z(x,h_N)=0$		
	Present	% Error	Present	% Error	FEM
1	524.51	0.34	524.53	0.33	526.29
2	2044.97	0.38	2045.18	0.37	2052.71
3	4356.66	8.69	4357.01	8.70	4008.32
4	4486.69	0.95	4487.40	0.97	4444.36
5	7497.46	0.42	7499.76	0.39	7529.14
6	11082.29	0.57	11086.91	0.53	11146.30
7	13248.62	9.92	13249.33	9.92	12053.30
8	15058.46	0.60	15066.59	0.55	15150.10
9	19288.29	0.83	19300.72	0.76	19449.40
10	22072.68	9.47	22074.24	9.48	20163.40

Table 3. Natural frequency for a six layers $BaTiO_3/CoFe_3O_4$ simply-supported beam with different magneto-electric boundary conditions

5 Conclusion

A model for the dynamic analysis of general multilayer magneto-electro-elastic beams has been presented. First-order beam theory and quasi-static magneto-electric assumptions have been used. After the preliminary solution of the magneto-electric state as a function of the mechanical variables, the equations of motions have been deduced. They involve the same terms of the pure mechanical problem with effective magneto-electro-elastic stiffness coefficients that inherently take the magneto-electro-elastic coupling into account. Besides additional terms comes into play because of the piezoelectric and piezomagnetic behavior, which involve the third-order spatial derivative of the transverse displacement. The magneto-electric loads enter the dynamic model as equivalent external axial force and bending moment per unit length and via the axial force and bending moment definition, which also affects the beam ends mechanical boundary conditions.

The solution for the free vibrations problem has been obtained and the results obtained validate the approach and prove its effectiveness.

References

[1] Nan C.W., "Magnetoelectric effect in composites of piezoelectric and piezomagnetic phases", *Physical Review B*, Vol. 50, No. 9, 1994, pp. 6082-6088.

[2] Eerenstein W., Mathur N.D. and Scott J.F., "Multiferroic and magnetoelectric materials", *Nature*, Vol. 442, No. 7104, 2006, pp. 759-765.

[3] Fiebig M., "Revival of the magnetoelectric effect", *Journal of Physics D: Applied Physics*, Vol. 38, No. 8, 2005, R123.

[4] Fetisov Y., "Magnetoelectric effect in multilayer ferromagnetic-piezoelectric structures and its application in electronics", *Bulletin of the Russian Academy of Sciences: Physics*, Vol. 71, 2007, pp. 1626-1628.

[5] Zhai J., Xing Z., Dong S., Li J. and Viehland D., "Magnetoelectric Laminate Composites: An Overview" *Journal of the American Ceramic Society*, Vol. 91, No. 2, 2008, pp. 351-358.

[6] Pan E. and Heyliger P.R., "Exact solutions for magneto-electro-elastic laminates in cylindrical bending", *International Journal of Solids and Structures*, Vol. 40, No. 24, 2003, pp. 6859-6876.

[7] Pan E. and Heyliger P.R., "Free vibrations of simply supported and multilayered magneto-electro-elastic plates", *Journal of Sound and Vibration*, Vol. 252, No. 3, 2002, pp. 429-442.

[8] Wang J., Qu L. and Qian F., "State vector approach of free-vibration analysis of magneto-electro-elastic hybrid laminated plates", *Composite Structures*, Vol. 92, No. 6, 2010, pp. 1318-1324.

[9] Chen J., Chen H., Pan E. and Heyliger P.R., "Modal analysis of magneto-electro-elastic plates using the state-vector approach", *Journal of Sound and Vibration*, Vol. 304, No. 3-5, 2007, pp. 722-734.

[10] Benjeddou A, Trindade M.A. and Ohayon R., "Piezoelectric actuation mechanisms for intelligent sandwich structures", *Smart Materials and Structures*, Vol. 9, No 3, 2000, pp. 328-339.

[11] Simões Moita J.M, Mota Soares C.M. and Mota Soares C.A., "Analyses of magneto-electro-elastic plates using a higher order finite element

model", *Composite Structures*, Vol. 91, No. 4, 2009, pp. 421-426.
[12] Annigeri A.R., Ganesan N. and Swarnamani S., "Free vibration behaviour of multiphase and layered magneto-electro-elastic beam", *Journal of Sound and Vibration*, Vol. 299, No. 1-2, 2007, pp. 44-63.
[13] Milazzo A., Benedetti I. and Orlando C., "Boundary element method for magneto electro elastic laminates", *Computer Modeling in Engineering and Sciences*, Vol. 15, No 1, 2006, pp. 17-30.
[14] Alaimo A., Milazzo A. and Orlando C., "Dynamic analysis of damaged magnetoelectroelastic laminated structures", *Int. Conf. on Boundary Element Techniques XI*, EC Ltd UK, 2010, pp. 324-329.
[15] Davì G. and Milazzo A., "A regular variational boundary model for free vibrations of magneto-electro-elastic structures", *Engineering Analysis with Boundary Elements*, Vol. 35, No. 3, 2011, pp. 303-312.
[16] Milazzo A., Orlando C. and Alaimo A., "On the shear influence on the free vibration behavior of magneto-electro-elastic beam", *ICCES*, Vol. 11, No. 2, 2009, pp. 55-61.
[17] Wang C.M and Reddy J.N., *Shear deformable beams and plates*. Elsevier, 2000.
[18] Milazzo A., Orlando C. and Alaimo A., "An analytical solution for the magneto-electro-elastic bimorph beam forced vibrations problem", *Smart Materials and Structures*, Vol. 18, No 8, 2009, pp. 850-862.
[19] Mindlin R.D. and Goodman L.E., "Beam vibrations with time-dependent boundary conditions", *ASME Journal of Applied Mechanics*, Vol. 17, No. 4, 1950, pp. 377-380.

Appendix - Coefficient vectors

The closed-form expressions of the coefficient vectors involved in the electric and magnetic potential formulas are given for the different electro-magnetic boundary conditions considered. Without loss of generality, the bottom surface conditions are fixed as in Eqs. (7.a) and (7.b).

The electro-magnetic boundary conditions for the case of imposed electric displacements and magnetic induction at the top surface are written as

$$D_z(x, h_N, \tau) = \tilde{D}_z = \xi_3 \qquad (A.1)$$

$$B_z(x, h_N, \tau) = \tilde{B}_z = \xi_4 \qquad (A.2)$$

By using Eqs (6), (7.a), (7.b), (A.1) and (A.2), after some manipulations, the following expressions are obtained for the coefficients of Eqs. (15.a) and (15.b)

$$\boldsymbol{\alpha}_u^{\langle k\rangle} = [\boldsymbol{SW}]^{\langle k\rangle} \qquad (A.3)$$

$$\boldsymbol{\alpha}_\vartheta^{\langle k\rangle} = \boldsymbol{S}^{\langle k\rangle}\left\{\boldsymbol{W}_{14}^{\langle k\rangle} h_k + \sum_{i=k+1}^{N} \boldsymbol{W}_{14}^{\langle i\rangle} t_i\right\} \qquad (A.4)$$

$$\boldsymbol{\alpha}_w^{\langle k\rangle} = -\boldsymbol{S}^{\langle k\rangle}\left\{\boldsymbol{W}_{14}^{\langle k\rangle} h_k + \sum_{i=k+1}^{N} \boldsymbol{W}_{14}^{\langle i\rangle} t_i\right\} \qquad (A.5)$$

$$\boldsymbol{\alpha}_{\xi i}^{\langle k\rangle} = \begin{cases} \mathbf{0} & i = 1,2 \\ -[\![\boldsymbol{S}^{\langle k\rangle}]\!]_{i-2} & i = 3,4 \end{cases} \qquad (A.6)$$

$$\boldsymbol{\beta}_u^{\langle k\rangle} = -\boldsymbol{\alpha}_u^{\langle k\rangle} h_{k-1} + \sum_{i=1}^{k-1} \boldsymbol{\alpha}_u^{\langle i\rangle}\, t_i \qquad (A.7)$$

$$\boldsymbol{\beta}_\vartheta^{\langle k\rangle} = [\boldsymbol{S}(\boldsymbol{W} + \boldsymbol{W}_{14})]^{\langle k\rangle} \frac{{h_{k-1}}^2}{2} - \boldsymbol{\alpha}_\vartheta^{\langle k\rangle} h_{k-1} + \sum_{i=1}^{k-1} \left\{\boldsymbol{\alpha}_\vartheta^{\langle i\rangle}\, t_i - [\boldsymbol{S}(\boldsymbol{W} + \boldsymbol{W}_{14})]^{\langle i\rangle} \bar{t}_i\right\} \qquad (A.8)$$

$$\boldsymbol{\beta}_w^{\langle k\rangle} = -\boldsymbol{\alpha}_w^{\langle k\rangle} h_{k-1} - [\boldsymbol{SW}_{14}]^{\langle k\rangle} \frac{{h_{k-1}}^2}{2} + \sum_{i=1}^{k-1} \left(\boldsymbol{\alpha}_w^{\langle i\rangle}\, t_i + [\boldsymbol{SW}_{14}]^{\langle i\rangle} \bar{t}_i\right) \qquad (A.9)$$

$$\boldsymbol{\beta}_{\xi i}^{\langle k\rangle} = \begin{cases} [\![\mathbf{I}]\!]_i - \boldsymbol{\alpha}_{\xi i}^{\langle k\rangle} h_{k-1} + \sum_{j=1}^{k-1} \boldsymbol{\alpha}_{\xi i}^{\langle j\rangle}\, t_j & i = 1,2 \\ \mathbf{0} & i = 3,4 \end{cases} \qquad (A.10)$$

where $t_i = h_i - h_{i-1}$, $\bar{t}_i = (h_i^2 - h_{i-1}^2)/2$ and the notation $[\![\cdot]\!]_j$ means the j-th column of a matrix.

The electro-magnetic boundary conditions for the case of imposed electric and magnetic potentials at the top surface are written as

$$\Phi(x, h_N, \tau) = \tilde{\Phi} = \xi_3 \qquad (A.11)$$

$$\Psi(x, h_N, \tau) = \tilde{\Psi} = \xi_4 \qquad (A.12)$$

By using Eqs (6), (7.a), (7.b), (A.11) and (A.12), the coefficients appearing in the Eqs. (15.a) and (15.b) are given by

$$\boldsymbol{\alpha}_{u}^{\langle k\rangle} = \boldsymbol{S}^{\langle k\rangle}\left[\sum_{j=2}^{k}|\boldsymbol{W}|_{\langle j-1\rangle}^{\langle j\rangle} - \mathbf{Z}\sum_{i=1}^{N}\left(\boldsymbol{S}^{\langle i\rangle}t_i\sum_{j=2}^{i}|\boldsymbol{W}|_{\langle j-1\rangle}^{\langle j\rangle}\right)\right] \tag{A.13}$$

$$\boldsymbol{\alpha}_{\vartheta}^{\langle k\rangle} = \boldsymbol{S}^{\langle k\rangle}\left\{\sum_{j=2}^{k}|\boldsymbol{W}|_{\langle j-1\rangle}^{\langle j\rangle}h_{j-1} + \mathbf{Z}\sum_{i=1}^{N}\boldsymbol{S}^{\langle i\rangle}\left[(\boldsymbol{W}+\boldsymbol{W}_{14})^{\langle i\rangle}\bar{t}_i - t_i\sum_{j=2}^{i}|\boldsymbol{W}_{14}|_{\langle j-1\rangle}^{\langle j\rangle}h_{j-1}\right]\right\} \tag{A.14}$$

$$\boldsymbol{\alpha}_{w}^{\langle k\rangle} = -\boldsymbol{S}^{\langle k\rangle}\left\{\sum_{j=2}^{k}|\boldsymbol{W}_{14}|_{\langle j-1\rangle}^{\langle j\rangle}h_{j-1} + \mathbf{Z}\sum_{i=1}^{N}\boldsymbol{S}^{\langle i\rangle}\left[\boldsymbol{W}_{14}{}^{\langle i\rangle}\bar{t}_i - t_i\sum_{j=2}^{i}|\boldsymbol{W}_{14}|_{\langle j-1\rangle}^{\langle j\rangle}h_{j-1}\right]\right\} \tag{A.15}$$

$$\boldsymbol{\alpha}_{\xi i}^{\langle k\rangle} = \begin{cases} -[\![\boldsymbol{S}^{\langle k\rangle}\mathbf{Z}]\!]_i & i=1,2 \\ [\![\boldsymbol{S}^{\langle k\rangle}\mathbf{Z}]\!]_{i-2} & i=3,4 \end{cases} \tag{A.16}$$

$$\boldsymbol{\beta}_{u}^{\langle k\rangle} = -\boldsymbol{\alpha}_{u}^{\langle k\rangle}h_k - \sum_{i=1}^{k-1}\boldsymbol{\alpha}_{u}^{\langle i\rangle}\,t_i \tag{A.17}$$

$$\boldsymbol{\beta}_{\vartheta}^{\langle k\rangle} = [\boldsymbol{S}(\boldsymbol{W}+\boldsymbol{W}_{14})]^{\langle k\rangle}\frac{h_k{}^2}{2} - \boldsymbol{\alpha}_{\vartheta}^{\langle k\rangle}h_k + \sum_{i=1}^{k-1}\left\{[\boldsymbol{S}(\boldsymbol{W}+\boldsymbol{W}_{14})]^{\langle i\rangle}\bar{t}_i - \boldsymbol{\alpha}_{\vartheta}^{\langle i\rangle}\,t_i\right\} \tag{A.18}$$

$$\boldsymbol{\beta}_{w}^{\langle k\rangle} = -[\boldsymbol{S}\boldsymbol{W}_{14}]^{\langle k\rangle}\frac{h_k{}^2}{2} - \boldsymbol{\alpha}_{w}^{\langle k\rangle}h_k + \sum_{i=1}^{k-1}\left(\boldsymbol{\alpha}_{w}^{\langle i\rangle}\,t_i + [\boldsymbol{S}\boldsymbol{W}_{14}]^{\langle i\rangle}\bar{t}_i\right) \tag{A.19}$$

$$\boldsymbol{\beta}_{\xi i}^{\langle k\rangle} = \begin{cases} \boldsymbol{\alpha}_{\xi i}^{\langle k\rangle}h_k + \sum_{j=1}^{k-1}\boldsymbol{\alpha}_{\xi i}^{\langle j\rangle}\,t_j & i=1,2 \\ [\![\mathbf{I}]\!]_{i-2} - \boldsymbol{\alpha}_{\xi i}^{\langle k\rangle}h_k - \sum_{j=1}^{k-1}\boldsymbol{\alpha}_{\xi i}^{\langle j\rangle}\,t_j & i=3,4 \end{cases} \tag{A.20}$$

where

$$\boldsymbol{Z} = \left(\sum_{i=1}^{N}-\boldsymbol{S}^{\langle i\rangle}\,t_i\right)^{-1} \tag{A.21}$$

3rd CEAS Air&Space Conference
21st AIDAA Congress

Dynamic Analysis of Magneto-Electro-Elastic Structures by a Meshless Approach

G. Davì and A. Milazzo
Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale, Università di Palermo, Viale delle Scienze, 90128 Palermo - Italy

Keywords: *Magneto-electro-elasticity, meshless methods, multifield dynamics*

Abstract

The present paper deals with a hybrid variational formulation for magneto-electro-elastic dynamic problems. Starting from a generalized hybrid functional, a suitable discretization scheme leads to a model involving generalized displacements only, produces frequency independent structural matrices and preserves their symmetry. The generalized stiffness and mass matrices are computed by boundary integrations of regular kernels only with the consequent computational advantages. The method is placed in the framework of the weak form meshless methods. Indeed, only a set of scattered points is needed for the variable interpolation, while a global background boundary mesh is merely used for the integration process. The model can be straightforwardly used for multilayer configurations. Numerical results for a damaged magneto-electro-elastic laminate are presented, showing the effectiveness of the method.

1 Introduction

Magneto-electro-elastic materials have recently emerged for application in sensors, actuators, transducers and in smart structures. They are able to convert energy into three different forms: magnetic, electric and mechanic. This feature seems to make these multifield composites potentially superior to other materials for the application in smart structures technology [1], where layers of magneto-electro-elastic material can be either attached to or embedded in the host structure to be controlled and actuated by a system of suitably arranged electrodes and poles. The effective design of magneto-electro-elastic devices requires the capability of correctly and reliably modelling the system's response taking the coupling effects into account, especially in dynamics which is the most important field of application (e.g. free vibration control).

Analytical solutions to single-layer and laminated magneto-electro-elastic solids are rather rare due to the complexity of the governing equations. More general problems are usually solved through numerical methods and the finite element (FEM) represents the widely employed technique [2-12].

Recently, also the boundary element method (BEM) has been extended to the study of piezoelectric [13-15] and magneto-electro-elastic [16-19] media. When dealing with anisotropic multifield dynamics by the

BEM, the lack of closed-form fundamental solutions leads to rather expensive numerical integration schemes. This drawback is often overcome by using the static fundamental solutions. By so doing the inertia terms are considered as body forces and the associated domain integrals are transformed into boundary integrals generally by the Dual Reciprocity Method [20]. Additionally, in the standard dynamic BEM the discrete structural operators loss their symmetry and definiteness properties resulting in less efficient models for dynamics. Concerning this issue, variational BEM formulations have been developed [21-24]. In this framework the authors originally proposed a hybrid variational formulation of BEM, called Displacement Boundary Method (DBM), devised to preserve the fundamental properties of the structural matrices [25-27]. As pointed out in reference [28], this approach presents some features which allow classifying it as a weak form meshless method [29].

In the present paper the DBM is applied to the magneto-electro-elastic problem [30], with the aim to investigate its accuracy and effectiveness for the computation of the dynamical response.

2 Governing equations

Let Ω denote a two-dimensional magneto-electro-elastic body lying in the x_1x_2 plane. The body is bounded by the contour line $\partial\Omega$. It is assumed that the magneto-electro-mechanical response does not vary along the x_3 direction and that the material behaves linearly.

By using suitable generalized variables a compact matrix notation is achieved for the multifield problem considered [30]. So, let us introduce the following generalized quantities for a two-dimensional problem:

– the generalized displacements $\boldsymbol{U}$, obtained collecting the displacement components u_i, the electric potential φ and the scalar magnetic potential ψ

$$\boldsymbol{U} = [u_1 \;\; u_2 \;\; \varphi \;\; \psi]^T \qquad (1.a)$$

– the generalized strain vector $\boldsymbol{\Gamma}$, containing the strain components γ_{ij}, the electric field components E_i and the magnetic field components H_i

$$\boldsymbol{\Gamma} = [\gamma_{11} \; \gamma_{22} \; \gamma_{12} \; -E_1 \; -E_2 \; -H_1 \; -H_2]^T \qquad (1.b)$$

– the generalized stress vector $\boldsymbol{\Sigma}$, containing the stresses σ_{ij}, the electric displacements D_i and the magnetic induction components B_i

$$\boldsymbol{\Sigma} = [\sigma_{11} \; \sigma_{22} \; \sigma_{12} \; D_1 \; D_2 \; B_1 \; B_2]^T \qquad (1.c)$$

– the generalized tractions vector $\boldsymbol{T}$, containing the traction components t_i, the normal electric displacement D_n and the normal magnetic induction B_n

$$\boldsymbol{T} = [t_1 \;\; t_2 \;\; D_n \;\; B_n]^T \qquad (1.d)$$

– the generalized body forces vector $\boldsymbol{F}$, containing the components of the mechanical body forces f_i , the electric charge density ω and the identically zero free current

$$\boldsymbol{F} = [f_1 \;\; f_2 \;\; -\omega \;\; 0]^T \qquad (1.e)$$

The proposed model is based on the extension to magneto-electro-elasticity of a hybrid displacement variational approach just proposed in the literature [25-28, 30], which is here briefly outlined for the considered problem. Let $\boldsymbol{U}$ be the generalized displacement field in the two-dimensional domain Ω, and let $\widetilde{\boldsymbol{U}}$ and $\widetilde{\boldsymbol{T}}$ be the generalized displacements and tractions on the boundary $\partial\Omega$. These variables are assumed to be independent from each other. The variational formulation of the magneto-electro-elastic problem is obtained by introducing the hybrid functional

$$\Pi = \int_\Omega \frac{1}{2}\boldsymbol{\Gamma}^T\boldsymbol{R}\boldsymbol{\Gamma}d\Omega - \int_\Omega \boldsymbol{U}^T\left(\boldsymbol{F} - \boldsymbol{\rho}\ddot{\boldsymbol{U}}\right)d\Omega - \int_{\partial\Omega}\left(\boldsymbol{U} - \widetilde{\boldsymbol{U}}\right)^T\widetilde{\boldsymbol{T}}d\partial\Omega - \int_{\partial\Omega_2}\widetilde{\boldsymbol{U}}^T\overline{\boldsymbol{T}}d\partial\Omega \qquad (2)$$

where $\boldsymbol{R}$ is the generalized stiffness matrix and $\boldsymbol{\rho}$ denotes the product of the material density by the 4×4 identity matrix in which the last two diagonal terms are replaced by zeros as quasi-

static modeling of electric and magnetic fields is assumed. In the Eq. (2) the overdot indicates the time derivative and the overline denotes prescribed quantities on the boundaries $\partial\Omega_1$ and $\partial\Omega_2$, where natural and essential boundary conditions are prescribed.

Assuming that linear strain-displacements relations hold and that electric and magnetic fields obey the gradient equations with respect to the corresponding potentials, after integration by parts, the stationarity of Π with respect to the independent fields $\boldsymbol{U}$, $\widetilde{\boldsymbol{U}}$ and $\widetilde{\boldsymbol{T}}$ provides the following set of Euler equations

$$\boldsymbol{\mathcal{D}}^T \boldsymbol{R}\boldsymbol{\mathcal{D}}\boldsymbol{U} - \boldsymbol{\rho}\ddot{\boldsymbol{U}} + \boldsymbol{F} = \boldsymbol{0} \quad on\ \Omega \tag{3.a}$$

$$\boldsymbol{U} = \widetilde{\boldsymbol{U}} \quad on\ \partial\Omega \tag{3.b}$$

$$\boldsymbol{T} = \widetilde{\boldsymbol{T}} \quad on\ \partial\Omega \tag{3.c}$$

$$\widetilde{\boldsymbol{T}} = \overline{\boldsymbol{T}} \quad on\ \partial\Omega_2 \tag{3.d}$$

where $\boldsymbol{\mathcal{D}}$ is the generalized equilibrium operator [30]

$$\boldsymbol{\mathcal{D}} = \begin{bmatrix} \partial/\partial x_1 & 0 & 0 & 0 \\ 0 & \partial/\partial x_2 & 0 & 0 \\ \partial/\partial x_2 & \partial/\partial x_1 & 0 & 0 \\ 0 & 0 & \partial/\partial x_1 & 0 \\ 0 & 0 & \partial/\partial x_2 & 0 \\ 0 & 0 & 0 & \partial/\partial x_1 \\ 0 & 0 & 0 & \partial/\partial x_2 \end{bmatrix} \tag{4}$$

Thus, by assuming that the generalized compatibility, the constitutive equations, the essential boundary conditions, and the initial conditions are satisfied, the stationarity of the generalized functional Π leads to the solution of the magneto-electro-elastic dynamic problem.

3 Discrete dynamic model

Let us consider a set of N randomly chosen nodes a part of which, namely N_B, lies on the boundary $\partial\Omega$, whereas the others, namely N_D, belong to the domain $\boldsymbol{\Omega}$ (see Fig. 1).

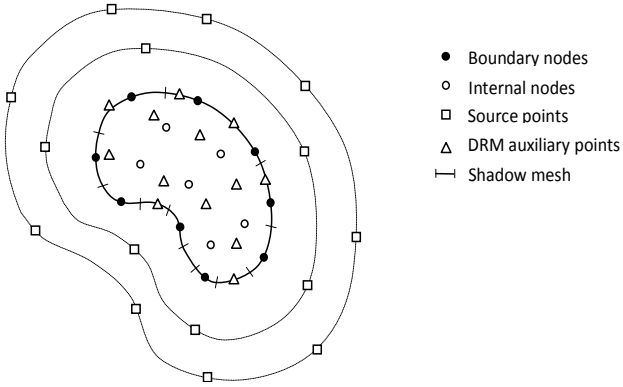


Fig. 1 Discretization scheme

3.1 Variable approximation

The generalized domain displacements $\boldsymbol{U}$ are approximated as a superposition of spatial trial functions $\boldsymbol{U}_k^*(P)$ weighted by time dependent coefficients $s_k(t)$ and one has

$$\boldsymbol{U} \cong \sum_{k=1}^{4N} \boldsymbol{U}_k^*(P)\, s_k(\tau) = \boldsymbol{U}^* \boldsymbol{s} \quad on\ \Omega \tag{5}$$

where P denotes the field point. In the present approach, the employed trial functions are the fundamental solutions of the static magneto-electro-elastic problem. They are the static magneto-electro-elastic response of the unbounded medium corresponding to generalized concentrated body forces applied at the point P_0, namely the source point. For each source point, four independent fundamental solutions are available: two related to the unit mechanical point loads acting along the reference axes, one to the unit point free charge density and one to the unit point current density. The expressions of these fundamental solutions can be found in Ref. [16] or [18]. The employed trial functions are defined by associating a source point with each nodal point and locating these source points outside the domain so as to deal with regular trial functions [22, 23, 25].

The boundary generalized variables are expressed in terms of their nodal values through spatial shape functions $\boldsymbol{\mathcal{N}}$ and $\boldsymbol{\mathcal{S}}$

$$\widetilde{\boldsymbol{U}} = \boldsymbol{\mathcal{N}}\boldsymbol{\Delta} = [\boldsymbol{\mathcal{N}}_1 \ \ \boldsymbol{\mathcal{N}}_2] \begin{bmatrix} \boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_2 \end{bmatrix} \tag{6}$$

$$\widetilde{\boldsymbol{T}} = \boldsymbol{S}\boldsymbol{P} \tag{7}$$

where $\boldsymbol{\Delta}$ and $\mathbf{P}$ are the nodal generalized displacements and tractions vector, respectively. The subscripts 1 and 2 refer to constrained and free generalized nodal displacements, respectively.

3.2 Generalized equations of motion

By using the introduced approximations, *i.e.* Eqs. (5), (6) and (7), the discretized expression of the generalized hybrid functional Π is obtained. Taking the discretized functional first variation with respect to the discrete variables $\boldsymbol{s}$, $\boldsymbol{\Delta}$ and $\mathbf{P}$ and applying the fundamental lemma of the calculus of variations, one obtains the following set of matrix equations

$$\left(\int_{\Omega} \boldsymbol{U}^{*T}\boldsymbol{\rho}\boldsymbol{U}^{*}d\Omega\right)\ddot{\boldsymbol{s}} + \left(\int_{\partial\Omega} \boldsymbol{U}^{*T}\boldsymbol{T}^{*}d\partial\Omega + \int_{\Omega} \boldsymbol{U}^{*T}\boldsymbol{Q}^{*}d\Omega\right)\boldsymbol{s} = \left(\int_{\partial\Omega} \boldsymbol{U}^{*T}\boldsymbol{S}d\partial\Omega\right)\boldsymbol{P} + \int_{\Omega} \boldsymbol{U}^{*T}\boldsymbol{F}d\Omega \tag{8.a}$$

$$\int_{\partial\Omega} \boldsymbol{\mathcal{N}}_2^{T}\boldsymbol{S}d\partial\Omega\,\boldsymbol{P} = \int_{\partial\Omega_2} \boldsymbol{\mathcal{N}}_2^{T}\overline{\boldsymbol{T}}d\partial\Omega \tag{8.b}$$

$$\int_{\partial\Omega} \boldsymbol{S}^{T}\boldsymbol{U}^{*}d\partial\Omega\,\boldsymbol{s} = \int_{\partial\Omega} \boldsymbol{S}^{T}\boldsymbol{\mathcal{N}}d\partial\Omega\cdot\boldsymbol{\Delta} \tag{8.c}$$

where $\boldsymbol{T}^{*}$ and $\boldsymbol{Q}^{*}$ are the matrices of the generalized boundary tractions and generalized domain loads associated with the trial functions collected in the matrix $\boldsymbol{U}^{*}$.

Due to the arbitrary choice of the shape functions $\boldsymbol{S}$, from Eq. (8.c) it results

$$\boldsymbol{U}^{*}\boldsymbol{s} = \boldsymbol{\mathcal{N}}\boldsymbol{\Delta} \quad on \quad \partial\Omega \tag{9}$$

By using Eqs. (5) and (9) a relation between the coefficient vector $\boldsymbol{s}$ and the generalized displacements $\boldsymbol{\Delta}$ is established. Actually, collocating Eq. (9) at the N_B boundary nodal points and Eq. (5) at the N_D domain nodes, one obtains

$$\begin{bmatrix}\overline{\boldsymbol{U}}_{\partial\Omega}^{*}\\ \overline{\boldsymbol{U}}_{\Omega}^{*}\end{bmatrix}\boldsymbol{s} = \boldsymbol{\Lambda}\,\boldsymbol{s} = \begin{bmatrix}\boldsymbol{\Delta}_{\partial\Omega}\\ \boldsymbol{\Delta}_{\Omega}\end{bmatrix} \tag{10}$$

where the elements of the collocation matrices $\overline{\boldsymbol{U}}_{\partial\Omega}^{*}$ and $\overline{\boldsymbol{U}}_{\Omega}^{*}$ are the values of $\boldsymbol{U}^{*}$ computed at the boundary and internal nodes, respectively. In Eq. (10), $\boldsymbol{\Delta}_{\partial\Omega}$ and $\boldsymbol{\Delta}_{\Omega}$ are the vectors collecting the generalized displacements of the boundary and domain nodes, respectively. From Eq. (10), one writes

$$\boldsymbol{s} = \boldsymbol{\Phi}\begin{bmatrix}\boldsymbol{\Delta}_{\partial\Omega}\\ \boldsymbol{\Delta}_{\Omega}\end{bmatrix} = [\boldsymbol{\Phi}_{\partial\Omega}\ \ \boldsymbol{\Phi}_{\Omega}]\begin{bmatrix}\boldsymbol{\Delta}_{\partial\Omega}\\ \boldsymbol{\Delta}_{\Omega}\end{bmatrix} = [\boldsymbol{\Phi}_1\ \ \boldsymbol{\Phi}_2]\begin{bmatrix}\boldsymbol{\Delta}_1\\ \boldsymbol{\Delta}_2\end{bmatrix} \tag{11}$$

where $\boldsymbol{\Phi} = \boldsymbol{\Lambda}^{-1}$. It is worth noting that Eqs. (9) and (11) imply that

$$\boldsymbol{\mathcal{N}} = \boldsymbol{U}^{*}\boldsymbol{\Phi} = \boldsymbol{U}^{*}[\boldsymbol{\Phi}_1\ \ \boldsymbol{\Phi}_2] = [\boldsymbol{\mathcal{N}}_1\ \ \boldsymbol{\mathcal{N}}_2] \tag{12}$$

Substituting Eq. (11) into Eq. (8.a), pre-multiplying it by $\boldsymbol{\Phi}_2^{T}$ and taking Eq. (8.b) into account, one obtains the generalized equations of motion for the magneto-electro-elastic solid

$$\boldsymbol{M}\ddot{\boldsymbol{\Delta}}_2 + \boldsymbol{K}\boldsymbol{\Delta}_2 = \int_{\Omega} \boldsymbol{\mathcal{N}}_2^{T}\boldsymbol{F}d\Omega + \int_{\partial\Omega_2} \boldsymbol{\mathcal{N}}_2^{T}\overline{\boldsymbol{T}}\,d\partial\Omega - \boldsymbol{\Phi}_2^{T}\overline{\boldsymbol{M}}\boldsymbol{\Phi}_1\ddot{\boldsymbol{\Delta}}_1 - \boldsymbol{\Phi}_2^{T}\overline{\boldsymbol{K}}\boldsymbol{\Phi}_1\boldsymbol{\Delta}_1 \tag{13}$$

where the generalized mass and stiffness matrices are defined by

$$\boldsymbol{M} = \boldsymbol{\Phi}_2^{T}\overline{\boldsymbol{M}}\boldsymbol{\Phi}_2 = \boldsymbol{\Phi}_2^{T}\left(\int_{\Omega} \boldsymbol{U}^{*T}\boldsymbol{\rho}\boldsymbol{U}^{*}d\Omega\right)\boldsymbol{\Phi}_2 \tag{14.a}$$

$$\boldsymbol{K} = \boldsymbol{\Phi}_2^{T}\overline{\boldsymbol{K}}\boldsymbol{\Phi}_2 = \boldsymbol{\Phi}_2^{T}\left(\int_{\partial\Omega} \boldsymbol{U}^{*T}\boldsymbol{T}^{*}d\partial\Omega + \int_{\Omega} \boldsymbol{U}^{*T}\boldsymbol{Q}^{*}d\Omega\right)\boldsymbol{\Phi}_2 \tag{14.b}$$

The matrix operators involved in the model, $\boldsymbol{M}$ and $\boldsymbol{K}$, are symmetric and frequency independent.

3.3 Generalized operators computation

Due to the use of source points external to the domain Ω, the generalized stiffness matrix $\boldsymbol{K}$ is computed by integrals performed on the boundary only, because the domain integrals appearing in Eq. (14.b) are trivially zero.

The generalized mass matrix $\boldsymbol{M}$ is defined in terms of domain integrals and then its numerical computation hypothetically calls for a background mesh in the domain. However the domain integrals defining the generalized mass matrix can be approximated to the boundary by using the Dual Reciprocity Method (DRM) [20].

Let us introduce a set of $M = M_B + M_D$ auxiliary points, M_B lying on the boundary and M_D belonging to the domain (see Fig. 1). Associating with each of these nodes four time independent auxiliary spatial functions $\boldsymbol{g}_k(P)$, the expression of the domain inertia term can be approximated as

$$\boldsymbol{\rho \ddot{U}} = \boldsymbol{\rho U^* \ddot{s}} \cong \sum_{k=1}^{4M} \boldsymbol{g}_k^*(P)\alpha_k(\tau) = \boldsymbol{G^* \alpha} \qquad (15)$$

where the unknown coefficients $\alpha_k(\tau)$ depend on time. Collocating Eq. (15) at the auxiliary points, a linear system linking the coefficients $\boldsymbol{\alpha}$ and $\ddot{\mathbf{s}}$ is obtained whose solution provides $\boldsymbol{\alpha} = \boldsymbol{\Xi}\,\ddot{\mathbf{s}}$ [30]. Thus, the inertia term is approximated by

$$\bar{\boldsymbol{M}} \cong \left(\int_\Omega \boldsymbol{U^{*T} G^*} d\Omega \right) \boldsymbol{\Xi \ddot{s}} \qquad (16)$$

For the integral term of the Eq. (16), the following reciprocity statement holds

$$\int_\Omega \boldsymbol{U^{*T} G^*} d\Omega = \int_\Gamma \left[\boldsymbol{T^{*T} V^*} - \boldsymbol{U^{*T} W^*} \right] d\Gamma \qquad (17)$$

where $\boldsymbol{V^*}$ and $\boldsymbol{W^*}$ are matrices whose k-th columns $\boldsymbol{v}_k^*$ and $\boldsymbol{w}_k^*$ are the generalized displacements and tractions due to the dummy generalized body forces $\boldsymbol{g}_k^*$. Consequently, the generalized mass matrix can be approximated by

$$\boldsymbol{M} = \boldsymbol{\Phi}_2^T \left(\int_\Gamma \left[\boldsymbol{T^{*T} V^*} - \boldsymbol{U^{*T} W^*} \right] d\Gamma \right) \boldsymbol{\Xi}\, \boldsymbol{\Phi}_2 \qquad (18)$$

where only boundary integrations are involved.

Due to the difficulty of determining the generalized displacement field associated with the dummy generalized body forces also when simple functions $\boldsymbol{g}_k^*$ are used, an alternative scheme for DRM provides that the generalized auxiliary displacements $\boldsymbol{v}_k^*$ are chosen *a priori* and then the associated generalized tractions $\boldsymbol{w}_k^*$ and body forces $\boldsymbol{g}_k^*$ are consequently derived. Actually, several authors concerned with the dual reciprocity BEM for anisotropic elasticity or piezoelectricity have employed this approach [31-33].

3.4 Discussion

The structure of the proposed model suggests that the approach can be thought as a meshfree model. Actually, the model involves only the displacements of a set of scattered nodal points introduced by the functional dicretization process of the functional Π. The influence matrices, i.e. the stiffness and mass matrices, are defined in terms of the trial functions and their construction does not call for assembly procedures based on a mesh definition.

The matrix operators can be evaluated by boundary integrations of regular functions with meaningful computational advantages. Their computation requires a background boundary mesh (*shadow mesh*), which does not introduce any degree of freedom and is used only as support for the integration process. For the generalized mass matrix the goal of its computation through boundary integrals is achieved in force of the approximation defined in Eq. (15). This allows the application of the DRM technique that transforms the mass matrix to the boundary, but generally destroys the symmetry properties. As a consequence, spurious complex eigenvalues could be found, which are not meaningful of the modal content of the dynamic response. Moreover, the mass matrix transformation may introduce

inaccuracies in the higher modes. However, these unwanted effects reduce with the approximation accuracy [30]. In conclusion, the generalized mass matrix transformation allows its computation without introducing a background domain mesh and performing the required integrations on the boundary background mesh defined for the calculation of the stiffness matrix. These computational advantages can be achieved tolerating the loss of symmetry in the generalized mass matrix.

It should be also noted that the proposed model evidences the characteristics of a displacement based finite element model. In this sense, one can assemble different domains by using standard finite elements assembly procedures. This approach allows dealing with multilayered structures.

4 Numerical results

For the application of the proposed method, a computer code with the following fundamental features has been implemented.

The set of boundary and internal nodes and the set of source points are chosen for the discretization process. For each discretization node a source point is introduced and located on lines emanating from the nodes and normal to the boundary. The source points are placed at a distance from the corresponding node given by r times the length of the element of the background boundary mesh nearest to the considered node [23, 25]. The transformation of the mass matrix is performed by setting as auxiliary displacement components cubic radial basis functions depending on the distance between the field point and the considered source point [30]. The background boundary mesh uses standard one-dimensional isoparametric elements over which Gaussian quadrature is performed to compute the stiffness and mass matrices. For the free vibration analysis standard condensation techniques is applied and standard eigenvalues routines are employed to compute the natural frequencies.

An extensive analysis of the influence of the principal parameters involved in the proposed method on its convergence and accuracy is given in Ref. [30] where the free vibration problem has been addressed.

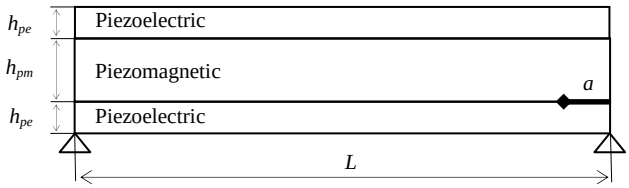


Fig. 1 Magneto-electro-elastic laminate geometry

Here the attention is focused on the computation of the time response in the framework of damaged multilayer structures. To this aim the different layers are connected along the interfaces through generalized spring elements, which are used to model the adhesive layers. Cracks are modelled as interfaces with infinite compliance, numerically accomplished through suitable penalty function. To integrate the generalized equations of motion the Houbolt method is employed. To avoid interface surfaces overlapping, an iterative procedure has been implemented for the solution at each time step [34].

A magneto-electro-elastic laminate undergoing a through the thickness magnetic field suddenly applied at time $\tau = 0$ is analyzed. The laminate consists of a piezomagnetic $CoFe_2O_4$ layer, having thickness $h_{pm} = 4\,mm$, embedded between two piezoelectric layers of $BaTiO_3$, having thickness $h_{pe} = 2\,mm$. An edge delamination crack of length $a = 2\,mm$ is considered between the piezoelectric and piezomagnetic layers as shown in Fig. 2. The laminate is simply-supported; its length is $L = 20\,mm$ while the materials properties are taken from Ref. [12]. Perfect bonding between the layers has been considered. The analyses have been performed by using a discretization scheme with $N_B = 48$, $N_D = 8$ and $M = 56$ for each layer.

Since the crack lies at a bimaterial interface the dynamic fracture behavior is characterized in terms of dynamic energy release rate G, normalized by the static energy release rate G_0, and in term of the modes phase angle Ψ to highlight the predominant fracture mode. It is worth noting that the energy release rate depends on the square of the stress intensity factor so that no compenetration occurrence can

be inferred from the time history of G. Thus, both the a linear analysis and a nonlinear analysis with contact are carried out to take the influence of the forbidden crack surface overlap on the dynamic fracture characteristic extimation into account.

The computed dynamic fracture mechanics parameters are given in Figs. 3 and 4 as function of the non-dimensional time τ/τ_0, where $\tau_0 = 0.286\,s$ is the first fundamental period of the structure. They have been evaluated by classical crack closure techniques, which are widely employed in finite element models. The comparison with the results obtained through a classical boundary element approach [35] shows a good agreement, proving that the present method is effective also to evaluate the time history of complex structures.

5 Conclusions

A meshless method for the dynamic analysis of two-dimensional magneto-electro-elastic solids has been presented. The proposed model is

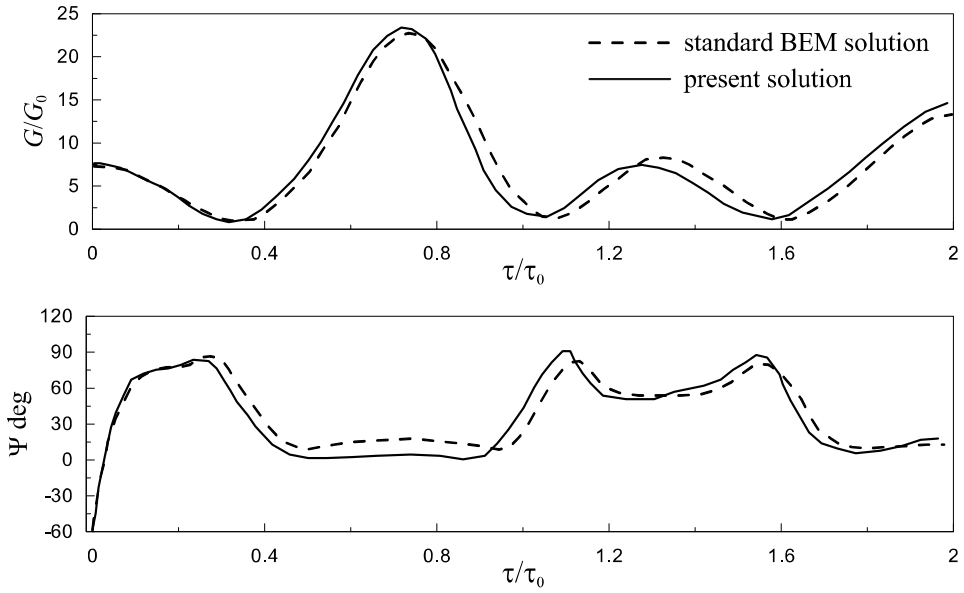


Fig. 3 Energy release rate and fracture modes phase angle: linear analys without contact

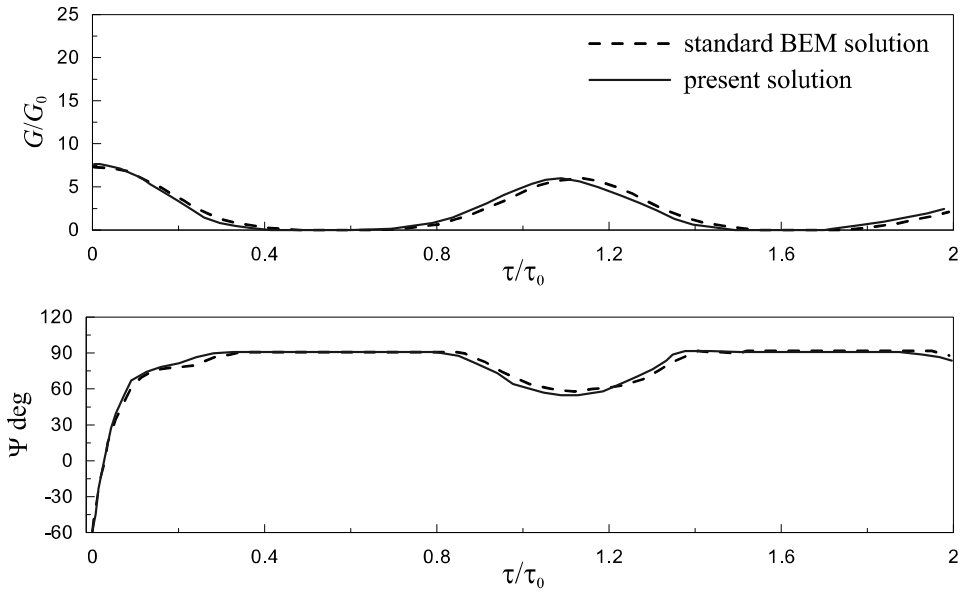


Fig. 4 Energy release rate and fracture modes phase angle: nonlinear contact analys

based on the discretization of a generalized hybrid functional by using as trial functions a set of magneto-electro-elastic non singular static fundamental solutions. This leads to frequency-independent structural matrices whose definiteness and symmetry properties are preserved. A boundary-only model is obtained by transforming domain integrals to the boundary through the dual reciprocity method. This implies the loss of symmetry in the inertial operator, but leads to meaningful computational advantages. The analyses performed for damaged multilayer laminates show that the method is effective to determine the structural time-histories of complex structures.

References

[1] Ryu J., Priya S., Uchinom K. and Kim H.E., "Magnetoelectric Effect in Composites of Magnetostrictive and Piezoelectric Materials", *Journal of Electroceramycs*, Vol. 8, No. 2, pp. 107-119.

[2] Saravanos D.A. and Heyliger P.R., "Mechanics and computational models for laminated piezoelectric beams, plates and shells", *Applied Mechanics Review*, Vol. 52, No. 10, 1999, pp. 305-320.

[3] Fernandes A. and Pouget J., "Analytical and numerical approaches to piezoelectric bimorph", *International Journal of Solids and Structures*, Vol. 40, No. 17, 2003, pp. 4331-4352.

[4] Vel S.S., Mewer R.C. and Batra R.C., "Analytical solution for the cylindrical bending vibration of piezoelectric composite plates", *International Journal of Solids and Structures*, Vol. 41, No. 5-6, 2004, pp. 1625-1643.

[5] Ballhause D., D'Ottavio M., Kröplin B. and Carrera E., "A unified formulation to assess multilayered theories for piezoelectric plates", *Computers and Structures*, Vol. 83, No. 15-16, 2005, pp.1217-1235.

[6] Pan E. and Heyliger P.R., "Free vibrations of simply supported and multilayered magneto-electro-elastic plates", *Journal of Sound and Vibrations*, Vol. 252, No. 3, 2002, pp. 429-442.

[7] Buchanan G.R., "Free vibration of an infinite magneto-electro-elastic cylinder", *Journal of Sound and Vibrations*, Vol. 268, No. 2, 2003, pp. 413-426.

[8] Lage R.G., Mota Soares C.M., Mota Soares C.A. and Reddy J.N., "Layerwise partial mixed finite element analysis of magneto-electro-elastic plates". *Computers and Structures*, Vol. 82, No. 17-19, 2004, pp. 1293-1301.

[9] Buchanan G.R., "Layered versus multiphase magneto-electro-elastic composites", *Composites Part B-Engineering*, Vol. 35, No. 5, 2004, pp. 413-420.

[10] Ramirez F., Heyliger P.R. and Pan E., "Free vibration response of two-dimensional magneto-electro-elastic laminated plates". *Journal of Sound and Vibrations*, Vol. 292, No. 3-5, 2006, pp. 626-644.

[11] Chen J., Chen H., Pan E. and Heyliger P.R., "Modal analysis of magneto-electro-elastic plates using the state-vector approach", *Journal of Sound and Vibrations*, Vol. 304, No. 3-5, 2007, pp. 722-734.

[12] Annigeri A.R., Ganesan N. and Swarnamani S., "Free vibration behaviour of multiphase and layered magneto-electro-elastic beam", *Journal of Sound and Vibrations*, Vol. 299, No. 1-2, 2007, pp. 44-63.

[13] Pan E., "A BEM Analysis of Fracture Mechanics in 2D Anisotropic Piezoelectric Solids", *Engineering Analysis with Boundary Elements*, Vol. 23, No. 1, 1999, pp. 67-76.

[14] Denda M. And Lua J., "Development of the boundary element method for 2D piezoelectricity". *Composites Part B-Engineering*, Vol. 30, No. 7, 1999, pp. 699-707.

[15] Davì G. And Milazzo A., "Multidomain boundary integral formulation for piezoelectric materials fracture mechanics". *International Journal of Solids and Structures*, Vol. 38, No. 40, 2001, pp. 7065-7078.

[16] Ding H.J. and Jang A.M., "A boundary integral formulation and solution for 2D problems in magneto-electro-elastic media", *Computers and Structures*, Vol. 82, No. 20-21, 2004, pp. 1599-1607.

[17] Li X.C. and Yao W.A., "Virtual boundary element-integral collocation method for the plane magnetoelectroelastic solids", *Engineering Analysis with Boundary Elements*, Vol. 30, No. 8, 2006, pp. 709-717.

[18] Milazzo A., Benedetti I. and Orlando C., "Boundary element method for magneto-electro-elastic laminates", *Computer Modeling in Engineering and Sciences*, Vol. 15, No. 1, 2006, pp. 17-30.

[19] Davì G., Milazzo A. and Orlando C., "Magneto-Electro-Elastic Bimorph Analysis by the Boundary Element Method", *Mechanics of Advanced Materials and Structures*, Vol. 15, No. 3-4, 2008, pp. 220-227.

[20] Partdrige P.W., Brebbia C.A. and Wrobel L.C. *The dual reciprocity boundary element method. Southampton*, Computational Mechanics Publications, 1992.

[21] De Figueredo T.G.B. and Brebbia C.A., "A new

hybrid displacement variational formulations of BEM for elastostatics", *Int. Conf. on Boundary Element Method*, Southampton: Computational Mechanics Publications and Berlin:Springer Verlag, 1989, pp. 47-57.

[22] Davì G., "A hibryd displacement variational formulation of BEM for elastostatics", *Engineering Analysis with Boundary Elements*, Vol. 10, No. 3, 1992, pp. 219-224.

[23] [45] Gaul L. Moser F. Non-singular symmetric boundary element formulation for elastodynamics. *Engineering Analysis with Boundary Elements*, Vol. 25, No. 10, 2001, pp. 843-849.

[24] Zhang J. And Yao Z., "The regular hybrid boundary node method for three-dimensional linear elasticity", *Engineering Analysis with Boundary Elements*, Vol. 28, No. 5, 2004, pp. 525-534.

[25] Davì G. And Milazzo A.. "A symmetric and positive definite variational BEM for 2-D free-vibration analysis", *Engineering Analysis with Boundary Elements*, Vol. 14, No. 4, 1994, pp. 343-348.

[26] Davì G. and Milazzo A., "A symmetric and positive definite BEM for 2-D forced vibrations", *Journal of Sound and Vibrations*, Vol. 206, No. 4, 1997, pp. 611-617.

[27] Davì G. and Milazzo A., "A new symmetric and positive definite boundary element formulation for lateral vibrations of plates" *Journal of Sound and Vibrations*, Vol. 206, No. 4, 1997, pp. 507-521.

[28] Davì G. and Milazzo A., "A meshfree method for transverse vibrations of anisotropic plates", *International Journal of Solids and Structures*, Vol. 40, No. 20, 2003, pp. 5229-5249.

[29] Liu G.R., *Meshfree Methods: Moving beyond the Finite Element Method*. CRC Press, 2009.

[30] Davì G. and Milazzo A., "A regular variational boundary model for free vibrations of magneto-electro-elastic structures", *Engineering Analysis with Boundary Elements*, Vol. 35, No. 3, 2011, pp. 303-312.

[31] Kögl M. and Gaul L., "Free vibration analysis of anisotropic solids with the boundary element method", *Engineering Analysis with Boundary Elements*, Vol. 27, No. 2, 2003, pp. 107-114.

[32] Kögl M. and Gaul L., "A boundary element method for transient piezoelectric analysis", *Engineering Analysis with Boundary Elements*, Vol. 24, No. 7, 2000, pp.591-598.

[33] Dziatkiewicz G. and Fedelinski P., "Dynamic analysis of piezoelectric structures by the dual reciprocity boundary element method", *Computer Modeling in Engineering and Sciences*, Vol. 17, No. 1, 2007, pp. 35-46.

[34] Alaimo A., Milazzo A. and Orlando C., "Boundary Elements Analysis of Adhesively Bonded Piezoelectric Active Repair" *Engineering Fracture Mechanics*, Vol. 76, No. 4, 2009, pp. 500-511.

[35] Alaimo A., Milazzo A. and Orlando C., "Dynamic analysis of damaged magnetoelectroelastic laminated structures", *Int. Conf. on Boundary Element Techniques XI*, Eastleigh: EC Ltd, 2010, pp. 324-329.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

The Effect of Laser Shock Peening on Fatigue Behaviour of Aluminium Alloy 6082-T6

Stefano Crestani, Goran Ivetic, Ivan Meneghin, Enrico Troiani
University of Bologna, Second Faculty of Engineering, DIEM - Aerospace Engineering
Via Fontanelle 40, 47121 Forlì, Italy

Keywords: *Laser Shock Peening, Fatigue, Aluminium Alloys*

Abstract

Laser Shock Peening (LSP) is an innovative surface treatment technique used to improve the fatigue performances of metallic components. This technique uses a high power density laser pulse to create a compressive residual stress field in the treated component; the established residual stress field is at the base of the fatigue performance enhancement. In this work, thick 6082-T6 aluminum alloy specimens were LSP treated and subsequently fatigue tested. An increase of fatigue life and surface hardness was encountered in respect to the untreated material. Besides confirming the results present in the literature, this experimental campaign served to confirm the capability of the University of Bologna to perform in-house an advanced surface treatment such as Laser Shock Peening.

1 Introduction

Laser Shock Peening (LSP) uses short laser pulses (1-50 ns) of high power intensity (order of GW/cm^2), fired at the surface of the metal component to be treated. The high energy laser beam vaporizes a superficial layer of the treated material forming locally high pressure temperature plasma. A transparent overlay (usually a thin water layer flowing on the metal surface) confines the generated plasma intensifying its pressure and inducing intense shock weaves directed toward the component thickness.

LSP applications usually use a "sacrificial" ablative layer (a coating painted on the surface) to form the plasma without damaging the metal surface; anyway, LSP applications exist where no ablative layer is used during the exposure to laser pulses and the treated specimen is emerged into water. This application of LSP, developed by Toshiba, is called Laser Peening without Coating (LPwC).

A good overview of LSP technology can be found in [1], [2] and [3].

In the work presented here, the LPwC approach has been used.

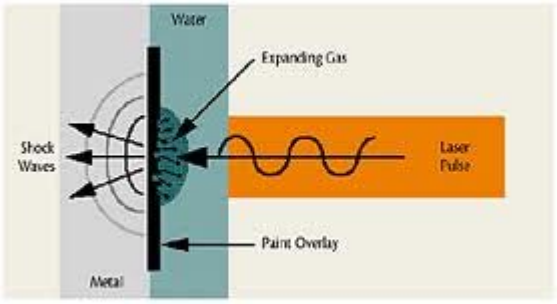


Figure 1. Laser Shock Peening

2 Experimental setup

The LSP treatment was performed at the DIEM Laboratories of the Faculty of Engineering of the University of Bologna. The specimens used for the analysis were made of aluminium alloy 6082-T6, with the geometry given in Fig. 1. The choice of the thickness of the specimen was governed by the fact that LSP treatment of relatively thick specimens does not require any special process optimization, as it would be in the case of thin specimens, where shock wave reflections are present. The shape of the specimen (three-point bending specimen) was chosen in according to already present experimental work done in the 1990s by the French scientists [4], which was used as a reference point for the evaluation of the performance of the LSP treatment carried out in the work presented here.

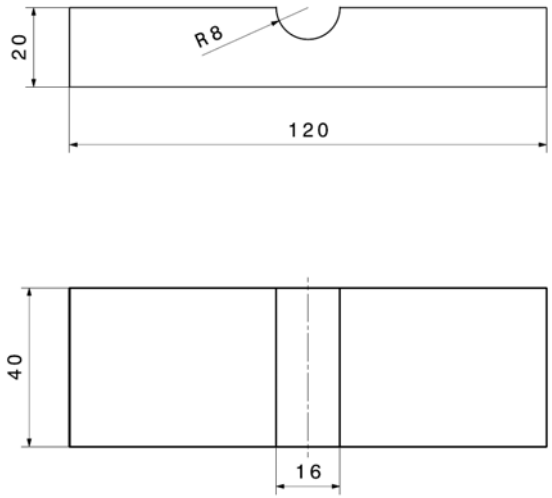


Figure 1. Specimen geometry

The treatment was performed using an Nd:YAG pulsed laser provided by Quanta Systems, with wavelength of 1064 nm, nominal pulse energy of 1.8 J, 8 ns pulse duration. The laser peen size was 2 mm in diameter with the distance between two successive spot centres of 0.25 mm; the spot overlapping permitted to covered the whole treated area. The experimental setup can be seen in Fig. 2.

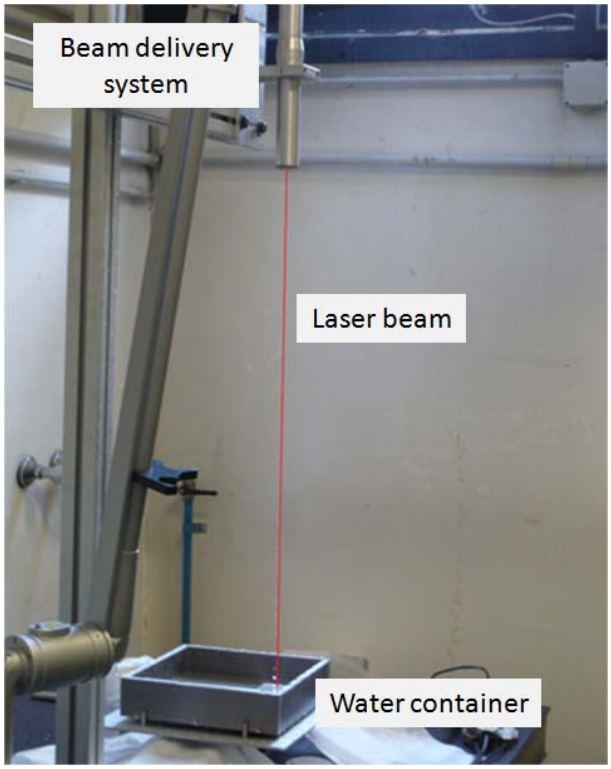


Figure 2. LSP treatment setup

Only the central zone of the specimen was treated, with the dimension of the treated area of 40x4 mm. Figure 3 shows the specimen after LSP treatment.

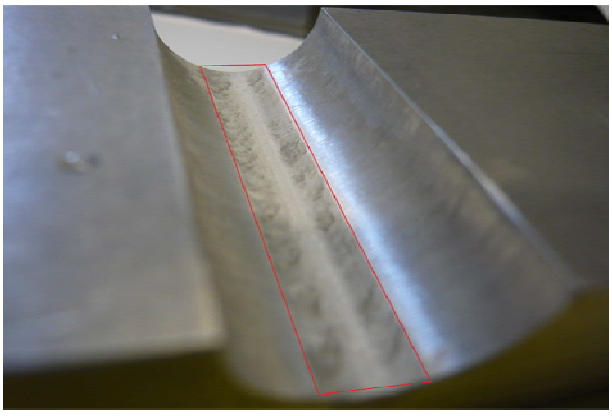


Figure 3. The treated zone of the specimen

After LSP treatment, three-point bending fatigue tests were performed at the Laboratories of the Forlì campus of the Second Faculty of Engineering of the University of Bologna. A 100 kN servo-hydraulic testing machine was used, with stress ratio R=0.1 and maximum stresses at the specimen notch ranging from 220

to 260 MPa. Figure 4 illustrates the specimen during testing.



Figure 4. Specimen during three-point bending test

In addition, Brinell hardness tests were performed in order to quantify the effect of LSP treatment on the increase of surface hardness of treated specimens.

3 Results

3.1 Fatigue tests

The total number of tested specimens was eight – i.e. four untreated specimens in order to obtain the baseline distribution and four LSP treated specimens. Table 1 illustrates the results obtained in the terms of fatigue lives of treated specimens for different nominal loads.

Table 1. Fatigue lives of treated specimens

σ max [MPa]	Cycles to failure – baseline	Cycles to failure – LSP treated	Fatigue life increase
220	11.30×10^4	32.15×10^4	2.84x
240	5.97×10^4	19.84×10^4	3.32x
250	2.85×10^4	9.54×10^4	3.34x
260	3.69×10^4	6.40×10^4	1.73x

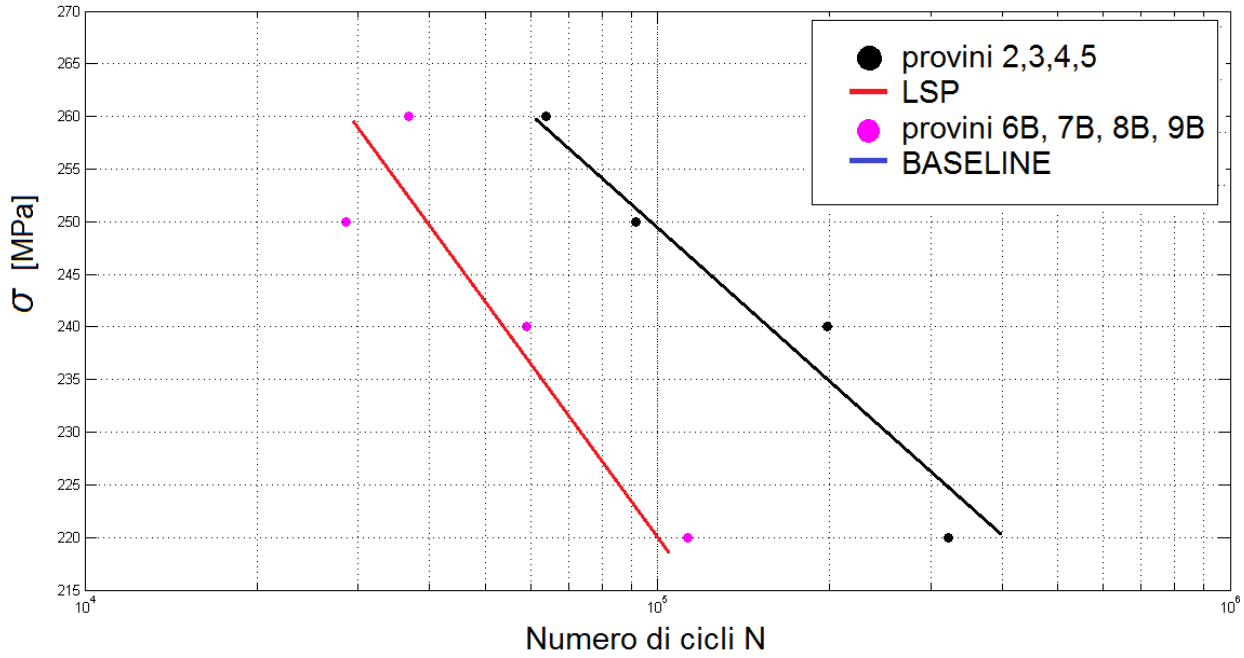


Figure 5 Maximum nominal stress versus fatigue lives of tested specimens

Even if eight specimens in total is not sufficient for construction of a statistically significant S-N curve, it is very clear from the tests performed that the influence of LSP treatment was very beneficial in the terms of the increase of observed fatigue lives. In Fig. 5, curves illustrating maximum stress vs. fatigue lives of specimens are given.

It is important to remember that the treatment itself did not undergo any additional

optimization of the LSP treatment parameters, meaning that there should be margins for further improvement of the already beneficial results.

3.2 Hardness measurement

The increasing in surface hardness has been measured using Brinell hardness tests. Total of five measurements were performed, two in the central, LSP treated section of the specimen, Fig. 6 and three on the untreated part of the specimen, Fig. 7. AS it can be seen in the figures, in order to accommodate the specimens to the hardness testing machine, the specimen was appropriately machined. This additional machining, however, did not interfere significantly with the LSP treated zone and the measured values should be very close to the exact ones. The obtained hardness values are given in Table 2.

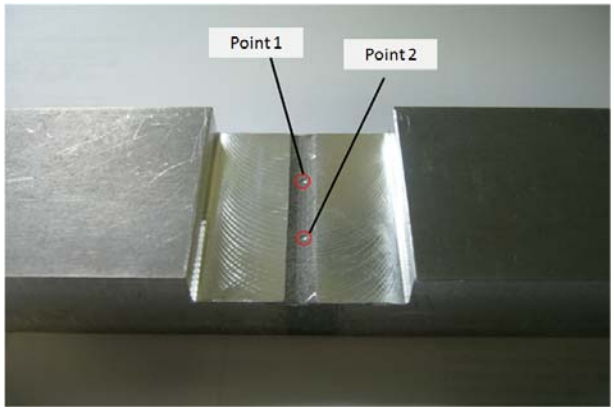


Figure 6. Hardness measurement points – LSP treated part of the specimen

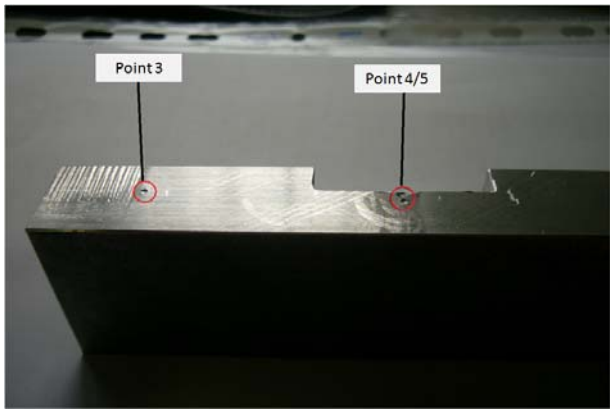


Figure 7 Hardness measurement points – untreated specimen

The increase in surface hardness is of order of magnitude of 50% which is in line with the previous findings presented in the literature [5]. A slightly increased hardness of the as machined measurement points in respect to the nominal Brinell hardness of 6082-T6 of 105 HB can be attributed to the machining operation itself.

Table 2. Brinell hardness measurement results

Measurement point	Surface state	Brinell hardness
1	LSP treated	178 HB
2	LSP treated	185 HB
3	As machined	123 HB
4	As machined	117 HB
5	As machined	120 HB

4 Conclusions

Concerning the results obtained in this experimental work, the following conclusions can be drawn:

- The fatigue results show that, the LSP treatment induced a high level compressive residual stresses that improved fatigue life, as suggested in reference [6], as well.
- In particular, the fatigue life of laser peened specimens was as much as three times higher than as-machined specimens (baseline), confirming the presence of a compressive residual stress field and the beneficial effect of the LSP treatment on the treated specimens.
- The Brinell tests confirmed the ability of LSP to increase surface hardness, thus improving wear capabilities of the treated specimens
- Finally, this experimental test has proven the capability of performing completely an advanced treatment such as LSP in-house, at the University of Bologna.

References

[1] Kruusing, A., *Handbook of Liquids-Assisted Laser Processing*, 1st ed., Elsevier, Oxford, 2008, Chap. 3.

[2] Ding, K. and Ye, L., *Laser shock peening: Performance and process simulation*, Woodhead Publishing, 1st ed., Cambridge, 2006, Chap. 2.

[3] Sano, Y., Mukai, N., Okazaki, K. and Obata, M., "Residual stress improvement in metal surface by underwater laser irradiation", Nuclear Instruments and Methods in Physics Research Section B, Vol. 121, 1997, pp. 432-436.

[4] Peyre, P., Fabbro, R., Merrien, P. and Lieurade, H.P., "Laser shock processing of aluminium alloys. Application to high cycle fatigue behaviour", Materials Science and Engineering A, Vol. 210, 1996, pp. 102-113.

[5] Zhang H., Zhang Y.-K. and Yu C.-Y., "Surface treatment of aluminium alloy by laser shock processing", Surface Engineering, Vol. 15 (6), 1999, pp. 454-456.

[6] Gomez-Rosas, G., Rubio-Gonzalez, C., Ocana, J.L., Molpeceres, C., Porro, J.A., Morales, M. and Casillas, F.J., "Laser Shock Processing of 6061-T6 Al alloy with 1064 nm and 532 nm wavelengths", Applied Surface Science, Vol. 256, 2010, pp. 5828-5831.

CEAS 2011

Solid Propellants and Launchers

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Relative comparison of additives-filled solid fuels and grain geometry effects on regression rate in hybrid rocket propulsion

L. Galfetti, L.T. De Luca, F. Maggi, L. Merotto, M. Boiocchi

Politecnico di Milano
Aerospace Engineering Department, Space Propulsion Laboratory (SPLab)
I-20156, Milano, MI, Italy

Keywords: *hybrid rocket engines, regression rate, solid fuels, metal hydrides, grain geometry*

Abstract

Some innovative solid fuels for hybrid rocket propulsion were tested in order to overcome the shortcoming of very low fuel regression rates, typical of conventional hybrid rocket fuels. Nano-Aluminum, Magnesium hydride and Lithium Aluminum hydride were used as fillers in a solid wax (SW-) matrix. Pure HTPB was used as reference fuel. The highest regression rates were found for solid wax added with 5% MgH_2 and with 6% $LiAlH_4$ mass fraction. The measured regression rate, under the operating conditions of the experimental setup, was approximately four times higher than that of the pure HTPB reference composition, tested under the same conditions. The non-filled fuel formulation was furthermore tested in different geometries (slab and radial) under the same combustion area, and in the same geometry (radial) with different combustion area.
Aim of the first part of this paper is to show through a relative comparison the regression rate increase of hybrid rockets fuels when peculiar energetic additives are considered. Aim of the second part is to show to what extent grain geometry effects affect the regression rate, in view of the extrapolation of results from a laboratory-scale setup to a full-.scale system.

1 Introduction

The need to obtain much higher regression rates of solid fuels designed for hybrid rockets is well known inside the propulsion community. Research activities are carried out, following different strategies [1], in several laboratories all over the world. The solid fuel regression rate is the most relevant parameter for the hybrid motor design; it is the key variable for the development and the future achievement of hybrid rocket propulsion technology. The regression rate is affected by the oxidizer injection process, oxidizer flux, turbulent and radiant heat transfer, pressure, solid grain geometry, gas-phase kinetics and, eventually, heterogeneous reactions occurring at the solid fuel surface [2]. To increase the regression rate a further research path concerns the effects of fuel additives. Metal additives can raise the theoretical flame temperature, produce a strong radiation flux from the metal oxide combustion products, reduce the blocking effect of the blowing pyrolyzed mass. These effects, due to addition and metal powders combustion, can determine higher fuel regression rates [3], compared with the regression rates of traditional non-metalized solid fuels.

Liquefying solid fuels were proposed by Karabeyoglu and other researchers [4-7];

regression rates are usually much faster than those of traditional hybrid fuels, because of droplet entrainment process caused by the fuel surface melting.

The problem of scale and geometry factor is a well known limitation for the hybrid propulsion development. The prediction of a hybrid rocket engine combustion process, characterized by complex interactions of several physical phenomena occurring simultaneously, has generally been based on simplified empirical methods and correlations. Predictions based on these correlations when applied to different-scale, different geometry and/or different operating parameters often lead to erroneous results. A study by Estey et al. [10] emphasizes the limitations of such parametric correlations aimed at extrapolation to different scales. In a more recent work by Swani and Gany [11], a theoretical similarity analysis is presented in order to derive proper scaling rules in hybrid rocket motors, and a fairly good agreement between the test-results and theoretical predictions is obtained. Nevertheless, further investigation is needed in order to assess the geometry contribution to overall fuel performance.

This paper has two main purposes: 1. to discuss the most recent experimental results obtained at the Space Propulsion Laboratory (SPLab) of Politecnico di Milano concerning the ballistic characterization of paraffin-based solid fuels doped with metal hydrides in a double slab geometry; 2. to examine the grain geometry effects, especially on the average regression rate behavior.

2 Materials and Methods

2.1 Fuel Formulations

The fuel formulations tested in this work are based on a solid wax (SW-) with average carbon atoms number 24. The main properties of the solid paraffin wax used as fuel binder are shown in

Table 1. The chemical formula is reported, with the theoretical density and the measured viscosity at 333 *K* and at 343 *K*. Specific heat at constant pressure, melting temperature and surface tension are also reported.

In order to overcome the problem of the poor mechanical properties typical of paraffin waxes, a strengthening structure based on a Poly-Urethane Foam (PUF) was used, as shown in Fig. 1. The melted paraffin is poured in the PUF structure in the mould, under vacuum conditions. The obtained fuel sample displays feasible mechanical properties.

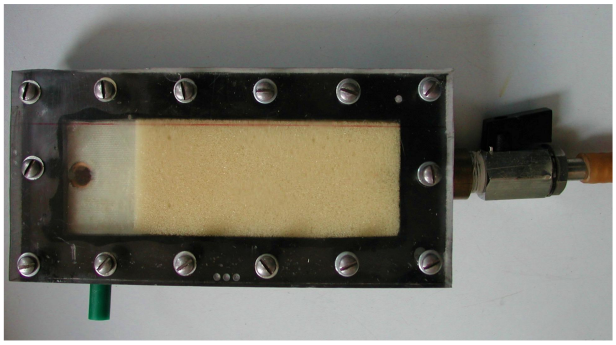


Fig. 1 PUF strengthening structure for paraffin-based fuel formulations.

Table 1 Binder main characteristics.

Solid Wax: $C_{24}H_{50}$	
Density (g/cm^3)	0.89
Measured viscosity at T = 333 *K* ($Pa\ s$)	0.89
Measured viscosity at T = 343 *K* ($Pa\ s$)	0.09
Specific heat at constant pressure, $C_{p,liquid}$ at T = 330 *K* ($J/mol\ K$)	772.5
Specific heat at constant pressure, $C_{p,solid}$ at T = 300-500 *K* ($J/mol\ K$)	601
Melting temperature, T_m (K)	324
Surface tension, σ, at T = 243 *K* (mN/m)	30.1

The energy increase of solid fuels for hybrid rockets applications is one of the main topics of

current research activities in the field of aerospace propulsion. The most common used fuel additive, Aluminum, could be replaced by a series of metal hydrides, with significant increases in the specific impulse (I_s), as shown in Table 2. I_s data reported in Table 2 are obtained with equilibrium calculations, with liquid oxygen as reference oxidizer and with expansion ratio of 40 for a combustion chamber operating at 7 MPa [12].

Among the selected hydrides of Table 2.1, the most effective seems to be AlH_3 (Aluminum hydride, also known as alane). It has a high combustion heat and gives low molecular weight gas. Replacing Al with alane always significant increases in the theoretical I_s are observed. Problems in using alane in solid fuels are the quite complex synthesis techniques, the thermal stability, and the short shelf life. Moreover, alane is not available commercially. For these reason, alane was not selected for this work.

Table 2 Physico-chemical properties and performance of selected hydrides for propulsion applications.

	AlH_3	$LiAlH_4$	MgH_2
MM *(g/mol)*	30.00	37.95	26.34
ρ *(g/cm^3)*	1.48	0.92	1.45
ΔH_f *(kcal/mol)*	-11.1	-28.0	-21.7
T_{dec} *(°C)*	~ 180	~ 137	~ 270
C_p @ 25°C *(cal/mol K)*	-	20.7	8.5
$I_{s,max}$ *(s) (source: [12])*	378	370	355
ΔI_s *(s) (Ref: HTPB)*	+ 23	+ 15	=

$LiAlH_4$ (Lithium Aluminum hydride) is a very strong reduction agent employed in organic and inorganic chemistry. It is a stable solid substance in dry atmosphere. Its stability, investigated using DSC techniques, is characterized by a three stage decomposition process. In the temperature range 154-161 °C half of the contained H_2 is released; in the range 197-227 °C approximately the 25% is released, between 580 and 586 °C all H_2 is released and the melting occurs. A very strong reaction, almost complete, is observed with water when the product is in form of powder, giving:

$$LiAlH_4 + 4H_2O \rightarrow LiAl(OH)_4 + 4H_2 \qquad (1)$$

MgH_2 (Magnesium hydride) is a solid substance, stable up to 125 °C. Under vacuum conditions, decomposition occurs in the temperature range 250-300 °C.

Fuel formulations investigated in this work are doped with nano-Aluminum powder (Alex) with average particle size of 100 nm, with a Magnesium hydride (MGH) powder, or with Litium Aluminum hydride (LAH). Both Magnesium hydride and Lithium Aluminum hydride have coarser particles (average particle diameter in the range 50-150 μm) than that typical of Alex, but higher reactivity. The nomenclature used for the tested fuel formulations, and the additive mass fraction included in each formulation is reported in Table 3.

Table 3 Fuel formulations composition.

Fuel	Additive type	Additive mass fraction
SW-Alex5	nAl	5%
SW-MGH5	MgH_2	5%
SW-LAH3	$LiAlH_4$	3%
SW-LAH6	$LiAlH_4$	6%

2.2 Experimental Set Up

The firing tests are performed using a slab/radial burner, shown in Fig. 2. The combustion chamber is equipped with a pressure transducer, a manometer, an oxygen and a nitrogen inlet system. Nitrogen is used for a quick extinction of the oxidation reactions after the oxygen shut off. Different sample

holders were designed and set up for testing different sample geometry: double slab or radial samples.

The fuel slabs have the following size: 50 x 10 x 4 mm. Radial samples have an internal port diameter of 4.8 mm, external diameter of 20 mm and length of 30 or 56 mm. A picture of typical slab and radial samples is shown in Fig. 3.

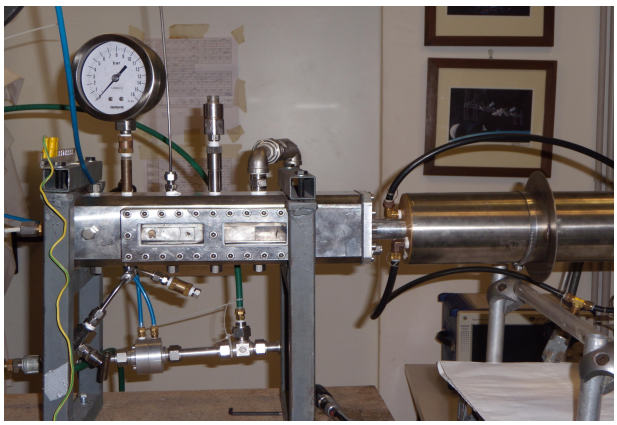


Fig. 2 Experimental set up used for this work.



Fig. 3 Slab geometry fuel samples (left) and radial fuel sample (right).

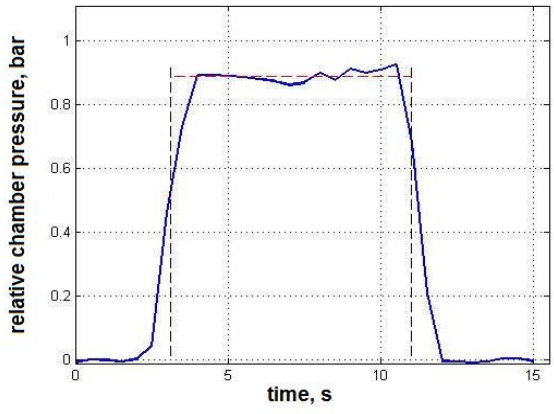


Fig. 4 Relative chamber pressure vs. time in a typical firing test.

The average regression rate (r_f) is measured from the burned fuel mass (Δm), the fuel density (ρ_f), the burning time (t_b) and the burning area (A_b), using the following equation:

$$r_f = \frac{\Delta m}{\rho_f t_b A_b} \qquad (2)$$

The pressure trace obtained during a firing test is used to determine the combustion time. A typical pressure trace recorded during a firing test is shown in Fig. 4.

3 Results Discussion

3.1 Double slab geometry tests

Firing tests in double slab configuration were performed with oxygen mass flux ranging from 150 to 350 kg/m^{2}s, at 1.5 bar operating pressure. Average regression rate vs. oxygen mass flux is shown in Fig. 5. The regression rate results obtained for a reference oxygen mass flux (150 kg/m^{2}s) are reported in Fig. 6.

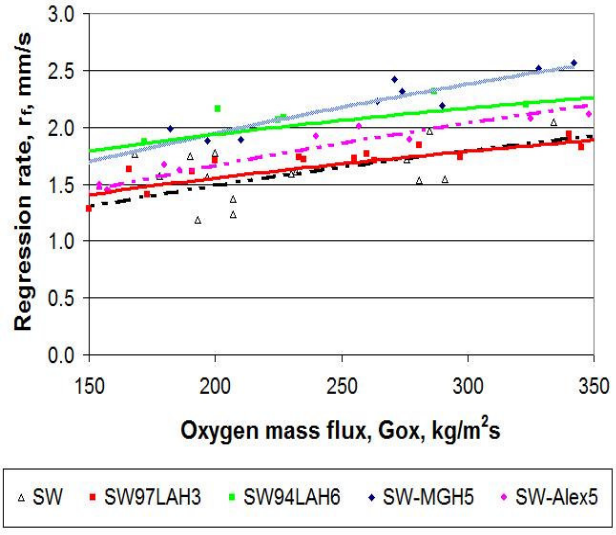


Fig. 5 Regression rate vs. oxygen mass flux for SW fuel formulation. Operating pressure: 1.5 bar.

Results show that all the SW-based fuels display higher regression rate when compared to the baseline formulation (pure HTPB).

Moreover, all the filled fuels display higher r_f when compared to the non-filled SW formulation.

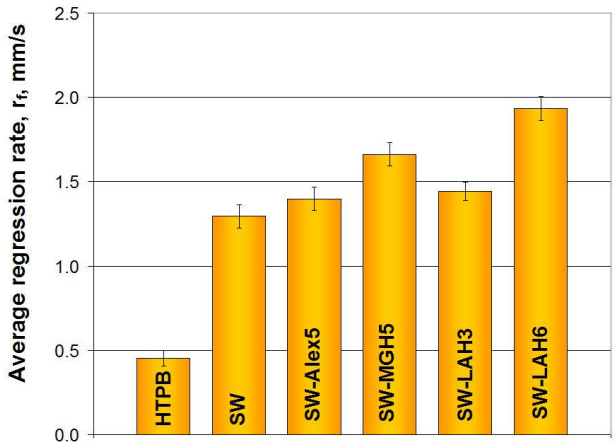


Fig. 6 Average regression rate comparison for SW fuel formulation. Oxygen mass flux: 150 kg/m^{2}s. Operating pressure: 1.5 bar.

The regression rate increase compared to the baseline formulation (pure HTPB) and to the non-filled solid wax fuel are reported in Table 4, for the reference oxygen mass flux. The best results are obtained with 5% MgH_2 and with 6% $LiAlH_4$ filled fuels: the regression rate increase compared to pure HTPB is +285% and +304%, respectively.

Table 4 Performance increase compared to pure HTPB and to SW, at 150 kg/m^{2}s oxygen mass flux.

Fuel type	r_f % variation vs. HTPB	r_f % variation vs. SW
SW	+187%	--
SW-Alex5	+224%	+13%
SW-MGH5	+285%	+34%
SW-LAH3	+219%	+11%
SW-LAH6	+304%	+40%

For all the tested formulations, a power law fitting the experimental data was obtained. The coefficients obtained from the fitting are reported in Table 5. LAH filled formulations display the lowest r_f increasing rate with increasing oxygen mass flux (n = 0.28 - 0.35), while for the other formulations the exponent is about 0.5.

Table 5 Coefficients of the power law fitting the experimental data.

Fuel type	$r_f = a\,(Gox)^n$	
	a	n
SW	0.15 ± 0.02	0.46 ± 0.05
SW-Alex5	0.12 ± 0.03	0.50 ± 0.03
SW-MGH5	0.15 ± 0.03	0.49 ± 0.03
SW-LAH3	0.25 ± 0.02	0.35 ± 0.05
SW-LAH6	0.45 ± 0.02	0.28 ± 0.04

The obtained results prove that metal hydrides are more effective in enhancing r_f than nAl is, due to the hydrides high hydrogen content. It is interesting to note that 3% LAH addition leads only to 13% r_f increase compared to SW, while a 40% increase is obtained when adding 6% LAH. This suggests that two different effects compete in determining the overall performance: on the one hand, a positive effect results from hydride addition, due to the high hydrogen content. On the other hand, the increased viscosity of the fuel melted layer, resulting from metal addition to the fuel formulation, is detrimental to ballistic performance.

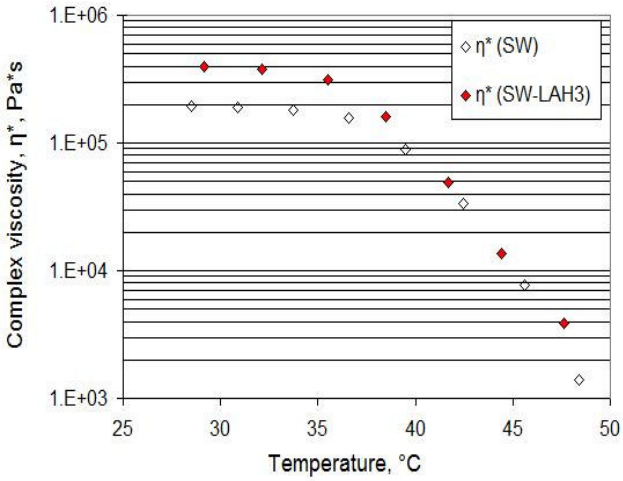


Fig. 7 Complex viscosity for SW and SW-LAH3 formulations. Strain = 1%.

The increased viscosity is confirmed by a rheological analysis, performed in a parallel-plate rheometer and shown in Fig. 7, in which the complex viscosity measured for SW and SW-LAH3 formulations is plotted vs. temperature. Complex viscosity is a frequency dependent viscosity function determined during forced harmonic oscillation of shear stress. It represents the angle between the viscous stress and the shear rate and is equal to the difference between the dynamic viscosity and the imaginary part of the complex viscosity (out-of-phase viscosity). Data reported in Fig. 7 show that the complex viscosity measured at low temperatures for SW-LAH3 is higher (about 4 x 10^5 Pa s) than that measured for non-filled SW (about 2 x 10^5 Pa s), thus confirming a more viscous behavior of SW-LAH3 formulations.

Fig. 8 shows the mixture ratio (O/F) for a pure SW-based fuel and a fuel added with 5% (mass fraction) Magnesium hydride. The O/F ratio of the wax fuel, filled with the metal hydride powder, ia close to the stoichiometric value (3.4), while the O/F ratio of the pure wax fuel is distributed in a O/F range spanning from 1.5 to approximately 4.5, being the stoichiometric value $O/F_{st} = 2.75$.

These results show the complex interactions occurring in the combustion chamber; several physical phenomena are involved, not only the traditionally investigated fluid dynamic, heat transfer and combustion processes, but also the thermophysical properties of the solid fuel (i.e., the viscosity of the melted layer at the fuel surface, affected by the metal powder mass fraction) and the important role of the O/F ratio, which affects the maximum temperature and the temperature distribution.

Increasing the Lithium Aluminum hydride mass fraction from 3% to 6% implies an increase of the heat released, whose effectiveness is much stronger of the decreased entrainment effect due to the melt layer increased viscosity. A 6% mass fraction leads to a measured regression rate increase of 40%; a 3% leads to an increase of only 11%. A similar comment could apply to the regression rate increase (34%, Tab. 4) obtained with 5% mass fraction of Magnesium hydride.

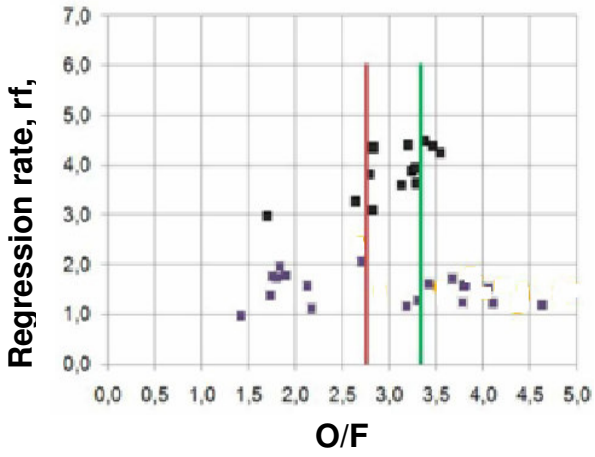


Fig. 8 Regression rate vs. O/F for SW fuel formulation in double slab geometry. Comparison between pure solid wax (larger squares, black) and solid wax added with MgH_2 (smaller squares, blue, in the lower part of the diagram).

3.2 Radial and slab geometry comparison

The main purpose of the comparison of radial and slab geometry results is not to investigate a similarity analysis (by identifying similarity laws of the most relevant controlling processes); it is to check the conditions of a laboratory-scale system to be used to simulate the behavior of a full-scale system.

For the hybrid rocket combustion chamber, it is known that the solid fuel geometry affects the flowfield, leading to the geometric similarity as a basic requirement to simulate the behavior of an untested rocket. It is also known that scaling rules require the scaling of the oxidizer flow rate in proportion to the port diameter and the use of the same oxidizer and fuel [12]. Under these conditions different scale systems are predicted to give an inverse dependence between r_f and D, constant values of mixture ratio O/F and characteristic velocity c*, approximately constant I_s and linear proportion between the thrust F, and the port diameter, D.

The 2D fuel geometry, which is very useful to investigate boundary layer, flame structure and fuel surface behavior, can be useful also to extend and use the 2D available test data to untested systems of a different geometry (typically radial) and a different size range ? What is the difference, in terms of regression

rate, between results of 2D slab and radial geometry, when the same fuel surface, oxidizer mass flux, operating pressure, injection and ignition systems are used ?

In order to assess the effect of geometry on the overall fuel performance, firing tests are carried out using different geometry SW samples. Double slab samples are compared to radial samples having 30 and 56 mm length. The burning area (A_b) is similar for double slab and 56 mm radial samples, while 30 mm radial samples have a lower burning area (-56%). All tests are performed using the same injection and ignition systems. A simple axial injector, located at the same distance from the grain head (same pre-chamber length for both double slab and radial configuration) is used. Ignition of the solid fuel sample is always successfully induced by a small charge of solid propellant ignited by a hot wire. Results are shown in Fig. , in terms of regression rate vs. oxygen mass flux. Tests were performed at 1.5 bar operating pressure, with oxygen mass flux ranging from 150 to 250 kg/m^{2}s.

A further comment concerns the Reynolds number. It is well known that hybrid rocket systems are affected more by physical than chemical processes. Turbulent flowfield, turbulent transport phenomena and diffusion flame characteristics have therefore greater significance than chemical aspects under the traditional operating conditions occurring in the hybrid rocket combustion chamber and, as a consequence, Reynolds number is one of the main controlling parameters to operate in similarity conditions. For results presented in Fig. 9, Reynolds number estimated at inlet conditions at the beginning of the test, is the same for double slab and radial geometry.

Results show that the highest regression rate is obtained with 56 mm long radial samples (up to 5 mm/s in the tested operating conditions, at 250 kg/m^{2}s oxygen mass flux), while 30 mm radial samples display r_f up to 4 mm/s. The lowest performance is obtained with double slab samples (up to 1.9 mm/s at 250 kg/m^{2}s oxygen mass flux).

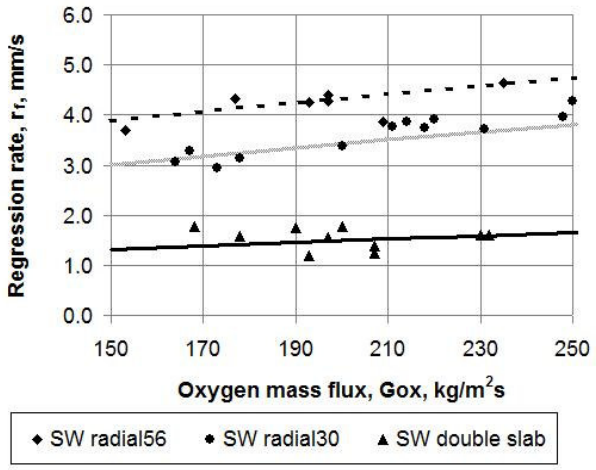


Fig. 9 Regression rate vs. oxygen mass flux for SW fuel formulation. Operating pressure: 1.5 bar.



Fig. 10 Regression rate vs. O/F for SW fuel formulation. Double slab (triangles) and radial 30 mm long (circles) configurations. Operating pressure: 1.5 bar.

The O/F ratio during each firing test was also measured in order to assess the effect of O/F ratio on ballistic performance. The O/F ratio was obtained from the fuel mass burned during each firing test, and the corresponding oxygen mass injected in the combustion chamber. Measured average regression rate is plotted vs. O/F in Fig. 10. The stoichiometric O/F value was obtained assuming the following reaction occurs during SW combustion:

$$C_{24}H_{50} + 73O_2 \rightarrow 25H_2O + 24CO_2 \qquad (2)$$

Examining the results reported in Fig. 10, it can be noticed that the results obtained using radial geometry (circles in Fig. 10) fall in a O/F range very close to the stoichiometric value (2.75, red line in Fig. 10), while using double slab geometry (triangles in Fig. 10) most of the data correspond to higher or lower O/F values. The higher regression rate values obtained for the radial configuration can be explained taking into account the radiative contribution to thermal exchange between flame and sample surface, which is likely to be higher in radial geometry, the O/F trend and the different geometry of the gas-phase diffusion flame established between the fuel products and the oxidizer. The boundary layer developing over the fuel surface is strongly affected by the slab or radial geometries.

A geometry factor can be defined to characterize the investigated geometry configurations; a simple definition is given by the regression rate ratio. It was calculated under the operating conditions tested, and the results are summarized in Table 6. A geometry factor of about 3 is obtained when radial and slab geometries having the same A_b are compared. When comparing two radial geometries having different length (A_b ratio equal to 2.28), a geometry factor of 1.3 is obtained. Results suggest that the geometry contribution to r_f , affecting flowfield, O/F ratio and the radiative contribution to heat exchange between flame and fuel surface due to radial geometry, are responsible for a three times r_f increase. For radial samples having A_b ratio of 2.28, an increase of 1.3 r_f is obtained in the tested oxygen mass flux range. This enhancement is likely due to the increased fuel mass flux caused by the larger combustion area.

Table 6 Computed geometry factor for slab/radial samples.

r_f ratio		
Gox range *(Kg/m^2s)*	A_b ratio =1, radial vs. slab	A_b ratio = 2.28, radial 56 mm vs. radial 30 mm
150-250	2.9	1.3

4 Concluding remarks

Some innovative solid fuels for hybrid rocket propulsion were tested in order to compare their average regression rate. A series of solid wax-based fuels, added with nano-Aluminum, Magnesium hydride and Lithium Aluminum hydride was investigated in a double slab setup.

The highest regression rates were found for solid wax added with 5% MgH_2 and with 6% $LiAlH_4$ mass fraction. The measured regression rate, under the operating conditions of the experimental setup, was approximately four times higher than that of the pure HTPB reference composition, tested under the same conditions (1.5 bar operating pressure, and oxygen mass flux ranging from 150 to 350 kg/m^2s). Two different effects appear to compete in determining the overall performance: on the one hand, a positive effect results from metal hydride addition, due to its high hydrogen content. On the other hand, the increased fuel viscosity, resulting from metal addition to the fuel formulation, is detrimental to ballistic performance due to the decreased entrainment effect.

SW fuels were tested in different geometries in order to investigate geometry effects for two different configurations. A geometry factor, defined by the average regression rate ratio, was found to be 3 for the same combustion area and different geometry (double slab and radial), and 1.3 for the same geometry (radial) and different combustion area, due to the different sample length (30 and 56 mm, respectively). Results suggest that the geometry contribution to r_f, affecting flowfield, O/F ratio and the radiative contribution to heat exchange between flame and fuel surface, is responsible for a three times r_f increase. For radial samples having a combustion area ratio of 2.28, an increase of 1.3 r_f is obtained in the tested oxygen mass flux range.

The message of this study is that the regression rate behavior of a solid fuel, even when it is filled with energetic additives, is strongly affected not only by operating conditions but also by fuel composition and thermophysical properties. The grain geometry

too strongly affects the regression rate. The obtained results stress the need to maintain a geometric similarity for the scaling of hybrid rocket motor.

References

[1] K. H. Shin, C. Lee, S.Y. Chang, and J.Y. Koo, "The Enhancement of Regression Rate of Hybrid Rocket Fuel by Various Methods", AIAA 2005-359.

[2] M.J. Chiaverini, N. Serin, D.K. Johnson, Y.C. Lu, and G.A. Risha, "Regression Rate Behavior of Hybrid Rocket Solid Fuels", *Journal of Propulsion and Power*, 16(1), 2000.

[3] G.A. Risha, E. Boyer, R.B.Wehrman, and K.K. Kuo, "Performance Comparison of HTPB-based Solid Fuels Containing Nano-sized Energetic Powder in a Cylindrical Hybrid Rocket Motor", 2002. AIAA 2002-3576.

[4] M.A. Karabeyoglu, "Combustion Instability and Transient Behavior in Hybrid Rocket Motors", in Fundamentals of Hybrid Rocket Combustion and Propulsion, edited by M.J. Chiaverini and K.K. Kuo, *Progress in Astronautics and Aeronautics*, Vol. 218, AIAA, Reston, VA, 2006, Chapter 9.

[5] I. Nakagawa, S. Hikone, and T. Suzuki, "A Study on the Regression Rate of Paraffin-based Hybrid Rocket Fuels", 2009. AIAA 2009-4935.

[6] S. Kim, J. Lee, H. Moon, H. Sung, J. Kim, and J. Cho, "Effect of Paraffin-LDPE Blended Fuel in Hybrid Rocket Motor", AIAA 2010-7031.

[7] M.A. Karabeyoglu, B.J. Cantwell, and D. Altman, "Development and Testing of Paraffin-based Hybrid Rocket Fuels", 2001. AIAA 2001-4503.

[8] M.A. Karabeyoglu, D. Altman, and B.J. Cantwell, "Combustion of Liquefying Hybrid Propellants: part 1, General Theory", *Journal of Propulsion and Power*, 18(3): 610-620, 2002.

[9] M.A. Karabeyoglu, and B.J. Cantwell, "Combustion of Liquefying Hybrid Propellants: part 2, Stability of Liquid Films", *Journal of Propulsion and Power*, 18(3): 621-630, 2002.

[10] P.N. Estey, D. Atman, and J. S. McFarlane, "An Evaluation of Scaling Effects for Hybrid Rocket Motors", AIAA-91-2517.

[11] R. D. Swani, and A. Gany, "Analysis and Testing of Similarity and Scale Effects in Hybrid Rocket Motors", Acta Astronautica, 52 (2003): 619-628, 2003.

[12] F. Martin, A. Chapelle, O. Orlandi, and P. Yvart, "Hybrid Propulsion Systems for Future Space Applications", AIAA 2010-6633.

[13] A. Gany, "Similarity and Scaling Effects in Hybrid Rocket Motors", in Fundamentals of Hybrid Rocket Combustion and Propulsion, edited by M.J. Chiaverini and K.K. Kuo, *Progress in Astronautics and Aeronautics*, Vol. 218, AIAA, Reston, VA, 2006, Chapter 12.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Random packing for propellant microstructure simulation

A. Baietta, F. Maggi, L. T. DeLuca
Aerospace Engineering Dept., Politecnico di Milano, Milan, Italy

Keywords: random packing, spheres, multimodal, microstructure, propellant.

Abstract

Composite solid propellants are heterogeneous materials. They are a mixture of different solid phases which arrange randomly during the mixing process. Generally speaking, macroscopic properties of heterogeneous materials are strictly dependent on microstructure. In the specific case, microstructure casts its influence on mechanical, combustion, as well as physical properties. Some modeling approaches include inhomogeneous effects but information on bulk microstructure is needed. Packing codes represent a viable methodology for the production of these datasets. The code implemented in the Space Propulsion Lab of Politecnico di Milano uses the Lubachevsky-Stillinger algorithm and can pack spheres inside cubes with periodic boundaries. The present paper introduces the reader to collision handling for spheres with growing diameters and it offers some preliminary insight in the search for a code stopping criterion.

1 Introduction

Composite solid propellants can be classified as heterogeneous materials. Their composition is similar to a mechanical mixture where micron-sized powders are homogeneously blended and kept together by a polymer matrix. If an average in space is committed on a bulk large enough, chemical and physical properties (such as composition or density) result in uniform values, within the capabilities of a good mixing and curing process. The same happens when mean combustion properties are evaluated (namely, burning rate). Nevertheless, the magnification of an optical microscope performed on a sample cross section reveals a nonuniform and chaotic displacement of ingredients. Akin observation is possible with a close-up of a steady-state burning process; combustion is locally unsteady and burning surface consumption is irregular.

Generally speaking, the buildup of properties at the microscale depends on the microscale and on how the single items arrange each other in space, as in other heterogeneous systems. Speaking of propulsion performance parameters, the nominal propellant composition represents the main driver for theoretical specific impulse. Nonetheless, global experimental properties such as burning rate, metal agglomeration, or elastic modulus are in general controlled by ingredient morphology. Focusing on burning rate, for example, the literature reports that the substitution of coarse oxidizer with a fine cut contributes to the increase of burning rate and pressure ballistic exponent [1].

Homogeneous models were and are sometimes used for global modeling. As an example, the combustion properties of composite solid propellants were early investigated by means of homogeneous combustion models, adapting the basic ideas of combustion models used for double-base propellants. Then, the concept of a different flame structure induced by the microstructure of

the burning material was gradually introduced in 1D models (Summerfield's GDF [2], Hermance [3], Beckstead's BDP [4]). Finally, two- or three-dimensional models that preserved heterogeneous features for combustion were developed for laminate or heterogeneous propellant combustion [5, 6, 7]. In the case of composite metalized propellants, the investigation of a correlation between heterogeneity and metal combustion was the subject of an experimental work by Povinelli [8]. Cohen was the author of a notable model for aluminum agglomeration and introduced the pocket concept as the region where agglomeration is favored [9]. More recently, a statistical approach was used for metal agglomeration model, requiring a model for heterogeneous propellants to be studied ([10, 11, 12]). For the combustion and agglomeration studies mentioned in this introduction, the evolution of model representation toward a more realistic approach required the use of 2D-3D numerical mockups for propellants.

2 Generation of random packs

The generation of data for the analysis of heterogeneous systems is a matter of spatial statistics and simulated packing or advanced experimental techniques. Density of sphere beads were largely investigated in the past by means of empirical methods spanning from tennis balls to calibrated microspheres. The works of Mason and Finney pioneered the spatial statistics of the granular matter with simple optical characterizations of spatial sphere aggregates [13, 14, 15, 16]. McGeary generated an experimental database by mixing spheres of various size and recording the volume occupation [17]. More recently, Aste and co-authors applied X-ray computed tomography (XCT) on packs of different natures (spheres, ellipsoids, random shapes) [18, 19]. Packing codes represent the numerical alternative for the representation of heterogeneous systems. These codes are based on different kinds of algorithms and have the scope to pack items together into a defined control volume. The shape of the packed objects span from spheres to arbitrary shapes, depending on the needs. The result of these type of codes is a representation of a container filled with packed objects which are arranged randomly in space. Typical code output consists of a list of particles with the respective collocation in the space. One family of packing codes appeals to non-dynamic algorithms where particles are placed in the space by a specific procedure. Bandera used an algorithm that placed spheres one-by-one, checking the presence of overlaps [20]. Yugong [21] and then Maggi [22] applied a collocation strategy based on the minimization of a target function. Out of the family of dynamic packing algorithm, a drop-down technique was adopted by Webb and Davies [23], a time-driven algorithm with inflatable particles and viscosity was applied by Rashkovskii. On the other hand, an event-driven packing algorithm was introduced by Lubachevsky and then used in several codes (e.g. Rocpack [24]). The results on packing inner structure depend on the the choice of the packing algorithm. Some methodologies, for example, do not reproduce correctly the reciprocal spatial arrangement that a group of particles would have in the reality. Only recently, the availability of tomography data made possible the comparison of simulations with numerically representative sets of data. A paper by Maggi et al. for example proposes an experimental validation for a code based on Lubachevsky algorithm, also comparing to the statistics of a drop-down code [25].

3 The packing code

The code here described is based on the algorithm proposed in 1990 by B. D. Lubachevsky, who worked first on billiard simulations [26] and then on random disk packing problems [27]. The idea is to have a number of spheres moving in a closed domain like a billiard simulation, but unlike the previous one, the diameter of the spheres grows linearly in time. Since the volume of the domain is finite, there will be a time when the spheres cannot grow any more, obtaining a jammed pack. The algorithm, inspired by the current advances in programming techniques for event-driven simulations in molecular dynamics, proposed a new approach to the problem; while earlier algorithms used an intrinsi-

cally sequential model in which the spheres were added one by one to an initial seed structure, Lubachevsky's algorithms was intrinsically concurrent and was closer in spirit to the procedures used experimentally to create amorphous solids.

As stated above, a time dependent simulation has to be run in order to obtain a single pack of spheres. A straightforward approach would discretize the status evolution in time and check if two particles were overlapping at the end of each time step. In this case, a collision should have occurred some time before, when the items got in touch, and appropriate measures should be implemented to deal with it. This approach, called time-driven molecular dynamics (TDMD), has many problems (as discussed in [28]) and a high computational cost. Lubachevsky proposed to use an event-driven approach (event-driven molecular dynamics, EDMD), where the simulation advancement is clocked by a sequence of events. EDMD algorithms are very efficient because, if there aren't external forces, every particle moves along a straight line between two collisions, making the identification of the exact time for the next event easy. Since a particle is more likely to collide with another one that is in its close vicinity than one that is far away, we can also split the domain in cells and consider only collisions between particles in neighboring cells.

Three types of events can be classified: the collision between two spheres, the transition across a domain boundary (for periodic boundary conditions) and the transition while passing from one cell to another.

4 Collisions

Collisions could be considered the heart of the Lubachevsky's algorithm since they are the driving event of the simulation. Cell or boundary transitions are only secondary events, they are introduced only in order to minimize computational time (cell transition) or define a finite domain (boundary transition)[1]. A collision, and in general any event, must be pointwise-in-time in order to fit within the general framework of the algorithm. This ensures no overlapping spheres are present at any time. In our case the classical elastic collision does not guarantee this property because each particle is growing and there is an outward surface velocity relative to the centroid. Depending on the diameter growth rate and the pre-impact centroid velocity, it may be necessary to add extra impulses in order to increment the rebound velocity and preserve from particle overlapping. However, extra impulses generate an increase of kinetic energy leading to related issues when sphere velocities are too high relevant to computer hardware precision and to the "control" of the simulation. If particle velocity increases too much, loss of precision due to computer finite algebra could cause a collision to be fully missed or inaccurately predicted, finding a collision time featured by overlapping spheres. The second issue does not jeopardize the code integrity but it affects the sensibility of the packing fraction from the diameter grow rate. Typically, diameter grow rate is used to drive the change of final pack density (small value cause more ordered packing [27]). In reality, the ratio between grow rate and particle mean speed appears to be more important. If the ratio becomes small, little growth occurs between collisions and it can become impossible to generate more ordered and denser packs. Kinetic energy increase caused by diameter growth is not important.

This chapter will focus on how different types of collision managements can cause a variation in the final result in terms of packing fraction. These changes are linked to different approaches used for the containment of kinetic energy.

4.1 Classical elastic collision

In a classical elastic collision between particles of the same diameter, the velocities of the spheres centers can be divided into component parallel and traverse to the line of centers. The traverse components are unchanged by collision, whereas the parallel components are exchanged. More

[1] Despite our code applies to periodic boundary conditions, it's important to note that in a real situation the necessary presence of a container causes local rearrangements of the spheres. This could influence the final value of packing fraction [17].

complex model ([29]) takes into account the possible difference in size (for multimodal analysis), although interaction between bigger and smaller sphere is something we have to study yet.

The relation between velocity before and after a collision can be expressed by the following equation,

$$m_1(\mathbf{v}_1 - \mathbf{v}_1^{(0)}) = \mathbf{J} \tag{1}$$

$$m_2(\mathbf{v}_2 - \mathbf{v}_2^{(0)}) = -\mathbf{J} \tag{2}$$

where $\mathbf{J}$ is the impulsive force exerted on particle 1 which also acts on particle 2 as the reaction force. The subscript 1 and 2 refer to the two spheres, the superscript (0) represents the values before the collision and $\mathbf{v}$ is the velocity of the sphere center. m is the mass. The unknown variables are the impulsive force $\mathbf{J}$ and the post-collisional velocities $\mathbf{v}$. We define the relative velocities between particle centers before and after collision,

$$\mathbf{G}^{(0)} = \mathbf{v}_1^{(0)} - \mathbf{v}_2^{(0)} \tag{3}$$

$$\mathbf{G} = \mathbf{v}_1 - \mathbf{v}_2 \tag{4}$$

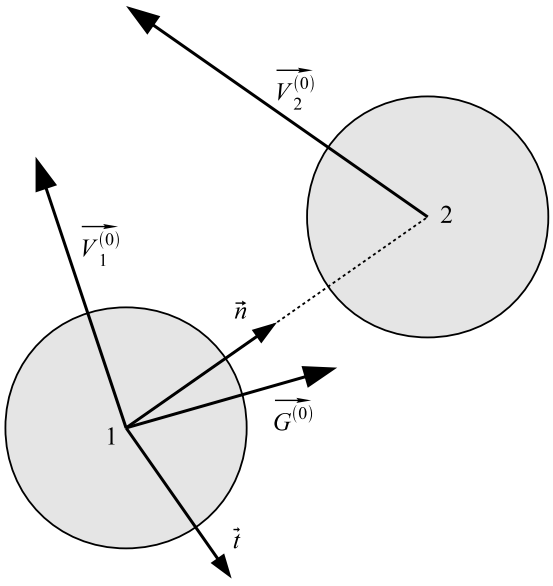


Figure 1: Relative motion of two spheres.

The post-collisional relative velocity $\mathbf{G}$ is related to the pre-collisional velocity $\mathbf{G}^{(0)}$ through the restitution coefficient e

$$\mathbf{n} \cdot \mathbf{G} = -e\left(\mathbf{n} \cdot \mathbf{G}^{(0)}\right) \tag{5}$$

The restitution coefficient e is assumed to be constant. A value of one allows for the conservation of the kinetic energy through the collision. Equations 1 and 2, with the definitions 3 and 4, become

$$\mathbf{G} = \mathbf{G}^{(0)} + \frac{m_1 + m_2}{m_1 m_2}\mathbf{J} \tag{6}$$

Multiplying by the unit vector $\mathbf{n}$ and substituting Eq. 5 into the left side of the above equation, we obtain the normal component of the impulsive force $\mathbf{J}_n$

$$\mathbf{J}_n = -\frac{m_1 m_2}{m_1 + m_2}(1 + e)\left(\mathbf{n} \cdot \mathbf{G}^{(0)}\right) \tag{7}$$

Assuming that the particles slide, the tangential component of the impulsive force $\mathbf{J}_t$ is

$$\mathbf{J}_t = \mu \mathbf{J}_n \tag{8}$$

where μ is the friction coefficient. The post-collisional velocities are

$$\mathbf{v}_1 = \mathbf{v}_1^{(0)} - (\mathbf{n} - \mu\mathbf{t})\left(\mathbf{n} \cdot \mathbf{G}^{(0)}\right)(1 + e)\frac{m_2}{m_1 + m_2} \tag{9}$$

$$\mathbf{v}_2 = \mathbf{v}_2^{(0)} + (\mathbf{n} - \mu\mathbf{t})\left(\mathbf{n} \cdot \mathbf{G}^{(0)}\right)(1 + e)\frac{m_1}{m_1 + m_2} \tag{10}$$

In this paper we use $\mu = 0$ and $e = 1$ unless specified.

4.2 Collision with diameter growth

In order to ensure a pointwise-in-time event, additional impulses must be added to account for diameter growth. In this case, after every collision a check is done on separation speed of the centers and, if necessary, an extra impulse is added to prevent overlapping. This approach however is not completely correct and it may cause insidious errors, depending on the implementation of the check on post-collisional velocity and on extra impulse addition. For a classical collision between sphere of constant diameter, necessary and sufficient condition for an impact is converging trajectories. However converging paths are only a sufficient condition if the diameter grows. A collision can occur even with diverging trajectories if the separating speed is less than the semi-sum of the diameter growth rate (Fig. 2), we dubbed this case *back impact*.

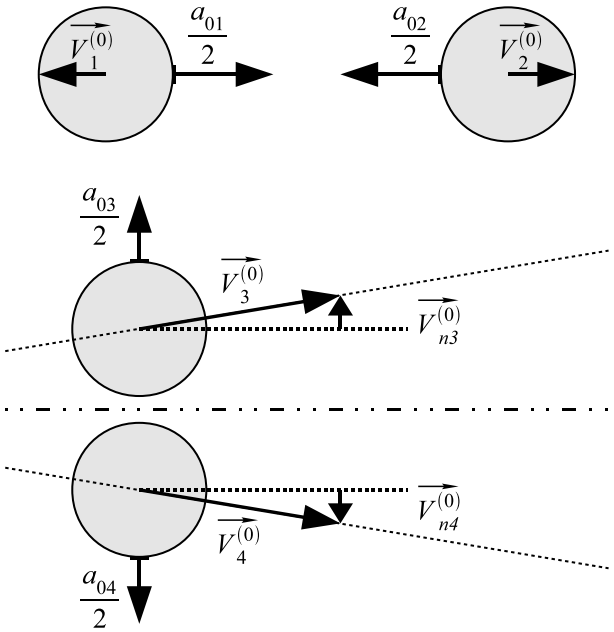


Figure 2: Collision with divergent trajectories, dubbed back impact.

It's important to note that if the diameter is constant and there is a collision, the pre-collisional velocities are on a converging path, while the post-collisional trajectories diverge. If the previous discussed formulas (Eq. 9 and Eq. 10) are applied to spheres of growing diameters and there is a back impact, post-collisional velocity on a converging trajectories are obtained and overlapping is inevitable. The original Lubachevsky's algorithm [27] circumvents the problem by adding always an impulse to each of the two spheres equal to the diameter growth rate, enough to ensure the absence of back impacts. This approach is too conservative and adds more energy than required; we propose here a different solution.

As a first step, we should define if the collision is a back impact, which happens if

$$\mathbf{n} \cdot \mathbf{G}^{(0)} \leq 0 \tag{11}$$

In this case, it is necessary to guarantee that the sphere centers are moving away from each other with a minimum velocity v_{min}, where

$$v_{min} > \frac{1}{2}\left(a_{01} + a_{02}\right) \tag{12}$$

a_{01} and a_{02} are the diameter growth rate of the two particles. The velocity to be added is

$$I = v_{min} + \mathbf{n} \cdot \mathbf{G}^{(0)} \tag{13}$$

The post-collisional velocity are

$$\mathbf{v}_1 = \mathbf{v}_1^{(0)} - I\frac{m_2}{m_1 + m_2}\mathbf{n} \tag{14}$$

$$\mathbf{v}_2 = \mathbf{v}_2^{(0)} + I\frac{m_1}{m_1 + m_2}\mathbf{n} \tag{15}$$

If Eq. 11 is false, the classical collision takes place and Eq. 9 and Eq. 10 are to be applied obtaining $\mathbf{v}_1^*$ and $\mathbf{v}_2^*$. Final velocity must guarantee v_{min}.

$$I = v_{min} - e\left(\mathbf{n} \cdot \mathbf{G}^{(0)}\right) \tag{16}$$

If $I < 0$ then $\mathbf{v}_1^*$ and $\mathbf{v}_2^*$ are the final post-collisional velocity, else it is necessary to increment the rebound speed.

$$\mathbf{v}_1 = \mathbf{v}_1^* - I\frac{m_2}{m_1 + m_2}\mathbf{n} \tag{17}$$

$$\mathbf{v}_2 = \mathbf{v}_2^* + I\frac{m_1}{m_1 + m_2}\mathbf{n} \tag{18}$$

5 More on the effect of the diameter growth rate

In our previous paper presented at EUCASS 2011, the influence of particle growth rate on packing fraction was investigated using a single-size pack. A different behavior between our development and a reference code [25] was attributed to a different collisional management. Since then, more detailed studies have been conducted on collisions and the results are presented here.

The pack is composed by 3000 spheres with the same diameter. The diameter growth rate a_0 is varied from 0.01 to 100 on a logarithmic increment. The stopping criteria is based on the variation of packing fraction across a computational step. A step is defined as $10 \cdot N$ iteration, where N is the number of particles (in this case a step consists of 30000 iteration). At the n-th pass the quantity

$$\epsilon = \frac{f_n - f_{n-1}}{f_n} \tag{19}$$

is computed and the code is stopped if $\epsilon < 10^{-7}$. Four different combination of v_{min} (minimum separation velocity between sphere centers after a collision) and e (restitution coefficient) are tested as reported in table 1.

Code	v_{min}	e
A	$2.00 \cdot (a_0)$	1.0
B	$1.02 \cdot (a_0)$	1.0
C	$1.02 \cdot (a_0)$	0.5
D	$1.02 \cdot (a_0)$	0.0

Table 1: Different combination of v_{min} and e used in the collision function.

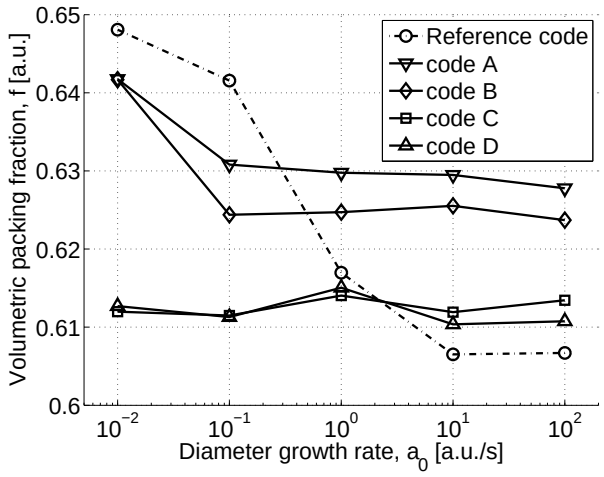


Figure 3: Dependence of volumetric packing fraction on diameter growth rate in single-size packs.

The results are presented in Fig. 3. Each curve in the plot is the average of three packing procedures with different initial random seed (positions and velocities are sampled from a random distribution of number in the range -1 to $+1$). The results are again well centered in the experimental range ($0.60 - 0.64$, as of [17]). The results obtained with code B are interesting since the volumetric packing fraction tends to 0.625, a value found by McGeary for the experimental packing of steel spheres. However, whereas the reference code can simulate a large span of packing fraction just changing the diameter growth rate, our code does not have large sensitivity on it. This was previously attributed to a different collision handling technique (EUCASS 2011), but it seems this is not the case. A different computational step for ϵ was then considered, since the reference code defines a step as N collisions instead of $10 \cdot N$ iterations, though finding results similar to Fig. 3. Conversely, the difference of behavior between the present code and the reference implementation did not find any explanation yet and will be further addressed in the future. Anyway, it should be noted that code behavior mismatch does represent only a peculiarity of the specific implementation and does not impair the quality of the final pack which effectively represents the code output.

6 Bimodal packs: searching for a stop criterion

As for every iterative procedure, the identification of a stopping criterion is fundamental to the packing simulation, since the algorithm tends asymptotically to displace particles in a close-pack ideal arrangement. Two types of stopping criteria are usually adopted. The first, the most common one, is based on the variation of a selected parameter across an iteration, compared to a chosen tolerance. This type is always used in order to limit the length of the simulation since it is pretty much a waste of time to continue if the parameter of interest does not change significantly any more (a more aggressive approach can be to limit the total number of iterations). A second type of stopping criterion targets a defined packing fraction. This method is more attuned to the nature of the simulation, and it builds a bridge between the problem physics and the simulation model. This type of criterion generally marks a separation between what an acceptable physical solution is and what goes beyond the experimental evidence.

Lubachevsky's algorithm is more a dynamic simulation than an iterative procedure, but it can be regarded as such if the packing fraction is considered. At every computational step[2] the

[2]The reader should note the use of "computational step" instead of "iteration". As discussed, the time of the simulation and the packing fraction could be the same across a single iteration but it's improbable that the same

spheres grow and the packing fraction increases toward a desired value. A simple stop criterion is already implemented in the code. If the variation of volumetric packing fraction ϵ across a computational step is less than a tolerance, the code stops. Different values of ϵ from 10^{-7} to 10^{-4} were considered. $\epsilon = 10^{-}7$ was chosen since it is a good compromise between the time it takes to run a simulation and the final packing fraction reached. Smaller values for ϵ do not chance packing fraction significantly but they greatly increase running time.

Before proceeding further it is necessary to say that there are two ways in which our code can be used. The first and most straightforward way is to define the desired packing fraction. This is the case when you want to generate a model for an existing propellant, the distribution of diameters and the final packing fraction are somehow known. In this situation, the achievement of the desired packing fraction stops the code and other stopping criteria are not needed. The second, and perhaps more interesting way in which the code can be used, is to compute the packing fraction that can be obtained from a particular distribution of diameters. In the latter case the definition of a suitable stopping criterion is essential to obtain a realistic result. In this section we report some observations made during simulation with bimodal packs.

The problem of estimating the final volume packing fraction of a known sphere distribution lies in the pack dependence on the manufacturing method. How and when spheres are added, how much and with which frequency the container is mixed or shaken, are parameters that influence the resulting pack. McGeary reports interesting results for binary packings. For bimodal powder distributions using the same coarse spheres, the smaller the diameter of fine spheres, the greater the packing fraction is going to be, since the finer material fills better the interstices left by the coarse one. However this result can be obtained only by one experimental technique [17]. First the coarse component is shaken by itself until minimum volume is obtained. Only at this point the fine component is added and the mixture is shaken until minimum volume is again obtained. Moreover vigorous shaking will generally produce a uniform mixture. Gentle agitation, on the other hand, can promote the segregation of particles of different sizes. Depending on the acceleration and frequency of the external shaking is possible to obtain big particle on the bottom or at the top (Brazil Nuts effect [30]) of the smaller ones [31, 32].

A series of simulations on bimodal packs of 20000 spheres was conducted. In every packs the biggest spheres have a diameter growth rate of 0.2 while the coarse-to-fine diameter ratio is 3.4 : 1. Volume fraction of fine component was varied across different values and the code was stopped when reaching the corresponding experimental packing fraction obtained by McGeary [17]. The variation of packing fraction ϵ when the code was stopped is reported in Fig. 4.

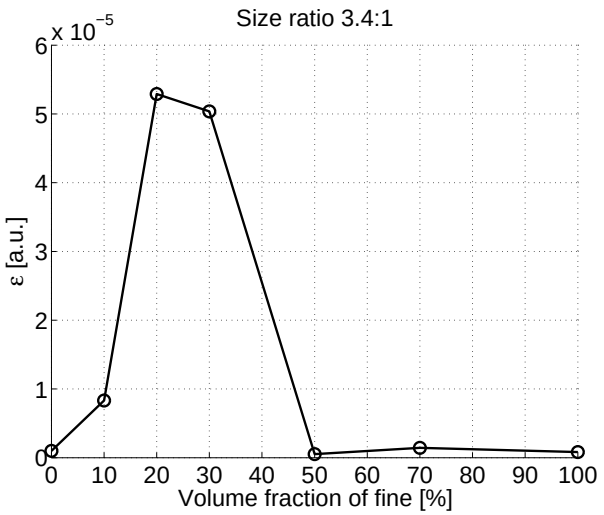


Figure 4: Value of ϵ for which McGeary's packing fraction is reached for different volume fractions of fine particles. Bimodal packs of 20000 spheres, the bigger ones have a growth rate of 0.2.

A peak well centered around 27% of fine fraction results from code runs, where McGeary proposes the ideal maximum attainable packing fraction of 0.86 for general bimodal packs. The stopping criterion based on $\epsilon < 10^{-7}$ seems good enough for monomodal pack (0 and 100% fraction of fine) and for packs with a large number

happens across a step.

of smaller spheres. However different behavior stems for the interval $10 - 40\%$ which can be explained if the nature of the packs is considered. Till 50% fraction of fine, the coarse component occupies most of the domain volume, while the fine one fills the interstices increasing the packing fraction. However, when more volume is occupied by the smaller spheres, the nature of the pack change. A small number of bigger spheres are scattered in a matrix of fine particles and the global pack presumably behaves as a monomodal one.

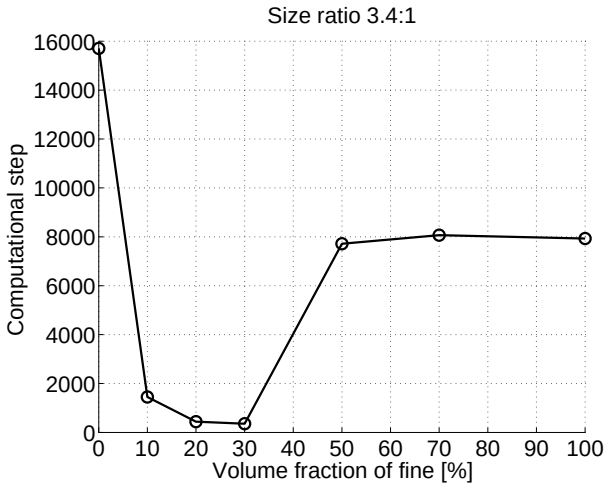


Figure 5: Number of computational step needed to reach McGeary's packing fraction for different volume fractions of fine particles. Bimodal packs of 20000 spheres, the bigger ones have a growth rate of 0.2.

The number of computational step required to reach the packing fraction proposed by McGeary are plotted for different mixture of fine and coarse particles in Fig. 5. Again two patterns can be recognized. For low percentages of fine $(10 - 30\%)$ a lower number of steps are necessary to reach the experimental data. One would expect that a more vigorous shaking is needed in order to achieve higher values of packing fraction. Nevertheless, from a computational point of view, less steps are needed. At this point only speculations are possible since McGeary does not report detailed information on experimental pack generation like time and intensity of vibrations. About the same number of iteration are needed for mixtures of more than 50% of fine, this again confirm the interpretation referring to the monomodal behavior.

7 Conclusion

A packing code based on Lubachevsky's algorithm is under development at the Politecnico di Milano. The code currently was tested only with packs of single-size spheres and with bimodal packs in a cubic domain with periodic boundary condition. However, data structure is designed as is to handle polydisperse realistic powder data.

Packs of 3000 single-size spheres were produced with different diameter growth rate and the resulting packing fraction confronted with results from a reference code. The difference from developed to reference code was previously attributed to different collision handling technique, since our earlier version added more than necessary rebound velocity. This however does not seem to be the case since different collision strategies were tested with similar results. Whereas the reference code can simulate a large span of packing fractions just changing the particle growth rate, the current version of the code does not have large sensitivity on it.

Bimodal packs with a size ratio of $3.4 : 1$ were also investigated and the data obtained were confronted with McGeary's experimental results. Two different patterns seems to emerge, for high values of fine volume fraction ($> 50\%$) the packs seem to behave as monomodal, while for lower values a bimodal behavior seems to prevail. In this second case the simple stop criterion based on ϵ cause an overestimation of packing fraction. Further simulation with higher size ratio will be carried out in order to recognize some patterns and design a new stopping criterion.

Some innovative features, despite being already implemented in the code, have not yet been tested such as the presence of particles with different densities. The target of the development can now be focused on code efficiency implementing less expensive forecasting algorithm for collisions (using octrees) or introducing a parallelization of the code through a shared memory paradigm (namely, OpenMP [33]).

References

[1] Kubota N., *Propellants And Explosives: Thermochemical Aspects Of Combustion*, 2nd ed., Wiley-VCH, 2007.

[2] Steinz J. A., Stang P. L. and Summerfield M., "The burning mechanism of ammonium perchlorate-based composite solid propellants", Technical Report 830, Department of Aerospace and Mechanical Sciences, Princeton University, 1969.

[3] Hermance C. E., "A model of composite propellant combustion including surface heterogeneity and heat generation", *AIAA Journal*, Vol. 4, No. 9, 1966, pp. 1629-1637.

[4] Beckstead M. W., Derr R. L. and Price C. F., "A model of composite solid-propellant combustion based on multiple flames", *AIAA Journal*, Vol. 8, No. 12, 1970, pp. 2200-2207.

[5] Jackson T. L. and Buckmaster J., "Heterogeneous propellant combustion", *AIAA Journal*, Vol. 40, No. 6, 2002, pp. 1122-1130.

[6] Maggi F., Galfetti L., Miccio F. and DeLuca L. T., "Flame structure simulation of non-aluminized composite propellants", *57th International Astronautical Conference*, number IAC-06-C4.P.3.03, pages 1-14, Valencia, Spain, October 2006, IAF.

[7] Genevieve P. and Brewster M. Q., "Modeling combustion in laminate propellants with fine ammonium perchlorate", *40th AIAA Aerospace Sciences Meeting and Exhibit*, number 2002-0781, Reno, NV, 2002.

[8] Povinelli L. A., "Effect on oxidizer particle size on additive agglomeration", TN D-1438, NASA, Lewis Research Center, Cleveland, OH, 1962.

[9] Cohen N. S., "A pocket model for aluminum agglomeration in composite propellants", *AIAA Journal*, Vol. 21 No. 5, 1983, pp. 720-725.

[10] Rashkovskii S. A., "Role of the structure of heterogeneous condensed mixtures in the formation of agglomerates", *Combustion, Explosion and Shock Waves*, Vol. 38, No. 4, 2002, pp. 435-445.

[11] Gallier S., "A stochastic pocket model for aluminum agglomeration in solid propellants", *Propellant, Explosives, Pyrotechnics*, Vol. 34, No. 2, 2009, pp. 97-105.

[12] Maggi F., Bandera A., Galfetti L., DeLuca L. T. and Jackson T. L., "Efficient solid rocket propulsion for access to space", *Acta Astronautica*, Vol. 66, No. 11-12, 2010, pp. 1563-1573.

[13] Bernal J. D., Cherry I. A., Finney J. L. and Knight K. R., "An optical machine for measuring sphere coordinates in random packings", *Journal of Physics E: Scientific Instruments*, Vol. 3, No. 5, 1970, pp. 388-390.

[14] Bernal J. D. and Mason J., "Co-ordination of randomly packed spheres", *Nature*, Vol. 188, 1960, pp. 910-911.

[15] Mason G., "Radial distribution functions from small packings of spheres", *Nature*, Vol. 217, 1968, pp. 733-735.

[16] Mason G. and Clark W., "Fine structure in the radial distribution function from a random packing of spheres", *Nature*, Vol. 211, 1966, pp. 957.

[17] McGeary R. K., "Mechanical packing of spherical particles", *Journal of the American Ceramic Society*, Vol. 44, No. 10, 1961, pp. 513-522.

[18] Aste T., Saadatfar M. and Senden T. J., "Geometrical structure of disordered sphere packings", *Physical Review E*, Vol. 71, No. 061302, 2005, pp. 1-15.

[19] Aste T., Saadatfar M., Sakellariou A. and Senden T. J., "Investigating the geometrical structure of disordered sphere packings", *Physica A*, Vol. 339, No. 1-2, 2004, pp. 16-23.

[20] Bandera A., "Combustion of metalized solid rocket propellants and motor performance", PhD thesis, Politecnico di Milano, 2009.

[21] Yugong W., Zhigang F. and Yuzhu L., "Bulk and interior packing densities of random close packing of hard spheres", *Journal of Materials Science*, Vol. 38, No. 9, 2003, pp. 2019-2025.

[22] Maggi F., "Heterogeneity effects on composite solid rocket propellants", PhD thesis, Politecnico di Milano, 2006.

[23] Webb M. D. and Davis I. L., "Random particle packing with large particle size variations using reduced-dimension algorithms", *Powder Technology*, Vol. 167, No. 1, 2006, pp. 10-19.

[24] Knott G. M., Jackson T. L. and Buckmaster J., "Random packing of heterogeneous propellants", *AIAA Journal*, Vol. 39, No. 4, 2001, pp. 678-686.

[25] Maggi F., Stafford S. and Jackson T. L., "Nature of packs used in propellant modeling", *Physical Review E*, Vol. 77, No. 046107, 2008, pp. 1-17.

[26] Lubachevsky B. D., "How to simulate billiards and similar system", *Journal of Computational Physics*, Vol. 94, No. 2, 1991, pp. 255-283.

[27] Lubachevsky B. D. and Stillinger F. H., "Geometric properties of random disk packings", *Journal of Statistical Physics*, Vol. 60, No. 5-6, 1990, pp. 561-583.

[28] Sigurgeirsson H., Stuart A. and Wan W., "Algorithms for particle-field simulations whit collisions", *Journal of Computational Physics*, Vol. 172, No. 2, 2001, pp. 766-807.

[29] Crowe C., Sommerfeld M., Tsuji Y., *Multiphase flows with droplets and particles*, CRC Press, 1998, Chap. 5.

[30] Rosato A., Strandburg K. J., Prinz F., Swendsen R. H., "Why the Brazil nuts are on top: Size segregation of particular matter by shaking", *Physical Review Letters*, Vol. 58, No. 10, 1987, pp. 1038-1040.

[31] Breu A. P. J., Ensner H.-M., Kruelle C. A., Rehberg I., "Reversing the Brazil-nut effect: Competition between percolation and condensation", *Physical Review Letters*, Vol. 90, No. 1, 2003, pp. 014302.

[32] Aste T., Weaire D. L., *The Pursuit of Perfect Packing*, 2nd ed., Taylor & Francis, 2008, Chap. 6.

[33] OpenMP Application Program Interface, version 2.5, 2005.

NEW ROBUST DESIGN GUIDELINE FOR IMPERFECTION SENSITIVE COMPOSITE LAUNCHER STRUCTURES

Richard Degenhardt [1,2]
[1] *DLR, Institute of Composite Structures and Adaptive Systems, Lilienthalplatz 7, 38108 Braunschweig, Germany*
e-mail: richard.degenhardt@dlr.de

[2] *PFH, Private University of Applied Sciences Göttingen, Composite Engineering Campus Stade, Germany*

Keywords: *Composits, Launcher structures, Design, Buckling, Stability, Imperfection Sensitivity*

Abstract

Up to today there are no design guidelines which allow designing imperfection sensitive CFRP structures in an efficient way. The upcoming EU project DESICOS (New Robust DESIgn Guideline for Imperfection Sensitive COmposite Launcher Structures) contributes to this aim by a new design procedure exploiting the worst imperfection approach efficiently by implementation of the Single Perturbation Load Approach. Preliminary studies already demonstrated high potential. This paper deals with the objectives of the DESICOS project, describes the line of actions of the new approach, and specifies the theoretical and experimental work to be carried out in order to meet the objectives.

1 Introduction

The Space industry demand for lighter and cheaper launcher transport systems. The upcoming EU project DESICOS (New Robust DESIgn Guideline for Imperfection Sensitive COmposite Launcher Structures), which will start in 2012, contributes to these aims by a new design procedure for imperfection sensitive composite launcher structures, exploiting the worst imperfection approach efficiently by implementation of the Single Perturbation Load Approach [1]. Currently, imperfection sensitive shell structures prone to buckling are commonly designed according the NASA SP 8007 guideline using the conservative lower bound curve. The guideline dates from 1968, and the structural behaviour of composite material is not considered appropriately, in particular since buckling load and imperfection sensitivity of shells made from such materials substantially depend on the lay-up design. This is not considered in the NASA SP 8007, which allows designing only so called "black metal" structures. Here is a high need for a new precise and efficient design approach for imperfection sensitive composite structures which allows significant reduction of structural weight and design cost. For most relevant architectures of cylindrical and conical launcher structures (monolithic, sandwich - without and with holes) DESICOS will investigate a combined methodology from the Single Perturbation Load Approach and a Specific Stochastic Approach which guarantees an effective and robust design. A recent investigation demonstrated, that an axially loaded unstiffened cylinder, which is disturbed by a large enough single perturbation

load, is leading directly to the design buckling load 45% higher compared with the respective NASA SP 8007 design [2]. Within DESICOS the new methods will be further developed, validated by tests and summarized in a handbook for the design of imperfection sensitive composite structures. The potential will be demonstrated within different industrially driven use cases. This paper deals with the objectives of the DESICOS project, describes the line of actions of the new approach, and specifies the theoretical and experimental work to be carried out in order to meet the objectives.

2 State of the art

2.1 Imperfection sensitivity

In Fig. 1 taken from [3], knock-down factors are shown for axially compressed cylindrical shells depending on the slenderness. The results of tests are presented by dots and show the large scatter. The knock-down factors decrease with increasing slenderness. The discrepancy between test and classical buckling theory shown in Fig. 1 has stimulated scientists and engineers on this subject during the past 50 years. These works focused on postbuckling, load-deflection behaviour of perfect shells, various boundary conditions and its effect on bifurcation buckling, empirically derived design formulas and initial geometric imperfections. Koiter was the first to develop a theory which provides the most rational explanation of the large discrepancy between test and theory for the buckling of axially compressed cylindrical shells. In his doctoral thesis published in 1945 Koiter revealed the extreme sensitivity of buckling loads to initial geometric imperfect-tions. His work received little attention until the early 1960's, because the thesis was written in Dutch. An English translation by Riks was published 1967 in [4].

Based on large test series in the 1950s and 60s the determination of lower bounds led to design regulations like NASA SP-8007 [1], but the given knock-down factors are very conservative. To improve the ratio of weight and stiffness and to reduce time and cost, numerical simulations could be used during the design process. The consideration of imperfections in the numerical simulation is essential for safe constructions. Usually, these imperfections are unknown in the design phase, thus pattern and amplitude have to be assumed.

In general, one can distinguish between loading imperfections and geometric imperfections. Both kinds of imperfections have a significant influence on the buckling behaviour and their state of the art is described in the following.

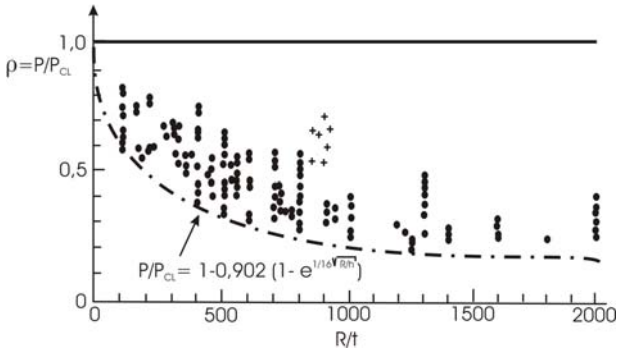


Fig. 1 Distribution of test data for cylinders subjected to axial compression [1]

Loading imperfections mean any deviations from perfect uniformly distributed loading, independent of the reason of the perturbation. Geier et al. tested composite cylindrical shells with different laminate designs [5], and they applied thin metal plates locally between test shell and supporting structure to perturb the applied loads and performed the so-called shim tests [6]. Later, numerical investigations were performed and compared to the test results; the importance was verified [7]. The need to investigate loading imperfections for practical use was shown for instance by Albus et al. [8] by the example of Ariane 5.

Geometric imperfections mean any deviations from the ideal shape of the shell structure. They are often regarded the main source for the differences between computed and tested buckling loads. Winterstetter et al. [9] suggest three approaches for the numerical simulation of geometrically imperfect shell structures: "realistic", "worst" and "stimulating" geometric imperfections. Stimulating geometric

imperfections like welded seams are local perturbations which "stimulate" the characteristic physical shell buckling behaviour [10]. "Worst" geometric imperfections have a mathematically determined worst possible imperfection pattern like the single buckle [11]. "Realistic" geometric imperfections are determined by measurement after fabrication and installation. This concept of measured imperfections is initiated and intensively promoted by Arbocz [12]; a large number of test data is needed, which has to be classified and analysed in an imperfection data bank. Within the study presented in this paper, real geometric imperfections measured at test shells are taken into account.

Hühne et al. [1] showed that for both, loading imperfections and geometric imperfections the loss of stability is initiated by a local single buckle. Therefore unification of imperfection sensitivity is allowed; systems sensitive to geometric imperfections are also sensitive to loading imperfections. Single buckles are realistic, stimulating and worst geometric imperfections.

Using laminated composites, the structural behaviour can be tailored by variation of fibre orientations, layer thicknesses and stacking sequence. Fixing the layer thicknesses and the number of layers, Zimmermann [13] demonstrated numerically and experimentally that variation of fibre orientations affects the buckling load remarkably. The tests showed that fibre orientations can also significantly influence the sensitivity of cylindrical shells to imperfections. Meyer-Piening et al. [14] reported about testing of composite cylinders, including combined axial and torsion loading, and compared the results with computations.

Designing a cylinder appropriately is very important because changing only the lay-up or stacking sequence for the same cylinder geometry and thickness can lead to significantly different buckling loads. Zimmermann [15] designed different cylinders with extreme structural behaviour. Hühne [1] selected some of them and performed additional studies. Within a DLR-ESA study one of these cylinder designs, which is most imperfection sensitive, was manufactured 10 times and tested. It allowed a comparison with already available results and enlarged the data base [2].

The buckling process, even under static loading, is highly dynamic. In order to measure the full scale cylinder deformations of that process during testing, the Institute of Composite Structures and Adaptive Systems of DLR has developed a 360° measurement method based on the ARAMIS concept, a high speed optical grating system developed by GOM GmbH (`Gesellschaft für Optische Messtechnik´) [16]. The development includes four self-sustaining fast systems which can measure the 3-dimensional deformation field of an object by applying fast digital cameras with a speed of max. 1,000 images/sec.

2.2 Single-Pertubation-Load Approach

Hühne [1] proposed an approach based on a single buckle as the worst imperfection mode leading directly to the load carrying capacity of a cylinder (cf. Fig. 2). Fig. 3 shows that the buckling load under a pertubation load larger than a minimum value P_1 is almost constant. A further increase of the pertubation load has no significant change on the buckling any more. The buckling load F_1 (cf. Fig. 3) corresponds to the design buckling load. This concept promises therefore to improve the knock-down factors and allows designing any CFRP cylinder by means of one calculation under axial compression and a single-perturbation load. Within a DLR-ESA study, this approach was confirmed analytically and experimentally [2]. However, there is still a need for further studies.

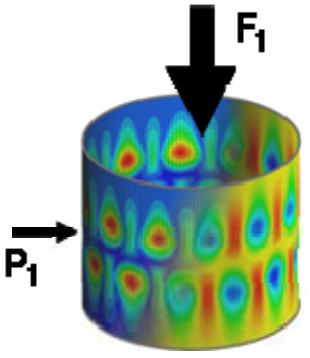


Fig. 2 Pertubation load mechanism

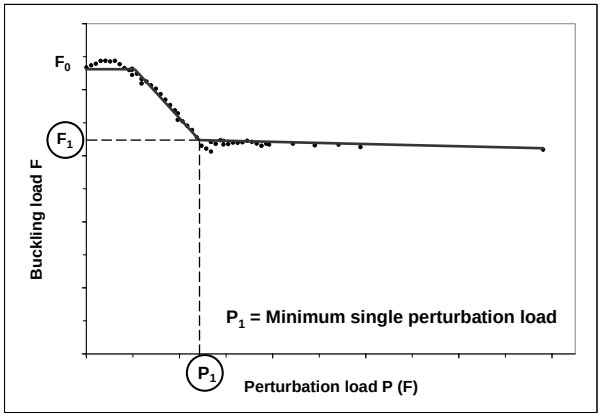


Fig. 3 Single perturbation load approach (SPLA)

2.3 Probabilistic research

In general, tests or analysis results are sensitive to certain parameters as boundary conditions or imperfections. Probabilistic methods are a possibility to assess the quality of results. The stochastic simulation with Monte Carlo (e.g. [17]) allows the statistical description of the sensitivity of the structural behaviour. It starts with a nominal model and makes copies of it whereas certain parameters are varied randomly. The random numbers, however, follow a given statistical distribution. Each generated model is slightly different, as in reality.

Nowadays, probabilistic simulations found the way into all industrial fields. In automotive engineering it is successfully applied in crash or safety (e.g. [18]). Klein et al. [20] applied the probabilistic approach to structural factors of safety in aerospace. Sickinger and Herbeck [21] investigated the deployable CFRP booms for a solar propelled sail of a spacecraft using the Monte Carlo method.

Velds [22] performed deterministic and probabilistic investigations on isotropic cylindrical shells applying finite element buckling analyses and showed the possibility to improve the knock-down factors. In addition, a set-up of a probabilistic design approach has still a lack of knowledge due to the incomplete base of material properties, geometric deviations, etc..

Arbocz and Hilburger [23] published a probability-based analysis method for predicting buckling loads of axially compressed composite cylinders. This method, which is based on the Monte Carlo method and first-order second-moment method, can be used to form the basis for a design approach and shell analysis that includes the effects of initial geometric imperfections on the buckling load of the shell. This promising approach yields less conservative knock-down factors than those used presently by industry.

2.4 Specific Stochastic Approach

Fig. 4 shows the variation (gray shaded band) of the buckling load resulting from its sensitivity to the scatter of the non-traditional imperfections (e.g. thickness variations). It demonstrates the need to cover this by the development of an additional knock-down factor ρ_2 in combination to the knock-down factor ρ_1 from SPLA.

An efficient design is feasible, if knowledge about possibly occurring imperfections exists and if this knowledge is used within the design process. Whereas the traditional imperfections are dealt with the SPLA, the non-traditional ones are taken into account by probabilistic methods, which enable the prediction of a stochastic distribution of buckling loads. Once the distribution of buckling loads is known, a lower bound can be defined by choosing a level of reliability. Degenhardt et al. [2] found less conservative knockdown factors than through the NASA-SP 8007 lower bound, by executing probabilistic analyses with non-traditional imperfections. This concept builds a decent foundation for extension in DESICOS.

The project work of DESICOS for the stochastic approach consists in checking which structural parameters substantially influence the buckling load and defining realistic limits for their deviations from the nominal values, in varying them within the limits and performing buckling load computations for these variations. The results are evaluated stochastically in order to define a guideline for the lower limits of the buckling loads within a certain given reliability. From these limits a knock-down factor is derived.

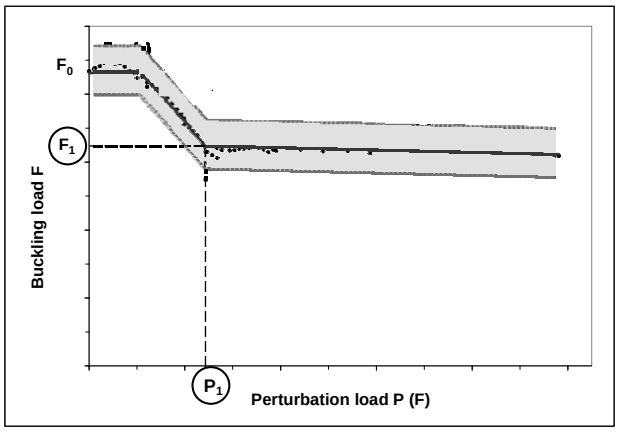


Fig. 4 Scatter of buckling load due to the scatter of non-traditional imperfections

2.5 Conclusions

From all this it becomes obvious that a great deal of knowledge is accumulated concerning the buckling of cylindrical shells under axial compression. However, the NASA guideline for the knock-down factors from 1968 is still in use, and there are no appropriate guidelines for unstiffened cylindrical CFRP shells. To define a lower bound of the buckling load of CFRP structures a new guideline is needed which takes the lay-up and the imperfections into account. Because there is a huge number of lay-up possibilities for each geometry, the new guideline must be different than the NASA SP-8007. This can be for instance a probabilistic approach [e.g. [23]) or the Single-Pertubation-Load approach as presented in [1]. Independent of the approach dozens of additional tests are necessary, in order to account for statistical scatter. In the following the idea of the DESICOS project is described which will try to combine both approaches.

3 DESICOS project

3.1 Objectives

The main objectives of DESICOS is realised by the implementation of the *Single Perturbation Load Approach* (SPLA) for buckling of imperfection sensitive structures. This approach considers the 'traditional' imperfections like the geometric and implicitly also the loading ones. In order to guarantee a robust design, the approach is combined with a specific stochastic procedure, which considers the 'non-traditional' imperfections like variations of thickness, material properties, etc. The outcome of the single perturbation load approach is the knock down factor ρ_1 and the outcome of the stochastic approach under consideration of specific aspects of composites (stacking, etc.) leads to the knock down factor ρ_2. Both factors are combined together in order to define the design load, cf. Eq. (1). This concept promises to exploit the full potential of composite structures in an efficient way.

$$F_{Design} = F_{perfect} * \rho_1 * \rho_2 \qquad (1)$$

ρ_1 = Knock down factor *Single Perturb. Approach*

ρ_2 = Knock down factor *Stochastic Approach*

3.2 Consortium

The DESICOS consortium merges knowledge from 2 large industrial partners (ADTRIUM-SAS from France and Astrium GmbH from Germany), one enterprise belonging to the category of SME (GRIPHUS from Israel), 2 research establishments (DLR from Germany and CRC-ACS from Australia) and 7 universities (Politecnico di Milano from Italy, RWTH Aachen, Leibniz University and the Private University of Applied Sciences Göttingen from Germany, TECHNION from Israel, TU-Delft from Netherlands and Technical University of Riga from Latvia). The large industrial enterprises and the SME bring in their specific experience with designing and manufacturing of space structures as well as their long grown manufacturing philosophies for high quality stiffened composite structures. The academic partners and the research organisations provide their special knowledge in methods and tool development as well as testing. This consortium composition assures the expected rapid and extensive industrial application of the DESICOS results.

3.3 Workpackages

The partners co-operate in the following technical work packages:

- WP1: Benchmarking on selected structures with existing methods: Benchmarks are defined for method evaluation purposes. The objective is the knowledge of the abilities and deficiencies of existing approaches.
- WP2: Material characterisation and design of structures for buckling tests: The first focus is on the design of structures which will be manufactured and tested in WP4. For that purpose, small specimens will be built and tested in order to characterise the specific composite material properties.
- WP3: Development and application of improved design approaches: In this workpackage new design approaches are developed, modelling and analysis strategies are derived. Finally, all methods are validated by means of the experimental results obtained from the other workpackages.
- WP4: Manufacture, inspection and testing of structures designed in WP 2: This workpackage deals with the manufacturing and testing of structures. The objective is to extend the data base on buckling of imperfection sensitive structures. Based on the designs from WP2 as input, a total of 14 (monolithic, sandwich, stiffened and unstiffened, cylindrical and conical) structures will be considered.
- WP5: Design handbook and industrial validation: WP5 comprises the final technical part; all the results of the project are assembled in order to derive the final design guidelines and to validate them as well as the new methods. The input is summarized in the improved design procedures, the documentation of the designs as well as the documentation of the experiments and their evaluated results.

3.4 Expected results

To reach the main objective, improved design methods, experimental data bases as well as design guidelines for imperfection sensitive structures are needed. The experimental data bases are indispensable for validation of the analytically developed methods. Reliable fast methods will allow for an economic design process. Industry brings in experience with the design and manufacture of real shells; research contributes knowledge on testing and on development of design methods. Design guidelines are defined in common, and the developed methods are validated by industry.

The results of DESICOS comprise:

- Material properties, measured according to the applicable standards
- Method for the design of buckling critical fibre composite launcher structures, based on the combined SPLA and stochastic procedures, validated by experiments
- Experimental results of buckling tests including measured imperfections, buckling and postbuckling deformations, load shortening curves, buckling loads
- Guidelines how to design composite cylindrical shells to resist buckling
- Reliable procedure how to apply the Vibration Correlation Technique (VCT) in order to predict buckling loads non-destructively by experiments
- Handbook including all the results
- Demonstration of the potential with different industrially driven use cases.

4 Summary

The main objective of the upcoming DESICOS project is the future design scenario for imperfection sensitive CFRP structures. The results comprise extended experimental data bases, improved design methods as well as design guidelines. More details can be found at www.desicos.de.

5 Acknowledgements

The project DESICOS is supported by the European Commission, Priority Space. The information in this paper is provided as is and no warranty is given that the information is fit for any particular purpose. The reader thereof uses the information at its sole risk and liability.

References

[1] C. Hühne, R. Rolfes, E. Breitbach, J. Teßmer, „Robust Design of Composite Cylindrical Shells under Axial Compression — Simulation and Validation", *Thin-Walled Structures* Vol. 46 (2008) pp. 947–962

[2] Degenhardt R., Kling A., Bethge A., Orf J., Kärger L., Rohwer K., Zimmermann R., Calvi A., "Investigations on Imperfection Sensitivity and Deduction of Improved Knock-Down Factors for Unstiffened CFRP Cylindrical Shells", *Composite Structures* Vol. 92 (8), (2010), pp. 1939–1946

[3] P. Seide, V.I. Weingarten, J.P. Petersen, NASA/SP-8007, Buckling of thin-walled circular cylinders, *NASA SPACE VEHICLE DESIGN CRITERIA (Structures)*, NASA ,Washington, DC, United States, September, 1965

[4] W.T. Koiter, On the Stability of Elastic Equilibrium. NASA-TT-F-10833, 1967

[5] B. Geier, H. Klein and R. Zimmermann, Buckling tests with axially compressed unstiffened cylindrical shells made from CFRP, Proceedings, *Int. Colloquium on Buckling of shell Structures*, on land, in the sea and in the air, J.F. Julien, ed.: Elsevier Applied Sciences, London and New York, 498-507, 1991

[6] Geier, B., H. Klein and R. Zimmermann, Experiments on buckling of CFRP cylindrical shells under non-uniform axial load, Proceedings, *Int. Conference on Composite Engineering*, ICCE/1 28-31 August 1994

[7] C. Hühne, R. Zimmermann, R. Rolfes and B. Geier, Loading imperfections – Experiments and Computations, *Euromech Colloquium 424*, Kerkrade, The Netherlands, 2-5 September 2001

[8] J. Albus, J. Gomez-Garcia, H. Oery, Control of Assembly Induced Stresses and Deformations due to Waviness of the Interface Flanges of the ESC-A Upper Stage, *52nd International Astronautical Congress*, Toulouse, France, 1-5 Oct 2001

[9] T. Winterstetter, H. Schmidt, Stability of circular cylindrical steel shells under combined loading, *Thin-Walled Structures*, 40, p. 893-909, 2002

[10] M. Pircher, R. Bridge, The influence of circumferential weld-induced imperfections on the buckling of silos and tanks, *Journal of Constructional Steel Research*, 57(5), p. 569-580, 2001

[11] M. Deml, W. Wunderlich, Direct evaluation of the 'worst' imperfection shape in shell buckling, *Computer Methods in Applied Mechanics and Engineering*, 149 [1-4], p. 201-222, 1997

[12] J. Arbocz, Jr. J.H. Starnes, Future directions and challenges in shell stability analysis, *Thin-Walled Structures*, 40, p. 729-754, 2002

[13] R. Zimmermann, Buckling Research for Imperfection tolerant fibre composite structures, Proc. *Conference on Spacecraft Structures*, Material & Mechanical Testing, Noordwijk, The Netherlands, ESA SP-386, 27-29 March 1996

[14] H.-R. Meyer-Piening, M. Farshad, B. Geier, R. Zimmermann, Buckling loads of CFRP composite cylinders under combined axial and torsion loading - experiment and computations, *Composite Structures* 53, 427-435, 2001

[15] R. Zimmermann, Optimierung axial gedrückter CFK-Zylinderschalen, Fortschrittsberichte VDI, Nr. 207, 1992

[16] Degenhardt R., Kling A., Klein H., Hillger W., Goetting Ch., Zimmermann R., Rohwer K., Gleiter A., "Experiments on Buckling and Postbuckling of Thin-Walled CFRP Structures using Advanced Measurement Systems", *International Journal of Structural Stability and Dynamics*, Vol. 7, No. 2 (**2007**), pp. 337-358

[17] C.A. Schenk, G.I. Schuëller, Buckling Analysis of Cylindrical Shells with Cutouts including Random Geometric Imperfections, WCCM V, 7-12 July 2002

[18] R. Reuter, T. Gärtner, Stochastische Crashsimulation mit LS-DYNA am Beispiel des Kopfaufpralls nach FMVSS 201, *17 CAD-FEM Users meeting*, Sonhofen, 1999

[19] R. Reuter, J. Hülsmann, Achieving design targets through stochastic simulation, *8th Int. MADYMO Users Conference*, Paris, 2000

[20] M. Klein, G.I. Schuëller, P. Deymarie, M. Macke, P. Courrian, R. S. Capitanio, Probabilistic approach to structural factors of safety in aerospace, Proceedings of the *International Conference on Spacecraft Structures and Mechanical Testing*, pages 679–693, Cépadués-Editions, **2-210, Paris, France, 1994

[21] C. Sickinger, L. Herbeck, E. Breitbach, Structural engineering on deployable CFRP booms for a solar propelled sailcraft, *54th Int. Astronautical Congress*, Bremen, Germany, 28 Sep.-3 Oct., 2003

[22] E. Velds, TUD/LR/ASCM graduation thesis, A deterministic and probabilistic approach to finite element buckling analysis of imperfect cylindrical shells, May 2002

[23] J. Arbocz, M.W. Hilburger, Toward a Probabilistic Preliminary Design Criterion for Buckling Critical Composite Shells, *AIAA Journal, Volume 43*, No. 8, pp. 1823-1827, 2005

CEAS 2011

New Concepts and Rapid Prototyping

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Construction of a CUBESAT using Additive Manufacturing

Franco Cevolini , Walter Holemans, Adam Huang
Italy, USA

Gilbert Moore, Stewart Davis
USA

Keywords: *additive manufacturing, aerospace design, cubesat, composite material*

Abstract

This paper examines the use of additive manufacturing and Windform XT to build a 2U CubeSat with an integrated warm gas Micro-Electro-Mechanical System (MEMS) propulsion subsystem with magnetic stabilization for CubeSat orbital altitude adjustment. The RAMPART project uses additive manufacturing techniques to build a custom satellite structure at a fraction of the time of current methods. Materials developed by CRP Technology for use in Formula 1 and NASCAR, with improved mechanical and thermal properties are being adapted for use in space. This paper describes manufacturing techniques and design guidelines used to build the structure of the 2U CubeSat, and the flexibility of design when using additive manufacturing.

1 General Introduction

The RAMPART CUBESAT (RApidprototyped Mems Propulsion And Radiation Test CUBEflowSATellite) project team consists of a group of volunteers brought together by Gil Moore to pursue the solution of several challenges facing the use of CUBESATs. The current nature of CubeSats requires that they be add-ons to current space flights, riding piggyback to an orbit location that may not be ideal for their intended mission.

The RAMPART project is designed to be deployed as a 2U CubeSat with the ability to utilize a MEMs propulsion unit to adjust the satellite's orbit after deployment from the launch vehicle. Walter Holemans of Planetary Systems Corporation conceived the idea of utilizing Additive Manufacturing (AM) to provide a flexible design path throughout the process, and to optimize the CubeSat until its final configuration.

Making use of Additive Manufacturing and Propulsion the RAMPART will utilize a "Design-Print-Fly" approach with the goal of raising its own apogee altitude to 1200 km, to allow the CubeSat to approach just below the equatorial inner Van Allen Radiation Belt, exposing the Windform XT and six other on board experiments to the energetic particle flux in that region. This will allow the testing of the material's performance as well as the performance of improved radiation-hardened Cubeflow circuits and new high performance solar cells. Details of the Cubeflow and particle test equipment were given in the paper entitled, "3D Printing and MEMS Propulsion for the

RAMPART 2U CUBESAT" by Gilbert Moore at Small Sat 20101.

1.1 Document text

The RAMPART satellite is a 2U CubeSat consisting of an upper BUS module, Lower Propulsion Module and deployed Solar Panels (Figure 1).

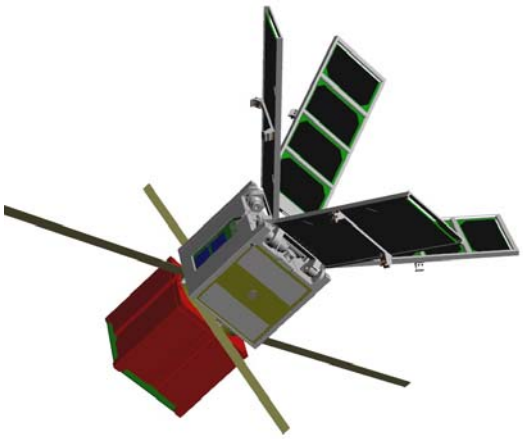


Fig. 1 RAMPART with panels deployed

The entire structure, built in Windform XT via Additive Manufacturing, is plated in a High Phosphorus Electroless Nickel to provide radar reflectivity for tracking purposes. The upper BUS module consists of several experiments as well as test solar cell panels mounted to the exterior of the module. An additional experiment related to the performance measurement of the Windform XT material has been placed at the top section of the BUS (Figure 2).

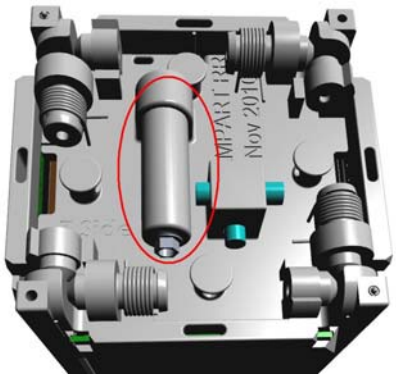


Fig. 2 Top of BUS module showing load cell circled in red

The load cell (Designed by Walter Holemans, Planetary Systems Corporation) measures the change in preload of Windform XT. A 400 lb compression load in the load cell will cause 400 lb tension and about 2,500 PSI stress in the Windform XT material. This integrated load cell is designed to measure creep or fracture as a result of exposure to radiation and thermal cycling over time.

RAMPART's Solar Panels will be spring–erected following the satellite's deployment from its launch vehicle.

The aft portion of the satellite consists of a Micro Electrical Mechanical (MEMs) propulsion system designed and developed by Dr Adam Huang of University of Arkansas. The RAMPART propulsion system incorporates a miniaturized resistojet thruster core with a microfabricated de Laval nozzle and integrated heater. Upstream of the nozzle/heater assembly is an injector fed by 3 miniature solenoid valves. Prior to arriving at the valves, the propellant passes through a 2-phase separator membrane where only the gas phase of the liquid propellant can pass through its micron-sized pores. The membrane also serves as a filter for preventing the valve seats from becoming contaminated with debris. The propellant used is the refrigerant R-134a, which is considered nontoxic and nonflammable. The compressed fluid nature of the R-134a provides relatively high self-pressurization for delivery throughout the thruster system.

Although resistojets do not represent the state-of-the-art in thruster technology and offer relatively low thruster efficiencies (RAMPART's design specific impulse of 90s), their simplicity and practicality apply well to pico- and nano-satellites. A key factor to the performance of the RAMPART propulsion system is the light weight and cell structure of the propellant tank made from Windform XT. Instead of utilizing state-of-the-art technology, RAMPART uses high propellant mass fraction to provide the required Delta-V (320m/s for RAMPART) for Low-Earth-Orbit in-plane orbital maneuvers. Since the dimensions and the weight of a CubeSat are its main constraints, the Windform XT allows these parameters to be

optimized through the usage of multiple and interconnected near-cubic cells. This offers the advantage of maximizing the propellant volume, integrated baffles provided by the interconnecting walls, and possibility of correlating material strength tests of a unit cell to the entire propellant tank design. One can easily compare this design to traditional pressure vessels (hemispherical caps) and realize a two times difference in the propellant storage volume.

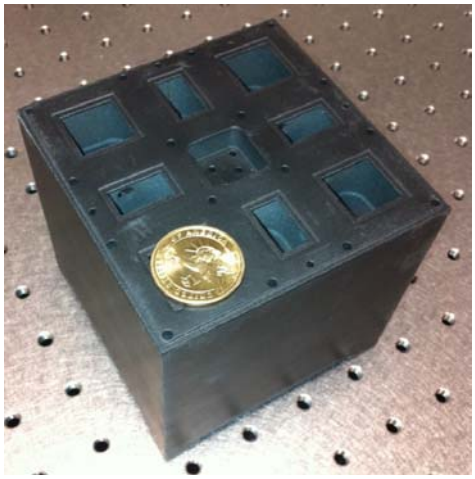


Fig. 2 Top of propulsion module

To determine material performance, test cubes similar in size to the baffled chambers, shown in Figure 3 and 11, were tested to various pressures. After physically testing the small test cubes that represented the chambers up to 600psi, an FEA model was developed by Whitney Reynolds of the U.S. Air Force Research Laboratory, Space Vehicles Directorate (AFRL/RV), to simulate the reactions of the small chambers and then correlate the results to a simulation of the final large propulsion unit.

Material Choice:

The use of Rapid Prototyping Plastics in Aerospace has been somewhat limited to prototype work with the exception of Nylon 11 developed by Boeing (ODM) and Nylon 12 utilized by Northrop Grumman. (Programs include F-18 and Global Hawk). These two materials are used in a production environment and are produced using Laser Sintering (LS) machines. The LS technology has been seen as the best opportunity for durable plastic materials with a consistency in manufacture. Selective Laser Sintering resolution for plastics is commonly run at 0.004" per layer utilizing a CO2 laser that can be adjusted to melt the plastics into a fully dense material.

Fused Deposition Modeling (FDM) by Stratasys, Multi Jet Modeling (MJM) by Objet, and Stereolithography (SLA) from 3D Systems are some of the Rapid Prototyping technologies examined. Each technology and material has benefits for prototyping. Key factors helped steer the material choice, including Heat Deflection temperature, UV exposure, and the requirement of plating to create to make the satellite radar reflective.

Windform was developed by CRP Technology based in Modena, Italy. This company began working on materials for the Laser Sintering (LS) systems in 1996 after their purchase of LS equipment from DTM. Their interest in the materials was driven by a customer request for an improvement in stiffness - specifically for wind tunnel applications in Formula One racing (Figure 4).

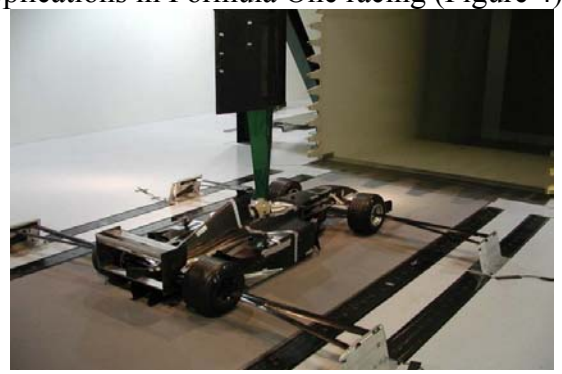


Fig. 4 F1 scale wind tunnel model equipped with Windform XT

Introduced for sale in 2004, Windform XT uses a base Polyamide and reinforces this with Carbon Microfibers. Racing teams were the first to begin using the material, and determined it to be a solution that could be utilized in "on car" applications. Windform is currently used by both F1 (Figure 5) and NASCAR teams to replace components that would typically require injection molded materials.

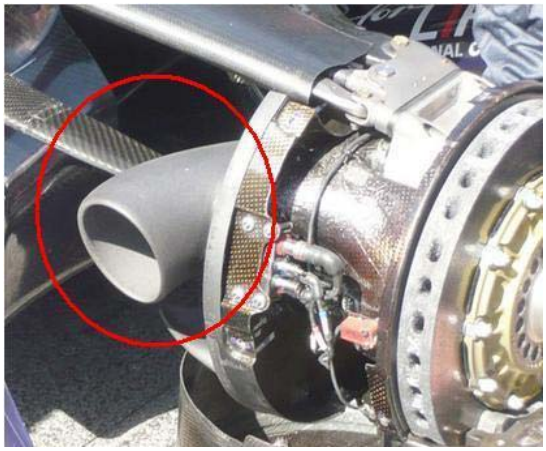


Fig. 5 Windform XT used as brake inlet duct on F1 Car

Based on observations of other materials, samples of Windform® XT were subject to Tensile

Test and the cross sections of the brakes were examined under an electron microscope. The micrographs showed the carbon microfibers were encapsulated by the Nylon base material. In addition little to no porosity was visible in the internal structure. (Figure 6)

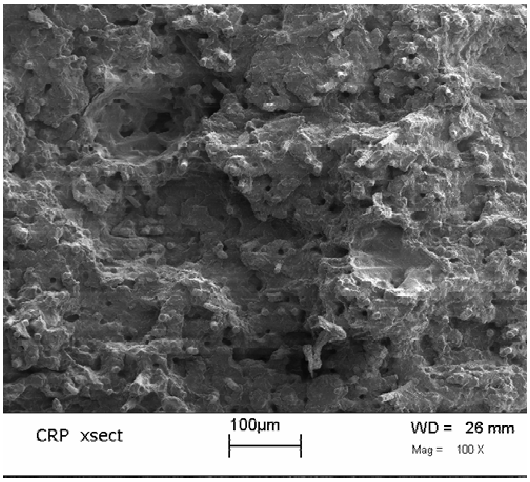


Fig. 6 Courtesy Paramount Industries

In addition, studies have shown that Windform XT performs in a predictable manner and has been compared to injection molded production materials to determine how it would react in exposure to extended temperature cycling. Having a predictable material is important in the examination of possible failure modes for design studies.

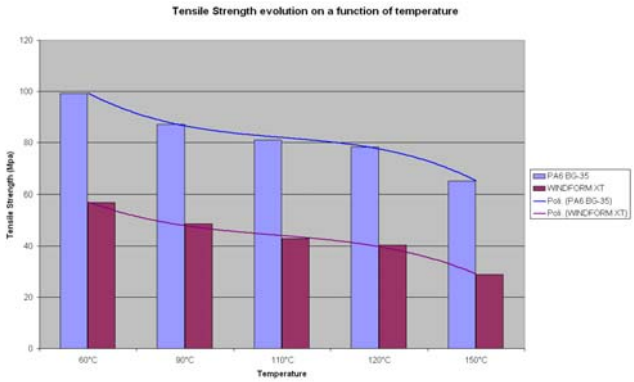


Fig. 7 Comparison of Windform XT to PA6 BG-35 injection molded plastic

The following key factors make Windform XT a good candidate for Additive Manufacturing for the CubeSat application.

1. Windform XT passes Outgas Screening, ASTM E-595.
2. Windform XT is produced in a manner that makes it dense.
3. Windform XT can be easily machined using conventional methods.
4. Windform XT has a good heat deflection temperature (HDT) relative to other RP materials. (HDT above 170C).
5. The base polyamide material has been proven to meet performance needs for other Aerospace applications.
6. Material batches are quality controlled, each coming with a Certificate of Conformance (COC).
7. Build volume of SLS system fits well with CubeSat applications= 381 x 330 x 457 mm (14.5 x 12.5 x 17.5 in).
8. Windform XT can be plated without the need for sealing agents.

Construction of BUS and Propulsion Module

Windform XT is used to build parts using the Laser Sintering process. This process can be described in several steps. (Figure 8)

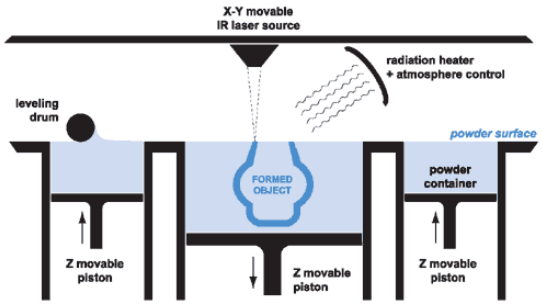


Fig 8 Illustration of LS process

Laser Sintering is a powder based Additive Manufacturing method. The process uses a layer system of building up a part with each layer of powder being sintered separately by a laser. As each layer is sintered it slowly constructs the part step by step. There are various grades of materials used each with their own characteristics, but each are used the same way within the LS machine.

The LS machine basically consists of three powder beds and a Laser. Two of the powder beds hold the feed powder and the third bed holds the part. The part bed is in the middle of the beds with the laser acting directly perpendicular to this bed. A roller is used to push the layers of powder over the part bed and all three beds have their own heater source. The process itself is a very simple, repeatable one.

The building of the parts is a repeatable two-step process.

Step 1

A roller is positioned beside one of the feed beds. This feed bed then raises a set amount (Usually < 0.1mm) and the roller pushes the raised powder across, covering the part bed with a Powder Layer.

Step 2

With the layer of powder present, the laser starts to trace out the desired shape of the part in the powder, melting the powder as it contacts the surface.

Step 3

Once this is done, the part bed drops down the set amount and the process continues from the opposite side, with the other feed bed raising and the roller distributing another layer of powder over the part bed, and the laser again traces the shape.

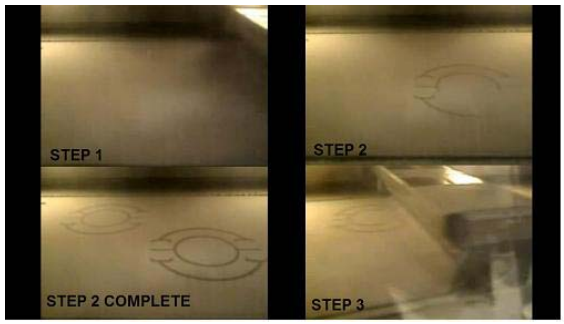


Fig 9 Illustration of LS process

The part is built up in slices, with each layer of powder representing a single slice of the part. As the laser melts the powder, each layer fuses together to grow a solid part. Because of the fact that the part is made up in layers, very complex shapes and design can be manufactured that would otherwise be impossible by conventional means.

LS Process benefits:

1. Layered technology is free from geometrical constraints: it is possible to build undercuts, hollow parts and internal ducts.
2. Opportunity to build directly functional "assembled mechanisms" (minimum clearance between parts 0.5 mm).
3. It is possible to build many different parts together in the same building volume.
4. Building time is not dependant on object geometry, but only on part volume and on build "height" (Z dimension).
5. Building time is very short (max 1-2 working days).
6. Lead time is very short (max 2-3 working days).
7. When a great dimensional accuracy is required, Windform parts can be machined.
8. To reduce mass and weight it is possible to create hollow parts with internal reinforcing structures

Items to consider when designing for SLS Process:

1. Building volume (360x310x380 mm). Larger objects are produced divided into parts and then bonded together with approved materials.
2. Minimum feasible wall thickness is 1 mm. Ideal value is 1.5-2 mm.
3. Minimum feasible details >=1 mm.

4. It is important to avoid enclosed volumes: hollow parts will need a hole or some type of gate to allow the removal of loose or un-melted powder.
5. When designing parts to be assembled together always keep a minimum 0.2 mm clearance between them.
6. Excessive wall thickness (>10 mm) can cause unwanted warping and shrinkage on parts.
7. Roughness after LS process is high (Ra 6 □m) and Z layering is visible. It is possible to smooth external surfaces to reduce it to Ra 1.5 □m.

Windform XT uses micro fibers as a reinforcing system and is similar to carbon lay-up techniques in that it is non-Isotropic. Similar design rules are applied when building parts to take advantage of the increased strength in one direction over the other resulting from the build process. In the case of the RAMPART BUS, the decision was made to produce the part with the Z build axis vertical to the satellite's longest axis after assembly. This allows the X and Y axes to be positioned to provide the most strength in the perpendicular axis holding the BUS together.

BUS Construction

The BUS underwent several revisions within the period of a few months. Walter Holemans was able to add experiments, create different deployment options, and develop wire routing and harness control devices utilizing design techniques to maximize the use of Additive Manufacturing. Increased complication was not a hindrance to the process and allowed for greater design freedom. The Solar Cell experiment (Figure 10) was easily tucked into the side of the BUS by only needing to implement a simple feature cut in the CAD system.

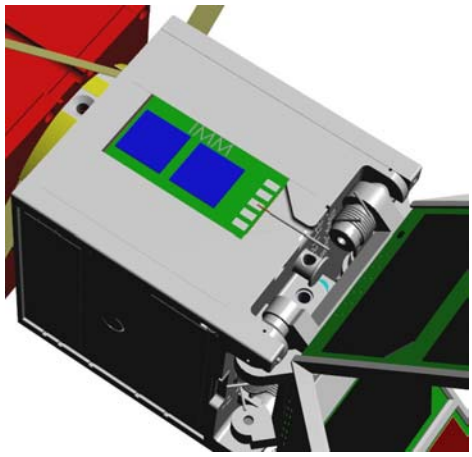


Fig 10 Solar Cell experiment added to RAMPART's BUS module

The Bus and Propulsion Modules were created in separate sections to allow the different teams to test and assemble them at different locations. This allowed the production of prototype modules to try ideas based on the current challenges and then adjust as needed. At the completion of the final design, the parts were grown at 0.004" (0.1mm) layers and then sanded to a smooth finish to aid in the plating process.

Utilizing the experimental and simulation data, the Propulsion module (Pressure Vessel) was oriented such that the Z build orientation was again parallel to the longest axis of the CubeSat. Due to the internal structures of the baffled cubes, (Figure 11)the module was subjected to an ultrasonic water bath to ensure no material was trapped in the chambers.

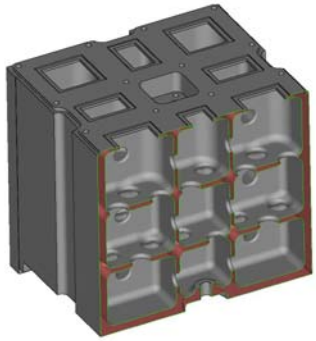


Fig 11: Cross section showing baffle chambers in propulsion module

The plating was completed by Quaker City Plating located in Whittier, California, under the direction of Frank Huizar.

Conclusion

The innovative use of Windform XT and Additive Manufacturing has shown that this type of technique will allow for fast adaptability and freedom of design. Working with the idea of utilizing additive manufacturing as a core principle, the team was able to modify, change, and add experiments without concern for having to develop tooling or modify an existing cube structure.

The development of RAMPART shows that this type of building technique adds value for standardized Cube Sat boards as well as for devices that do not fit easily "in to the box." As electronics and sensor systems become smaller and more complex, there is a push for developing more complex and advanced CubeSats. In the development of greater complexity, Additive Manufacturing opens the possibility to adapt the structure to carry new sensors, or optics. Utilizing CAD and additive manufacturing, the internal structure of the Cube Sat can be built to adapt to the components, instead of the other way around. This allows the use of the CAD system to adapt to new technology as it is introduced. This eliminates the concern for legacy systems that are being held on the shelf, for technology that is becoming obsolete. A new optic or sensor can be fitted immediately and the benefit deployed to the field as quickly as possible.

In addition, standard components can be placed into a CAD library that will allow for parametric generation of portions of the satellite. Additive Manufacturing of the RAMPART BUS module shows a clear mixing of standard board modules and customizing for solar panels , wire routing, and the addition of the load cell. Another observed benefit was the reduction in fasteners and ease of assembling the RAMPART CubeSat. Positioning features and "snap fit" devices can be incorporated into the design methodology to speed assembly. The extensive baffle design in the tank structure of RAMPART leverages this consolidation of components, removing the need for a complicated assembly.

References

[1] Moore. G., "*3D Printing and MEMS Propulsion for the RAMPART 2U CUBESAT*" Small Sat, Logan UT, August 2010.

Conceptual design of a rechargeable air thruster for small aerial vehicles

Lorenzo Olivieri
Aerospace Engineering Master course, University of Padova, Italy

Marco Chiaradia[1], Gabriele Rodeghiero[2], Alessandro Francesconi[3]
[1]Aerospace Engineering Master course, University of Padova, Italy
[2] Astronomy Master course, University of Padova, Italy
[3] Ph.D Researcher, CISAS 'G. Colombo', University of Padova, Italy

Keywords: *Cold Gas Actuator, Balloon, Rechargeable*

Abstract

A lot of research in Aerospace Engineering aims both to miniaturization and development of economic projects. In this scenario the increasing employment of Lighter Than Air Vehicles such as blimps and balloons requires a new generation of attitude controllers with lower requirements in terms of mass and overall resources. This document analyzes the optimization of an innovative Rechargeable Cold Gas Actuator which is able to collect the propellant directly from the surrounding environment. In the first part of the paper the conceptual design of the actuator is introduced. After that, the optimization of thrust and total impulse parameters and to the development of a simplified technological demonstrator is discussed, and the last section analyzes future applications and niches of the marketplace.

1 Introduction [1][2]

In the last decade, there has been a considerable interest in the development of small-scale air systems. This is due to the gain in mass, cost, performance, reliability and operational flexibility that reduction in vehicle size could offer. Small scale and light-weight aircrafts could in fact adapt to a large variety of commercial, scientific, police and mapping applications as well as remote observation of hazardous environments that are inaccessible to ground vehicles. One of the key issues in developing autonomous small-scale lightweight systems is related to the ability of providing accurate stabilization and maneuverability. This justifies the attention to compact, low-thrust (in the range 10-1000 mN) and accurate actuators.
Among the technologies considered for application to light-weight and small-scale systems, cold gas thrusters represent a valuable candidate under evaluation: their working principle consists on ejecting gas through a nozzle at high speed in order to create force. Main advantages of cold-gas thrusters are the simplicity, ease of operations, low-cost and suitability to miniaturization. On the other hand, such actuators have two main drawbacks: the need of carrying propellant on board and the very low specific impulse (below 60 s). Considered the above mentioned advantages and limitations, the objective of this paper is to evaluate the possibility of exploiting the cold-gas technology in a way that overcomes the need of carrying the propellant on-board. Being the atmosphere the operational theatre,

propellant could directly derive by ambient air, after proper increase of its energy through active heating and/or compression.

In more detail, the paper presents a numerical evaluation of an air actuator containing three main parts (Fig.1): in the first, the gas is directly collected from the external environment and is energized by a group of micro-compressors, then it flows into a tank split into two stages separated by automatic valves, with heaters to warm it; in the last, compressed air is ejected from the second stage through a nozzle, finally delivering the thrust.

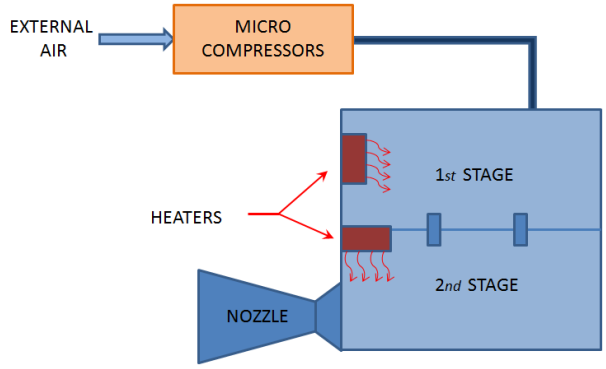


Fig. 1: Actuator layout

Careful design of the system (including the two-stage operating cycle) makes it possible to maximize the peak thrust or the total impulse, according to the application requirements.

The remainder of this paper is organized as follows. Next section is dedicated to the conceptual scheme of the actuator and to the individuation of the performance design drivers, and in sections 3 and 4 the development of the numerical model and the optimization process are presented, to realize an high-performance actuator for terrestrial and other planets applications (section 5). At the end of section 4 a simplified actuator is presented as a technological demonstrator.

2 Actuator conceptual design

The optimization of the actuator performances follows a inch by inch logic: starting from the functional scheme depicted in Fig.1, the design parameters are analyzed and their influence on performance are discussed.
A good choice can be to use a three compressors configuration in the pneumatic circuit: the pressure increase allows the utilization of a supersonic nozzle with a gain in terms of thrust and impulse. The nozzle shape is also another design parameter, and the diameter of the valves and the nozzle throat are important, too. Heaters can be implemented in the two stages of the actuator, and their effect will be analyzed in the simulations.
At last, the most important design parameters are the actuation profiles: the opening and closing pressure values of the valves influence the other parameters and the performance of the actuator.

2.1 Design Drivers

In the actuator optimization process, the design drivers are the elements that can be modified in order to have a different dynamic behavior of the system: numerical simulations collect all the results of the combination of these variables and a quick analysis of the resulting performance parameters gives the final values of the design drivers.

The next paragraphs present the main design parameters and analyze the possible improvements to the basic layout (visible in fig. 1).

2.1.1 Flow diameters

The two most important flow dimensions of the actuator are the flow area of the valves between the first and the second stage of the reservoir and the throat area of the nozzle.

The first design driver, the equivalent flow diameter of the valves, is the diameter of a circular section with the same area of the total flow section of the valves between the two stages. The simulated values of this parameter are listed in Table 1.

About the nozzle throat, the section is an independent parameter in the flow governing equations: assuming a supersonic flow in the nozzle (see next sections), the throat area is the

critical area. A high value of this parameter means a high thrust and a short impulse; a small section, on the other hand, gives a lower thrust and a longer impulse. The simulated values are visible in Table 1.

Table 1: Valve (left) & nozzle (right) diameter values in the simulations

Diameter $m \cdot 10^{-3}$	Diameter $m \cdot 10^{-3}$
4	3
6	4
8	6
10	8
12	10

2.1.2 Compressors

An interesting solution is the implementation of a three compressors layout, with the third in series of two pumps in parallel. Because of the low prevalence (not over 150% of the incoming pressure for commercial miniaturized compressors) of a two parallel pumps configuration, with the third compressor the higher pressure reached allows the implementation of a supersonic nozzle and provides a higher thrust.

The first estimation of the maximum pressure the compressors can reach demonstrate that the implementation of this layout is desirable: with a power consumption increment of less than 50% by the third compressor, the pressure in the vessel grows of about 100%.

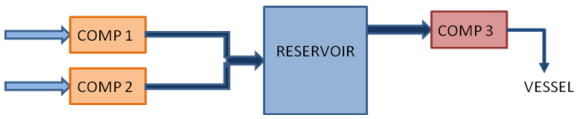


Fig. 2: pumps configuration (in red the added compressor)

The main disadvantages of implementing a three compressors system are the loss of redundancy, this configuration has only one pneumatic circuit for all the compressors instead of two separate circuits in the case of simple parallel, and the presence of a little reservoir upstream the third pump, as visible in Fig.2.

The small reservoir is important to stabilize the pressure between the first two compressors and the third: without this rebalancing chamber the pressure in the pipes would manifest higher gradients and the compressors would work in harder conditions.

2.1.3 Supersonic Nozzle

The direct consequence of the implementation of the third compressor is the possibility of using a supersonic nozzle instead of a subsonic or sonic one.

The thrust given by the nozzle is calculated with the relation:

$$S = \dot{m} \cdot u + A_e \cdot (p_e - p_{ext}) \qquad (1)$$

A_e is the final section of the nozzle, and u the speed of the fluid in that section. With a subsonic nozzle (M<1) the pressure in the final section is the same of the environment, and the speed u is minor than the speed of sound. The supersonic nozzle gives a higher Mach number, so the speed increases, and the pressure p_e can be different from the external pressure; in this nozzle the thrust is higher.

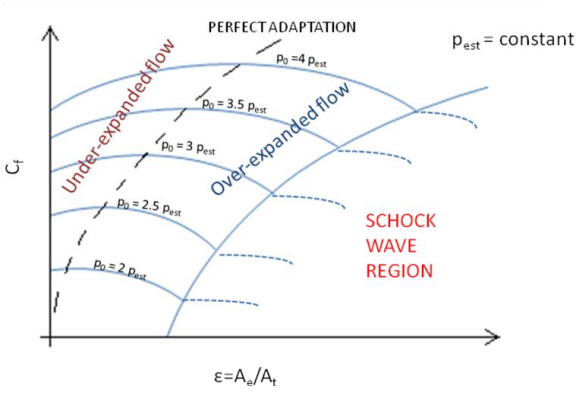


Fig. 3: simplified C_f graph

The actuator nozzle must avoid the shock wave region, and should give a nearly constant thrust, so it has to work in under expanded flow: as visible in the force coefficient graph in Fig. 3, with a fixed value of the expansion coefficient ε the thrust coefficient values have a lower

variation in this condition, resulting in a flattening of the thrust.

The design in under expanded flow is also important for a flexibility principle: in case of extraordinary events the nozzle can work in over expanded conditions, elaborating longer impulses, although with a lower performance.

In the simulations the nozzle design parameter is the inverse of the expansion coefficient ε due to numerical reasons: the nozzle throat is a valve, so in the simulations the value of the throat area is 0 when there isn't flow, and A_t when the nozzle is open, and the ratio ε can give numerical errors. To avoid this problem the model implements the ratio $r = A_t/A_e=1/\varepsilon$: the values of r lie between 0 and 1.

In next table the simulated values of the expansion ratio are listed.

Table 2: simulated values for the nozzle

ε	r=1/ ε
1.01	0.99
1.15	0.87
1.30	0.77
1.50	0.67

The pressure in the vessel can't reach values over 5 times the external pressure, and the selection of the values of Table 2 do not consider higher expansion ratios to avoid unnecessary simulations.

2.1.4 Stages Size

The stages size influences the duration of the impulses, a bigger vessel contains more propellant than a little one, but this gain is balanced with the longer time to fill the reservoir with the compressors.

The stage volume therefore seems to be an important design parameter: its influence on the actuator efficiency will be evaluated and presented in section 3.

2.1.5 Heaters

The implementation of two heaters, one for each stage, could influence the performance of the actuator. It is not simple to foresee the effect of the heaters on the thermodynamic of the system, and the simulations are the easiest way to determinate the benefit of this solution. The simulated values of the incoming power from the heaters in the simulations are of 0 W, 1 W, 3 W and 5 W for each one, with the hypothesis of an efficiency of 100% in the power transfer from the heater to the gas.

In some applications at low temperatures (see section 5) the heaters are necessary: for example, with the pressure increase, the water vapor (or the methane, in Titan atmosphere) can reach saturation and start to condensate in the stages and in the pipes.

2.1.6 Actuation Profiles

At last, the most important design drivers in the optimization process are the actuation profiles, that can be described through the relative values of the pressure differentials for opening and closing the valves (see eqn.2). These parameters influence and are influenced by all of the other design drivers, and their effect can be evaluated only with the simulations.

The opening and closing values of the nozzle can modify the dynamic of all of the actuator: high values can give short impulses of high thrust, low values give lower thrusts with a lower recharge time. A high opening value and a low closing value can create the longest impulse, but with a very long recharge time. The closing value also influence the expansion ratio of the nozzle: a high internal pressure can be elaborated by a nozzle with a higher expansion coefficient. The internal valve is important, too: a high open value could avoid the opening of the valve because of the too much high pressure to reach, on the other hand a low opening value can give a short impulse.

In the next table the simulated values are listed. The values represent the ratio of relative pressure on front pressure:

$$p = (p_{back} - p_{front}) / p_{front} \qquad (2)$$

Table 3: simulated values

Internal Valve		Nozzle	
Open %	Close %	Open %	Close %
30	10	150	100
50	30	180	150
70	50	200	180
90	70	230	210
110	90	260	250
130	110	300	280
		350	300

3 Numerical Model

The numerical model is developed in Matlab™ Simulink language. It simulates a system of three compressors, two in parallel and one in series, and two vessels. Between the vessels a valve is implemented, and at the exit of the second stage the model simulates a supersonic nozzle. In all the components the model represents the gasdynamics and the thermodynamic laws, finally calculating the thrust of the nozzle. To reduce the computational heaviness, only the supersonic case is analyzed; if the flow is subsonic or with shock waves, the model gives a warning with a -1 N value of the thrust, in order to simplify the simulation data analysis. The launch program can be modified to change the design parameters and the environmental characteristics. The simulations campaign is split in different analysis, due to the high number of values of the design drivers (more than 2 millions possible combinations at each environmental condition). Each simulation also follows an optimization process to reduce the possible values of the design parameters and faster the analysis

3.1 Numerical Model Components

The main part of the numerical model is formed by a system of pre-designed advanced Simulink components that can be combined to the desired configuration. The reservoir elements represent complex subsystems (with variable volume or variable mass) and contain the analysis of the thermal exchanges and the mixture of fluids. The valve elements are designed as convergent nozzles, and evaluate the fluid motion and its gasdynamic characteristics. The combination of these two elements and the implementation of a supersonic nozzle and a pumps subsystem allows the analysis of the performance of the cold gas actuator.

3.2 Optimization

The analysis of the actuator was performed with a simulation campaign. The chosen design parameters gives 2.073.600 possible combinations, not counting the different environmental conditions. To reduce the computation time and the results analysis, the simulations are split in different sections at constant ambient temperature and pressure. Each simulation also followed an optimization process to reduce the possible values of the design parameters and faster the analysis. If the general simulation would be launched, calculating 1 second of elaboration each run, the total elaboration time would be of more of 2 Ms, that are about 24 days of CPU work. This procedure would be repeated for 6 times due to the different environmental conditions. Working with a quad-core processor the simulation could be split in 3 parts, each one on a single CPU with the 4th dedicated to the PC routines, but the total time could be only reduced to 48 days, 8 each simulation.

The two solutions to avoid this problem are to improve the launch model in order to avoid useless simulations and to decrease the number of variables. About the first solution, the model implements a control loop that, if the value of the thrust is -1 (i.e. the flow is not allowed), skips the simulations with worst conditions (i.e. a bigger expansion ratio or a lower opening pressure). A preliminary simulation campaign demonstrates a reduction of ~20% of the number of simulations.

The reduction of the number of variables is possible after a first simulation: the results demonstrate that the implementation of a supersonic nozzle with high expansion ratios is not really useful, because it needs higher pressures to elaborate higher thrust of short duration, reducing the total impulse. Other parameters that can be simplified are the actuation profiles of the valves. Last, the effect of the heaters seems to be negligible, due to their low power (max. 5 W each one); using more powerful heaters would increment the power budget, dumping the performance of the actuator.

4 Results

The main results of the simulation are the maximal thrust and the total impulse of the actuator. These parameters are normalized by the power consumption, as an indication of the actuator performance. In Table 4 the main results are presented.

4.1 Simulation Results

The simulation campaign provided this evaluation of the optimized actuator configuration. The required average power is of about 30 W at ground and lowers with the altitude, and the tank has two stages both of 1 dm^3. The internal pressure range lies between 0 and 3 bar. This configuration employs three compressors, two in parallel and one in series downstream of them; the implementation of a supersonic nozzle is not useful, too, within the pressure ranges of the actuator. At last, the effect of the heaters seems to influence the flow only for low values of temperature and pressure (i.e. altitude > 10 km in Standard Atmosphere), but their power consumption is more significant than the variation of total impulse.

It came out that the impulse is dependent from external air pressure and temperature; not considering a short transient phase at the switch-on of the actuator, at ground conditions an impulse of 0.24 Ns can be reached with long pulses of low thrust (< 1 N), suitable for modulated controls, or shorter pulses of 0.02 s with higher thrusts (about 8 N). At higher altitude, with environmental conditions of 500 mbar and 250 K, the impulse drops to about 0.12 Ns, due to the reduction of temperature and pressure. Higher altitudes lead to a further decline of the total impulse and of the required power. In Table 4 and in Fig. 4 the trend of the total impulse and of the maximal thrust are normalized on the power consumption.

Other results of the simulation campaign demonstrate that there is not an univocal solution to maximize the performance, and if the two stages configuration is useful under 10 km of altitude, a single stage solution is more

Table 4: simulation results

Environmental Conditions			Power Cons., W	Total Impulse		Max. Thrust	
Pressure, mbar	Temperature, K	Earth Altitude, m		Ns	Normalized, Ns/W	N	Normalized, N/W
1000	290	0	34	0.2359	0.0069	7.70	0.2265
500	250	5000	21	0.1213	0.0058	3.30	0.1571
250	220	10000	15	0.0206	0.0013	0.72	0.0480
100	215	15000	13	0.0241	0.0018	0.62	0.0477
50	215	20000	12	0.0084	0.0007	0.17	0.0142
*25**	*220*	*25000*	*12*	*0.0013*	*0.0001*	/	/

* *Out of the working range*

functional at higher altitudes: it should be implemented a by-pass system to allow the two stage actuator to work in these conditions.

The total impulse lowers with the altitude, from 0.24 Ns at 0 km to 10^{-3} at 25 km; the analysis also demonstrates that the highest evaluated altitude (25 km) is out of the working range of the actuator.

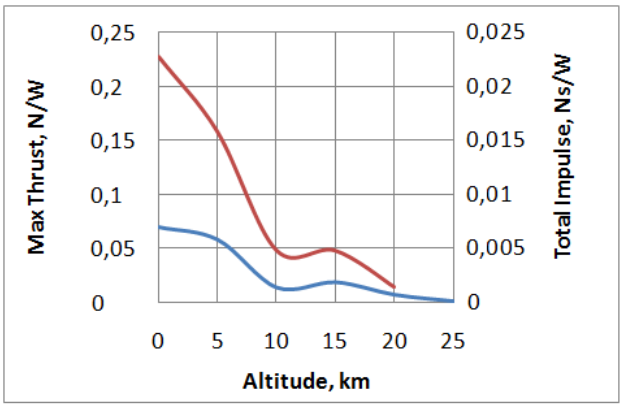


Fig. 4: normalized total impulse (blue) and thrust (red) at different altitudes

Last, the variation of the geometric dimensions of the vessel is directly correlated to the total impulse intensity with a quadratic proportion, but the mean thrust is nearly independent from the volume and remains almost constant.

4.2 SCRAT Technological Demonstrator

Finally, the theoretical model has been validated using data collected in the environmental test of the SCRAT (Spherical Compact Rechargeable Air Thruster) technological demonstrator, launched on the BEXUS 10 balloon [1].

SCRAT is a simple version of the proposed actuator, using only two compressors and no heaters; it was tested through the different atmosphere layers reached during the flight, from ground to about 25 km of altitude.

The flight campaign [3] led to collect a total of 101 thrusts which span between 10^{-2} and 10^{-4} N, with a pronounced clustering about the 10^{-3} N scales; about the total impulse, the results lies between 10^{-2} Ns and 10^{-3} Ns.



Fig. 5 SCRAT launch on BEXUS 10

If the numerical values of the performance parameters seem to stand out from the foreseen performances, due to the different design, the total impulse trend is comparable with the simulation results.

5 Conclusions

The comparison with a simplified model of a small propeller with the same power consumption demonstrates that the rechargeable cold gas actuator has a lower mean thrust, but it is in line with the maximal thrusts (i.e.: 2 N for the actuator, 2.7 N for the propeller at 5 km): in terms of theoretical performance a propeller is better than the rechargeable gas actuator at these altitudes, because of its higher impulse.

The main cons of propellers can be listed in three categories: the gyroscopic behavior, the effect of boundary conditions on the real flow, the bulk of the propeller and of the mechanical transmission. With the cold gas actuator, these limitations aren't so stringent, and the implementation can be easily adapted to the design requirements.

An interesting application could be the control of balloons such as the InfraRed Hot Air Balloons (Montgolfiere Infra Rouge - MIR), developed by the French CNES (Centre National d'Etudes Spaciales). The MIR [4] is a new concept of hot air balloon: it can fly up to several months, changing altitude flight between day (30 km) and night (about 18 km).

The actuator could be used in the attitude control of the MIR, with different performance during the day and during the night, but it also could be used to control the navigation: the actuator could work each morning as a thruster, using the air collected at the lower altitude to deliver an higher thrust at the day-navigation altitude to change the rout. On terrestrial applications like MIR, the actuator could be used for small airships attitude control: next figure show the maximum controllable speed of gusts at different altitudes. The numerical values are calculated for small airships (Area 10 m^2, C_D 0.5).

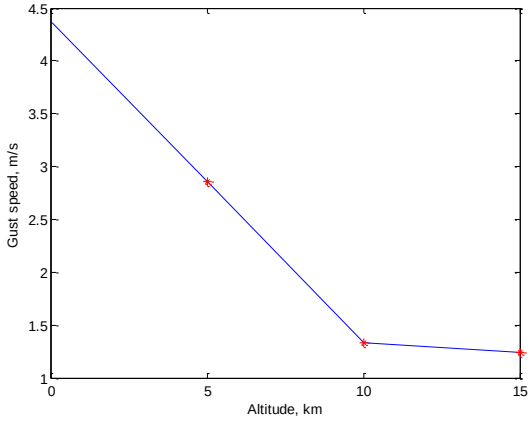


Fig. 6: Max. controllable gust speed

Space exploration could involve balloons on planets and moons with atmosphere; if on Mars the actuator should be improved to work in the thin atmosphere of carbon dioxide, on Venus the area of the troposphere most similar to Earth is above an altitude of 50 km, near the tropopause, with a temperature between 290 and 300 K and a pressure of about 1 bar. Soviet program VEGA (1986) involved two balloons, dropped onto the planet's darkside and deployed at about 56 km, to make atmospheric measures and to demonstrate the capability of balloons in planet explorations. Due to the high wind speed, on Venus the rechargeable cold gas actuator could be used only for attitude control of airships, with performances similar to Earth's.

Last, Titan has many similarities with the Earth: due to the presence of organic molecules and water, Titan can be described like a pre-biotic Earth. Cassini-Huygens, was the first mission to deeply explore the biggest moon of the giant planet, but not the last: the Titan Saturn System Mission, an ESA - NASA future project, will be dedicated to the exploration of Titan and will continue Cassini-Huygens work. Its launch is not jet decided, but the main payloads are well defined: an orbiter, a lander, and a small balloon floating around the moon at about 10 km of altitude, carried by winds.

In the balloon exploration of Titan, the implementation of a cold gas actuator can be evaluated. About the environmental conditions [5], the pressure is higher than on earth under 20 km, with a better performance of the actuator: a first evaluation, using the developed numerical model with a 10 km altitude (87 K, 1250 mbar) gives a maximal thrust of 9 N, a total impulse of 0.4 Ns for pulses of about 0.1 s.

Due to the low temperature of Titan atmosphere the methane can condense during the compression cycles: to avoid the accumulation of the methane in the first vessel, a valve is required on the bottom of it, to remove the liquid.

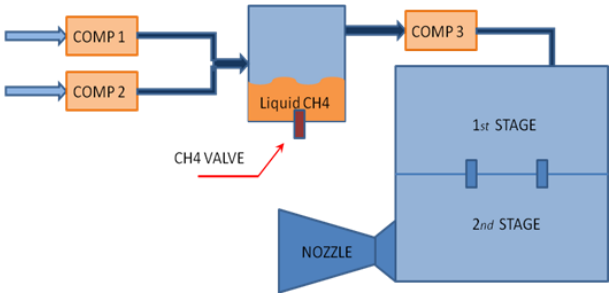


Fig. 7: Titan configuration of the actuator

The cold gas actuator is a good choice also in terms of the airship mission design: the two stages can be filled on Earth with the helium to inflate the balloon on Titan, using a reverse valve in the pneumatic circuit, with a gain in term of economize mass and volumes. After the inflation the reservoirs would be used in the normal configuration with Titan atmosphere.

Looking to the composition of the moon atmosphere, another improvement can be analyzed: the Methane concentration on Titan is at the lower side of the flammable mixture, so it is possible to modify the actuator in order to transform it in a combustion engine. This

improvement of the thruster should be better analyzed, but a preliminary design should implement a small oxygen tank near the second stage of the vessel and some heaters to avoid the methane liquefaction. In case of small requests of thrust, the actuator could work like a cold gas actuator, with the analyzed performances, but if the required thrust would be higher, the oxygen tank would give the oxidant to transform the actuator in a sort of pulse jet engine.

References

[1] Marco Chiaradia, Gabriele Rodeghiero, Lorenzo Olivieri et al., "SCRAT Student Experiment Document", University of Padova, 2010

[2] NASA/CR—2004-213345, *Airships for Planetary Exploration*, Anthony Colozza Analex Corporation, Brook Park, Ohio

[3] Marco Chiaradia, Gabriele Rodeghiero, Lorenzo Olivieri et al., "Analysis of data retrieved by the SCRAT experiment during the 2010 ESA BEXUS 10 Campaign", University of Padova, 2011

[4] "MIR – Montgolfiere Infra Rouge", 20th ESA SYMPOSIUM ON EUROPEAN ROCKET AND BALLOON PROGRAMMES AND RELATED RESEARCH, Poster Session.

[5] Robert H. Brown, Jean-Pierre Lebreton, J. Hunter Waite, (Editors), *Titan from Cassini-Huygens*, 2009

Functional analysis and Rapid Manufacturing of re-designed nanosatellites

F. De Crescenzio, M. Fantini, F. Piergentili, F. Persiani
DIEM, II Faculty of Engineering (Forlì - FC), University of Bologna, Italy

F. Santoni
School of Aerospace Engineering, University of Rome "La Sapienza", Italy

Keywords: *Rapid Prototyping, Small satellites, Design for Manufacturing*

Abstract

This paper presents the design, the functional analysis and the manufacturing of EPSILON (Experimental Plastic Satellite. Innovative, Light, Off-the-shelf, Nano), a cube shaped nanosatellite. The aim is to demonstrate that rapid prototyping techniques, which are being diffused in the design and development of industrial products, can be effectively applied to re-engineer and manufacture nanosatellites with complex structures and tight mission requirements.

EPSILON's main body has been completely re-designed to be entirely built in plastic. It integrates the interfaces that host the payload and the subsystems following a DFM (Design for Manufacturing) approach. A FEM analysis was also carried out to evaluate the functionality of such system and a first prototype has been manufactured by FDM (Fused Deposition Modeling) technique.

As a result, building the main body of the satellite in a single plastic part, time, costs and weight can be significantly reduced compared to conventional processes, based on machining and connecting several aluminum machined parts.

1 General Introduction

Rapid prototyping has been originally used to create physical prototypes directly from CAD data early in the product's development cycle to reduce the Time to Market [1][2]. Such additive technique allows designers to produce fully working prototypes of new products to test ergonomics and functionalities before starting the traditional manufacturing process [3]. Rapid prototyping techniques can be also exploited for the direct manufacturing of a final products, being considerably more cost and time effective than traditional approaches under a certain number of units [4].

Nevertheless, whereas rapid prototyping and direct manufacturing are widely in use in the industrial field [5], in literature there are only a few examples of their effective application to produce the complex structures in aerospace industries [6]. Chang et. al use SLA (Stereolitography) samples to optimize the design of aircraft torque tubes in a Concurrent Design Approach [7]. Concerning the satellite, it has been illustrated that using SLS (Selective Laser Sintering) to create a prototype, but not the final product, of a small satellite structure,

allows assembly and design problems to be detected earlier [8].

In the aerospace field, the emerging small satellites market has been developing since the 1990s making possible space missions also to satisfy academic research and scientific needs [9][10][11][12].

Nowadays, these space systems are built by maximizing the use of existing components and commercial off-the-shelf technologies while minimizing development efforts. Therefore, small satellites are being widely developed given their functional and operational characteristics and their low costs. One of the reasons for choosing small satellites is also that the launch cost is proportional to their weight (a classification of small satellites according to their mass is shown in Table 1).

Classification	Mass
Mini	100 – 500 kg
Micro	10 – 100 kg
Nano	1 – 10 kg
Pico	0.1 – 1 kg

Table 1 – Classification of small satellites according to their mass

The aim of this project is the design optimization and the construction of EPSILON (Experimental Plastic Satellite: Innovative, Light, Off-the-shelf, Nano), a cubic-shaped nanosatellite (10 cm per side), weighing about 1 kg with a plastic main body completely manufactured by rapid prototyping technique. Such a manufactured structure would be the final product to be directly experimented in the real space environment.

The design of the case has been developed in order to fit the CubeSat standards and the NARCISO mission requirements that will be described in the following sections. Moreover, it has been conceived considering the capabilities of FDM (Fused Deposition Modelling) RP Technology (Figure 1).

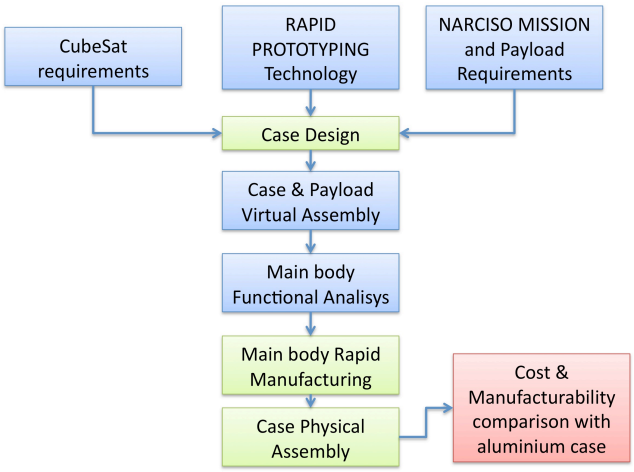


Figure 1 The design-flow of EPSILON

The rapid prototyping technology has been adopted in this project for two main reasons. The first reason is related to the geometric complexity of the satellite's structure that would be nearly impossible to produce by traditional CNC machines as a single element. The second one regards the weight of the structure that should be kept as low as possible to reduce the cost of the launch. This approach leads to a completely new design of the case that hosts the payload and the electronic boards. The case has been verified through a virtual assembly and a functional analysis. The main body of the satellite case has been manufactured and a physical assembly has been performed. Finally, a cost and manufacturability comparison with an aluminum case is provided.

2 EPSILON (Experimental Plastic Satellite: Innovative, Light, Off-the-shelf, Nano)

EPSILON can be considered a standalone nanosatellite because it respects the CubeSat standards and can be launched by the standard CubeSat dispenser (www.cubesat.org).

The CubeSat Project was developed by California Polytechnic State University, San Luis Obispo and Stanford University's Space Systems Development Lab with the primary

mission to provide low cost access to space for small payloads [13][14][15].

The first CubeSats were manufactured by Stanford University students improving their curriculum and capabilities by hands-on education and providing them with the opportunity to work on a low cost real space mission. The project spread through universities and research centers all over the world and a few years after the first idea, CubeSats manufacturing arose [16][17]. CubeSats are cube shaped picosatellites with a nominal length of 100 mm per side. General features of all CubeSats are that each single satellite may not exceed 1 kg mass and the center of mass must be within 2 cm of its geometric center. The structure of the CubeSat must be strong enough to survive maximum loading defined in the testing requirements and cumulative loading of all required tests and launch. The CubeSat structure must be compatible with the standard launcher interface of the flight proven deployment system Poly Picosatellite Orbital Deployer or P-POD, [18][19][20]. This means that CubeSat should be provided with four longitudinal rails which must be smooth, and edges must be rounded to a minimum radius of 1 mm.

Electronic systems must be designed taking into account that no electronics may be active during launch to prevent any electrical or RF interference with the launch vehicle and primary payloads. Moreover, CubeSats with rechargeable batteries must be fully deactivated during launch or launched with discharged batteries.

Compared to the previous projects based on CubeSat standards, the main contribution of this work deals with the innovative approach for manufacturing the plastic structure of EPSILON by means of FDM rapid prototyping technique. Moreover, the application presented in the paper refers to the design of an EPSILON used as payload of a larger microsatellite: UNISAT-5.

In early nineties the Group of Astrodynamics of the "Sapienza" University of Rome (GAUSS), established the UNISAT program at the Rome School of Aerospace Engineering with the aim to design, manufacture and launch small educational satellites, thus involving students, researchers and professors in a real space project. In the framework of this program four satellites have already been launched (respectively in 2000, 2002, 2004 and 2006) from Baykonour Cosmodrome using the DNEPR launch vehicle, [21][22].

The efforts to reduce costs, weight and size have always aimed to adapt innovative, terrestrial Commercial Off-The-Shelf (COTS) technologies to the space environment.

The UNISAT experience showed that it is possible to use terrestrial technologies in orbit (UNISAT-3 launched in June 2004 is still operative), including solar cells, electronic components and sensors. This achievement together with the recent development in rapid prototyping industrial techniques led us to consider this new approach in designing a satellite structure configuration suitable for the realization of a low cost satellite. Moreover, it can also be used as part of a fleet or a formation of satellites being easily adaptable to different missions.

As an application case study, this paper proposes a coordinated mission with a UNISAT series satellite, to which EPSILON is connected by a tether or a boom, as described in the next section.

3 NARCISO mission and payload

In the configuration presented in this paper UNISAT-5 and EPSILON will take pictures of each other in order to evaluate aging and possible damage to EPSILON's external structure and UNISAT-5 solar arrays due to space environment exposure. Both satellites create the NARCISO mission, named after the well-known mythological character Narcissus, who loved mirroring himself in the lake water. The orbit is LEO (Low Earth Orbit) sun synchronous ranging between 500 km and 800 km with thermal cycle between -20 and 60 °C.

In this configuration the EPSILON satellite will communicate with the UNISAT-5 mother satellite and behaves as a tip mass. The proposed mission is experimental and makes it possible to test this kind of satellite in a real

space mission. The EPSILON satellite provides attitude stabilization to the main satellite exploiting gravity gradient features. The separation from the main satellite is achieved by exploiting one spring which is released by thermal cut.

The EPSILON satellite main payloads are a camera to take pictures of the UNISAT-5 satellite and a magnetometer to measure the Earth's Magnetic Field far from UNISAT satellite influence. At the same time a camera boarded on the UNISAT-5 satellite and oriented towards the EPSILON satellite will take pictures of the experiment, providing an effective evaluation of the ABS structure and solar cells in orbit aging.



Figure 2 – UNISAT-4 Camera

The camera proposed is the same boarded on previous UNISAT-4 and is sketched in Figure 2. It is a COTS camera usually implemented for terrestrial industrial applications. The field of view and resolution can be changed by using different optics. The camera can be programmed by the user by means of a simple "High level" visual programming block diagram environment. It measures 6x6 cm in size and 50 g in weight. Power consumption in operative mode is 1.5 W. The magnetometer on board EPSILON is an AP539 fluxgate magnetometer by Applied Physics Systems.

4 EPSILON design

Rapid Prototyping allows the manufacturing of complex shape models which cannot be achieved through conventional subtractive processes. The chance to manufacture the whole satellite structure in a single work session, without the usual metal constraints due to processing procedures and without any additional costs for small accurate details, provides best performances in the design of small satellites, which are usually very difficult to assemble. A machined aluminum structure would be composed by several parts to be assembled together and would also increase both costs and building time (much longer is needed for machine programming and set-up). Moreover, in terms of mass (density of ABS = 1.05 g/cm³; density of aluminum = 2.7g/cm³), an ABS structure is usually much lighter than an aluminum one leading to a saving of more than 50% in weight.

Based on these considerations, EPSILON's main body structure was designed to be completely manufactured in ABS as a single element (Figure 3).

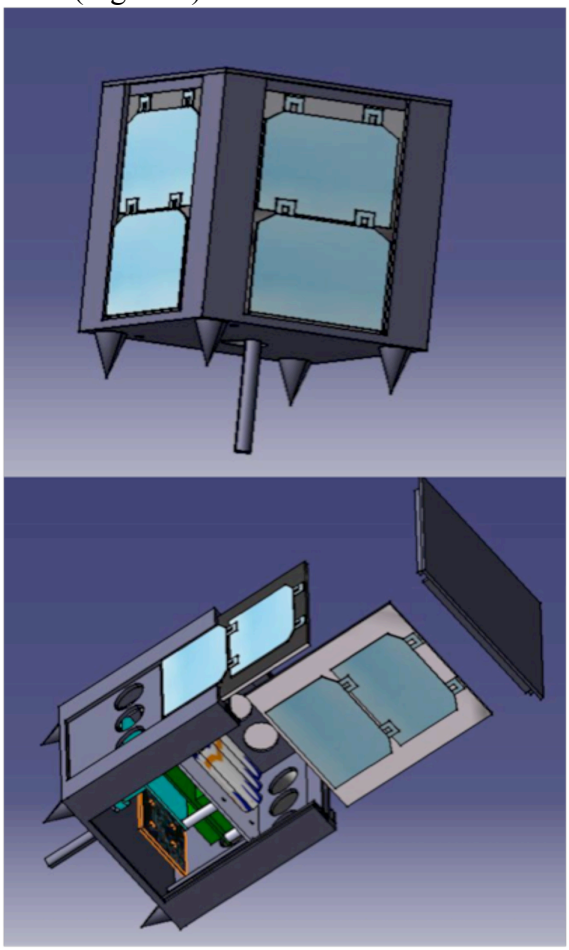


Figure 3 – EPSILON assembled view (up) and disassembled view (bottom)

The four solar arrays can be inserted into plastic tracks; they are blocked by the upper plate, which is screwed to a small aluminum

plate placed in an internal track. This allows the structure to be completely closed by means of four screws. The lower plate hosts four cone-shaped tools to interface with the mother satellite and an ABS cylinder to be connected to the boom edge. The internal layout of payloads is given by a single block comprising batteries, electronic board, camera and magnetometer. The whole block can be assembled and tested autonomously before being installed on the satellite.

All parts are screwed to a single central aluminum plate, guaranteeing thermal inertia to the batteries. Electronic boards, camera and magnetometer are screwed to the aluminum plate. The radio is integrated into the electronic board and is placed on the bottom part of the aluminum plate (towards the mother satellite) to avoid transmission interference.

The whole part is fixed by four screws to a plastic support. A hole is drilled in the lower part of the EPSILON's structure to allow the camera an external view.

EPSILON can be fixed to the main UNISAT satellite through a boom 240 cm long. The boom is essentially a harmonic steel spring, already used in previous UNISAT payload manufacturing deployed by means of thermal cut [23].

The boom can be substituted by a tether for tether dynamical motion experiments or can be completely removed; when removed, EPSILON is free to move and experiments can be performed on satellite formation flying.

The configuration with boom deployed provides an increase in the moments of inertia of the mother satellite associated to axes perpendicular to the boom itself. The moments of inertia increase is equal to the square of the boom length (about 2.5 meters) per EPSILON weight (about 1 Kg); this means an increase in transverse inertia moments of about 6.25 Kgm^2.

The satellite inertia moments on transverse axes value were evaluated to about 0.2 Kgm^2, and 0.3 Kgm^2 along the boom axis. Thus new moments of inertia achieved by boom deployment are 6.45 Kgm^2 (transverse axes) and 0.3 Kgm^2 (boom axis).

The EPSILON on-board data handling system is based upon a microprocessor controlling the camera and magnetometer payload through an asynchronous serial connection. It monitors the satellite operations, measuring batteries and solar array voltage, current and temperature. Images, magnetometer and sensor data are collected in a flash memory and transmitted when commanded by the UNISAT-5 mother satellite. In this way the whole EPSILON is seen as a single peripheral payload by the UNISAT-5 satellite.

Communications between EPSILON and UNISAT-5 are performed using two S-band transceiver modules, with integrated modem and antenna. The modules' size is 3x4 cm and they can be directly soldered on printed circuit boards as normal integrated circuits. An identical module is hosted on board the UNISAT-5 satellite to assure the communication. All of the electronic components are commercial off the shelf, with operative temperature ranging from -20 to 80 degrees.

The EPSILON power system is based upon triple junction, low efficiency solar cells and NiCd batteries. Two serial solar cells are glued to movable carbon or glass fiber supports. Only the four lateral panels and the upper one are provided with solar cells, whereas the bottom tray, where the separation device is positioned, has no solar cells.

Low efficiency triple junction solar cells have an efficiency of about 22%. These cells and the panelization procedure have already been tested on board UNISAT-3 showing their capability to withstand environmental conditions with launch loads and in orbit [24].

The solar cells are connected, through a blocking diode to three series-connected NiCd (600mAh) batteries.

Each solar array is capable of providing a power given by:

$$P = \eta\, Q\, A$$

With a solar surface, A, of (26 cm^2), a solar flux constant of 1 365 W/m^2 and an efficiency, η, of 22%, a total power of about 0.78 W per cell is generated. Similar values have been achieved by the analysis of UNISAT-3 in orbit

collected data [25]. This leads to a single solar array maximum power of about 1.56 W.

Subsystem	Power [W]	T/T_{orb}	Power consumption [W]
Electronics	**1**	**0.6**	**0.6**
Communication	**1**	**0.2**	**0.2**
Camera	**1.5**	**0.02**	**0.03**
Magnetometer	**1**	**0.1**	**0.1**
Average Power consumption			**0.93**

Table 2 – EPSILON power consumption

The EPSILON power consumption is sketched in Table 2.

These values are obtained considering that the electronics has been realized to be switched on only under sun illumination (about 60% of the orbit), communication data transmission lasts about 20 minutes, the camera obtains one picture per orbit and is switched on for 2 minutes, while the magnetometer obtains data with a frequency of 0.1 Hz.

The EPSILON bottom tray always faces the Earth due to gravity gradient attitude stabilization and then it is illuminated by the Sun only close to eclipses for short times. The worst illumination condition is achieved when only one satellite panel is directly exposed to the Sun. In this configuration the generated power is 1.56 watt.

The power budget shows that the worst illumination condition guarantees a margin of about 0.5 W.

Batteries are needed to supply the system peak power when the radio is transmitting or the camera is taking pictures. The maximum instantaneous power consumption reaches 3.5 W, but this condition lasts no more than two minutes corresponding to 18 mAh of battery discharge. During transmission with camera off, for twenty minutes, power consumption is about 2 W corresponding to battery discharge of about 45 mAh. Thus the total battery discharge per orbit is 63 mAh leading to a battery depth of discharge of 10%, guaranteeing more than three years of operative battery life.

5 Manufacturing

The manufacturing process started converting the CAD model of the main body in STL format, which is the standard interface used in the rapid prototyping field.

The tessellated triangular mesh (Figure 4) has a sufficient number of faces to describe the shape of the model surface (672 038 points and 1 344 160 triangular faces).

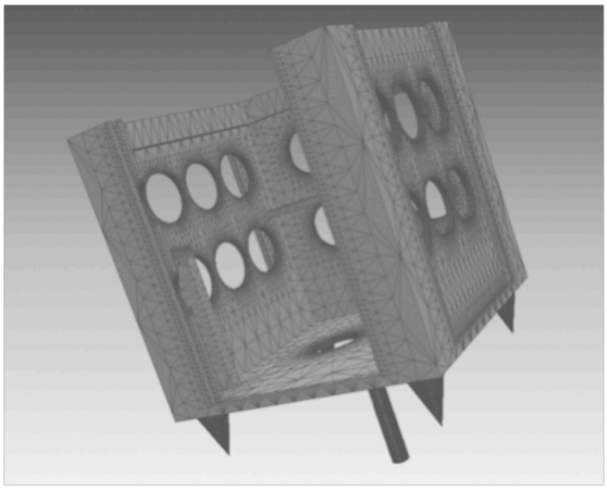


Figure 4 – EPSILON STL model

The physical model of the EPSILON's structure was then generated by means of Stratasys Dimension SST (Soluble Support Technology). This rapid prototyping system is based on the FDM technique. It deposits fused ABS P400 jet (Acrylonitrile Butadiene Styrene plastic material) and soluble support material to sustain the prototype under construction. An automated support removal process for hands-free model completion is also provided. It is an agitation system using hot water and soap bath (at 70 °C) to automatically wash away the support structures. The layer thickness is 0.254 mm and the maximum build size is 203 x 203 x 305 mm (height).

The ABS filament is fed through a heating head in a semi-molten state and then extruded through a nozzle. Layer by layer, the semi-molten filament is deposited onto the partially constructed model. During the cooling process in the working room, the semi-molten filament adheres to the preceding layer already deposited. The extrusion temperature of ABS is

usually 270°C while the envelope temperature (the temperature of the air in the working room around the model) is 70°C. Therefore, it can be assumed that these kinds of models could survive LEO (Low Earth Orbit) thermal environment (60 degrees maximum) for a time compatible with the usual low cost satellite lifetime.

The Dimension SST is also provided with a pre-processing software (Catalyst) that makes it possible to import the STL file, orient and slice the model, generate any needed support structures and calculate the head's path to build up the part. Since the mechanical behavior of the prototype manufacture by additive FDM technology depends on the building direction, in particular by layer-to-layer adhesion, two different orientations of the STL model during the building process were simulated: vertical (building direction along longitudinal axis) and horizontal (building direction along transversal axis).

This system allows two different fill options: solid and sparse. In the first case, each cross-section through the model is filled with ABS material. In the second one, the interior part of the model is not filled solid. The model is built up with thin inner and outer walls while the interior part is replaced with a kind of honeycomb structure. Since the sparse interior structure results entirely enclosed when the model is completed, there is no visible difference in a prototype built with these two different options.

As a result of these two different approaches, solid fills are stronger and heavier while sparse fills are weaker and lighter. Moreover, the sparse fill option saves on material and therefore speeds up the construction process.

As a consequence, the STL model of the EPSILON's structure was processed by Catalyst with both fill options (solid and sparse) for each orientation of the STL model, and the results are compared in Table 3.

The environmental loads that EPSILON will encounter during its lifetime are mainly directed along longitudinal axis, therefore, the horizontal position (building direction along transversal axis) was chosen during the building process. This choice assures the highest tensile strength along longitudinal axis. Regarding to the fill options, since the structure is mainly characterized by thin walls, there is no appreciable difference between them in terms of both material consumed (and therefore material cost and weight) and build time. For that reason, solid fill option was preferred to obtain a stronger structure. This option also reduces out gassing due to the vacuum environment in space.

Fill Option	Model Material *	Support Material	Build Time	Weight est. densitv	Material Cost
Vertical position (building direction along longitudinal axis)					
Solid	172.96 cm³	146.00 cm³	26 hr 59 min	181.608 g	74.82 €
Sparse	164.46 cm³	146.00 cm³	26 hr 31 min	172.683 g	71.97 €
Saving %	4.91 %	//	1.73 %	4.91 %	3.81 %
Horizontal position (building direction along transversal axis)					
Solid	174.16 cm³	141.40 cm³	24 hr 39 min	182.868 g	77.44 €
Sparse	154.35 cm³	141.40 cm³	23 hr 26 min	162.068 g	73.17 €
Saving %	11.37 %	//	4.94 %	11.37 %	5.51 %

* small variation of consumed model material in solid fill option between vertical and horizontal position are due to the additive building process based on layer thickness of 0.254 mm.

Table 3 – Comparison between solid and sparse fill

The digital model of the structure (bounding box dimensions: 160 mm high, 100 mm wide, 100 mm deep) was then directly prototyped in a single work session with 394 slices. Completion of the build process took 24 hr 39 min while 174.16 cm³ of ABS and 141.40 cm³ of support material were consumed (≈78.00 € is the total cost under educational conditions). The process was then completed by washing the model to remove all the support material. The weight of the plastic structure is 182.868 g and compared to the estimated weight of the same structure manufactured in aluminum (466.992 g) a load reduction of more than 60% is observed.

Finally, this first prototype of the EPSILON's structure was used for the subsystem integration test and for fit check purposes. Figure 5 shows the assembly with the four solar arrays inserted into the designed holding tracks.

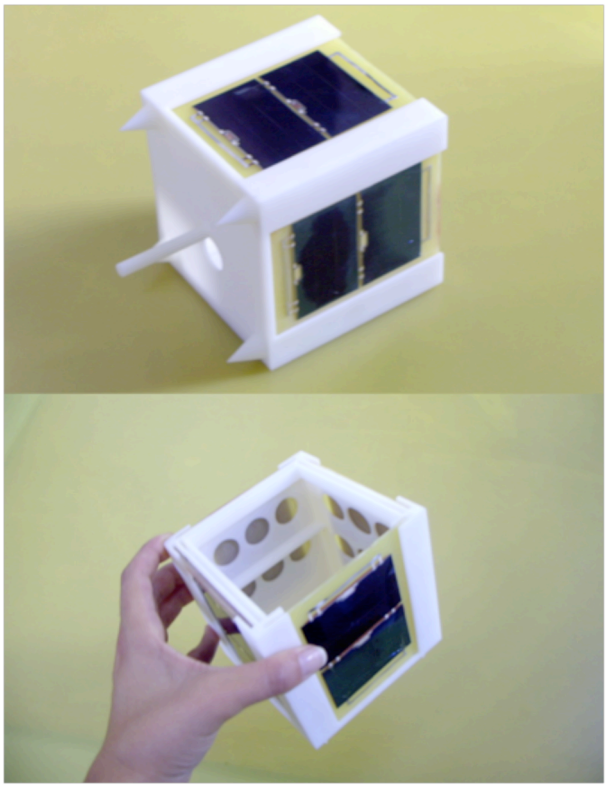


Figure 5 – EPSILON prototype external view (up) and internal view (bottom)

6 Functional Analysis

The functional requirements to be fulfilled by this type of space systems are related to the environmental loads it will encounter during its lifetime.

Loads for space systems in Low Earth Orbit essentially depend on the static and dynamic loads over the structure during the launch phase and to vacuum and solar flux during the orbital phase.

Regarding the static and dynamic loads, a FEM analysis was carried out in order to verify the functional constraints of the system designed.

It is well known that parts fabricated using FDM technology reveal anisotropic properties due to the additive fabrication process and show the lowest tensile strength along the building direction caused by layer-to-layer adhesion. By this reason, the results of FEM analysis depend on prototype orientation during the layer-by-layer deposition step. To simplify FEM analysis as well as to assure maximum conservative results, the ABS prototype was assumed to be isotropic and the worst mechanical characteristics among the ones evaluated by Bellini and Güçeri were chosen [26].

According to the launcher manual, static loads are due to acceleration, which can have a maximum value of 6g (about 60 m/s^2) [27].

Mass distribution adopted for static analysis simulation is sketched in Figure 6-a. Masses were distributed on the upper surface of the EPSILON satellite structure and on the supports for the internal components. The launcher acceleration is directed as shown by the arrows along z-axis from bottom towards up.

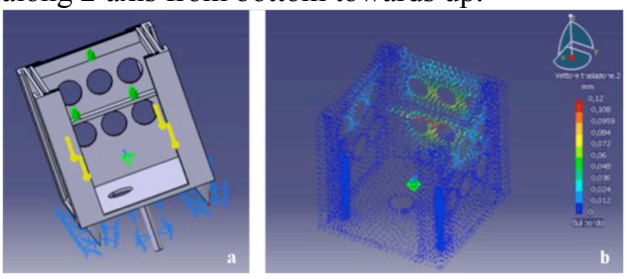


Figure 6 a) Mass distribution and launch acceleration direction. b) Shift field results by static loads effect analysis

On the basis of such masses and acceleration distribution the shift field was evaluated and the results are sketched in Figure 6-b, the maximum shift is limited to 120 μm.

The vibrational requirements are related to the structure vibration frequency. In particular, according to the launcher manual the lowest resonant structure vibration frequency must be greater then 20 Hz.

For this purpose a FEM analysis was carried out to evaluate the first ten vibration frequencies.

Structure constraints were simulated by joints on the cones which are blocked on the mother satellite in launch configuration and on the cylinder which is connected to the boom.

The first ten resonant vibration frequencies of the EPSILON main structure are sketched in Table 4. It can be observed that the lowest frequency is 194 Hz so it is about one order of magnitude larger than that required by the launcher. This is sufficient to state that the satellite structure manufactured in ABS is fully compliant with launcher requirements, thus it is able to survive to launch static and vibration loads.

Mode	Frequency Hz
1	194.1
2	203.3
3	402.8
4	407.4
5	457.6
6	491.6
7	515.2
8	547.1
9	630.2
10	635.5

Table 4 – First ten resonant vibration frequencies of the EPSILON main structure

Vacuum survivability was tested by keeping a few ABS samples (with different shapes and realized with the solid fill option) in a vacuum chamber, used for space environment simulation, for three days. No cracks or holes due to outgassing were noticed on the samples, thus results of the vacuum survivability seem to confirm the feasibility of a Cubesat in ABS.

The thermal loads seem to be not significant because the temperature, measured by nanosatellites in Low Earth Orbit, ranges from -20°C to 60°C [28]. The maximum temperature is lower than the ABS Heat Deflection Temperature (HDT) that is 76°C at 264 psi and 90°C at 66 psi (Glass Transition (Tg) temperature is 104°C) [29]. These values are conservative enough since in Low Earth Orbit the pressure is about 10^{-8} Pa. However, one of the aims of the planned mission with EPSILON is to evaluate the actual behavior of the ABS material due to Low Earth Orbit thermal cycling.

All of these considerations on system functional analysis seem to confirm the feasibility of the proposed system and its capacity to survive typical space systems loads.

7 Impact of System Architecture and Manufacturing on System Design

In the design of experimental aerospace applications several requirements need to be met. Concerning EPSILON, besides the functional requirements previously described, three main design factors were also identified. These are, respectively, system reliability, system manufacturability and the cost of production.

Since EPSILON is one module of a mission system and considering the difficulties in launching experimental satellites, the system must be designed to minimize the causes of failures. For this reason the architecture of EPSILON is strongly integrated since there is one single part, the main body, performing a huge number of functions: case, carter, support for external and internal components, connection tracks, etc. It is known that by reducing the number of parts, the number of interactions between parts is also reduced, avoiding both the predictable and the incidental interactions which can cause failures. The external case of EPSILON is composed of 10 parts, including the connection elements:

- 1 main body;
- 4 solar arrays;
- 1 upper panel;
- 4 screws.

By comparing this plastic structure to the more complex system obtained by manufacturing EPSILON's main body in aluminum, a reduction in number of components of about 70 % can be achieved. Consequently, the complexity of the interfaces between modules and parts is significantly decreased. This concept also allows excellent product performances to be achieved, such as compactness and lightness in the EPSILON's structure. Moreover, the construction of the EPSILON's main body as a unique ABS part presents several advantages in terms of

manufacturability. The system design follows the majority of design principles in the Design for Manufacturing (DFM) guidelines list [30]. Reducing the total number of parts is the first rule presented in these guidelines, since the main challenge is considered to be the reduction of manufacturing costs, process time, engineering time, project handling and, in general, the complexity of production and assembly procedures.

In order to measure the efficiency of the EPSILON's design in terms of costs, Figure 7 shows the comparison of the total direct cost needed to produce the EPSILON's structure by FDM technique and by the traditional approach, mainly based on the connection of aluminum machined parts.

In detail, in a simplified cost model for manufacturing EPSILON's structure in aluminum there is an amount of fixed costs, given by the CAM (Computer Aided Manufacturing) procedures, which comprises the software and the engineering costs, and the machine set up costs. The variable costs are given by the machine operational cost, the cost of the material and the assembly cost. For the ABS version the fixed costs can be approximated to 0 since the RP operations are highly automated, except the assisted step to find the best orientation of the STL model before allowing the slicing process to start.

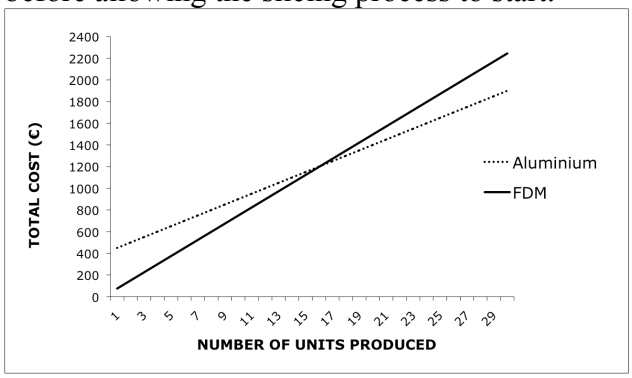


Figure 7 – Total manufacturing costs for number of units produced

Finally, the total cost and the time needed for producing one or few units are lower than the estimated cost and time when using alternative technologies. This is true for a limited number of units, as needed in this experimental application.

To design for ease of fabrication is a further rule presented in the DFM guidelines and the use of rapid prototyping allows minimization of the overall manufacturing cost also because, in this specific kind of application, excessive tolerances and finishing machining operations are not needed in the output of the FDM machine.

Therefore, manufacturing the satellite by means of rapid prototyping techniques has a significant impact on the overall mission performance and not only on the single system performances, such as the low weight and the compact architecture.

In the field of rapid prototyping, besides Fused Deposition Modeling (FDM), other kinds of layer manufacturing processes are available, such as stereolithography (SLA), selective laser sintering (SLS), solid ground curing (SGC), three-dimensional printing (3DP), and laminated object manufacturing (LOM). Several benchmarking studies were carried out by comparing costs and properties of different RP technologies [31][32].

Results showed that the technologies typically considered appropriate for use as low-cost 3D printers are FDM, LOM, and 3DP. Whereas 3DP and LOM are advantageous in terms of material costs, these processes produce prototypes with lower mechanical properties than the FDM approach. On the other hand SLA, SLS and SGC can produce prototypes with better mechanical properties but also feature a more expensive process.

The relatively low cost of FDM systems and their ability to build one small part at a time economically are important elements. An additional reason for choosing this technology is the mechanical properties of the ABS thermoplastic material used in the FDM process. Therefore, FDM can be considered a good balance between costs and mechanical properties.

Other possible alternatives for the production of such a system in the aerospace industry are based on nanomaterials. For example, high strength-to-weight ratio and multifunctionality, carbon nanotube reinforced polymer composites may provide a valid solution for manufacturing

the EPSILON's structure [33][34]. The most obvious obstacle in using this kind of technology is the cost that would not be compatible with the low budget available for this kind of experimental aerospace nanosatellite.

8 Conclusions

This paper describes the development of a nanosatellite whose main body structure is designed to be completely manufactured in plastic by means of the FDM technique.
The first prototype of the EPSILON's structure has already been manufactured and used to test the integration of all subsystems that will be hosted on board. After minor design improvements needed to better refine the whole nanosatellite, the final version of the plastic main body structure will be produced by the same process. Once boarded with all necessary components, EPSILON will then be ready to be launched in orbit.
It is important to recall that the final (and most important) test will be performed in space in real operative conditions according to the NARCISO mission when EPSILON is launched with the UNISAT-5 satellite. This kind of mission was planned also to evaluate the behavior of EPSILON in LEO (Low Earth Orbit) radiation and a thermal environment since the camera boarded on UNISAT-5 and oriented towards EPSILON will collect pictures to provide an effective evaluation of the plastic structure aging.
The approach described in this paper could lead to innovative manufacturing processes of small satellite structures opening new perspectives in the aerospace field. In this way, it would be possible to manufacture the whole plastic structure of small satellites in a very short time with an effective reduction of both cost and weight with respect to the traditional manufacturing processes of aluminum machined structures. Moreover, it would also be possible to realize a fleet of small satellites easily, based on the same modular structure.

References

[1] Ulrich, K. T. and Eppinger, S. D. (2004), *Product Design and Development*, Third Edition, McGraw-Hill.

[2] Guangchun, W, Huiping, L., Yanjin, G. and Guoqun, Z. (2004), "A rapid design and manufacturing system for product development applications", *Rapid Prototyping Journal*, Vol. 10, No. 3, pp. 200-206

[3] Evans, M. A. and Campbell, R. I. (2003), "A comparative evaluation of industrial design models produced using rapid prototyping and workshop-based fabrication techniques", *Rapid Prototyping Journal*, Vol. 9, No. 5, pp. 344-351.

[4] Hopkinson, N. and Dickens, P.M. (2001), "Rapid prototyping for direct manufacture", *Rapid Prototyping Journal*, Vol. 7, No. 4, pp. 197-202.

[5] Tennyson, S., McCain, G., Hatten, S. and Eggert, R. (2006), "Case study: promoting design automation by rural manufacturers", *Rapid Prototyping Journal*, Vol. 12, No. 5, pp. 304-309.

[6] Rochus, P., Plesseria, J.-Y., Van Elsen, M., Kruth, J.-P., Carrusc, R. and Dormalc, T., (2007), "New applications of rapid prototyping and rapid manufacturing (RP/RM) technologies for space instrumentation", *Acta Astronautica*, Vol.61, Issue 1-6, pp.352-359. doi:10.1016/j.actaastro.2007.01.004

[7] Chang, K.-H. and Bryant, I.H., (2004), "Concurrent design and Manufacturing of Aircraft Torque Tubes", *Journal of Materials Processing Technology*, Vol. 150, Issue 1-2, pp 151-162. doi:10.1016/j.jmatprotec.2004.01.033

[8] Quincieu, J., Robinson, C., Stucker, B. and Mosher, T. (2005), "Case study: selective laser sintering of the USUSat II small satellite structure", *Assembly Automation*, Vol. 25, No. 4, pp. 267-272.

[9] Rycroft, M. and Crosby, N. (2002), Smaller Satellites: Bigger Business?: Concepts, Applications and Markets for Micro/Nanosatellites in a New Information World, Springer.

[10] Barnhart, D.J., Vladimirova, T. and Sweeting, M.N. (2007), "Very-Small-Satellite Design for Distributed Space Missions", *Journal of Spacecraft and Rockets*, Vol. 45, No. 1, pp. 1294-1306.

[11] Alger, M.W.R. and Kumar, K.D., (2007), "A design and development of Pico- and Femto-satellites", *International Journal of Manufacturing Research*, Vol. 3, No. 1, pp. 48-79.

[12] Kramer, H. and Cracknell, A., (2008), "An overview of small satellites in remote sensing", *International Journal of Remote Sensing*, Vol. 29, No. 15, pp. 4285-4337.

[13] Heidt, H., Puig-Suari, J., Moore, A.S., Nakasuka,

S. and Twiggs, R.J. (2000), "CubeSat: A New Generation of Picosatellite for Education and Industry Low-Cost Space Experimentation", *Proceedings of the 14th AIAA/USU Conference on Small Satellites, Logan, UT, USA, August, 2000*.

[14] Puig-Suari, J., Turner, C. and Ahlgren, W. (2001), "Development of the standard CubeSat deployer and a CubeSat class PicoSatellite", *Aerospace Conference Proceedings*, 2001, IEEE, Vol. 1, pp. 347-353.

[15] Nason, I., Puig-Suari, J. and Twiggs, R., (2002), "Development of a family of picosatellite deployers based on the CubeSat standard", *Aerospace Conference Proceedings*, 2002, IEEE, Vol. 1, pp. 457-464.

[16] Funase, R., Takei, E., Nakamura, Y., Nagai, M., Enokuchi, A., Yuliang, C., Nakada, K., Nojiri, Y., Sasaki, F., Funane, T., Eishima, T. and Nakasuka, S. (2007), "Technology demonstration on University of Tokyo's pico-satellite "XI-V" and its effective operation result using ground station network", *Acta Astronautica*, Vol. 61, No. 7-8, pp. 707-711.

[17] Akagi, J.M., Tamashiro, T.N., Iwami, R.T., Akagi, J.T., Shiroma, W.A. and Cardenas, J.M. (2008), "CubeSat-Based Disaster Detection and Monitoring Systems", *Proceedings of the 6th Responsive Space Conference, Los Angeles, CA, USA, April–May, 2008*.

[18] Nason, I., Creedon, M. and Johansen, N. (2002), "CUBESAT P-Pod Deployer Requirements", available at: http://cubesat.atl.calpoly.edu/media/Documents/Launch%20Providers/ppod_mk1_icd.pdf (accessed 12 December 2008).

[19] Toorian, A. (2004), "Poly Picosatellite Orbital Deployer Interface Control Document" available at: http://cubesat.atl.calpoly.edu/media/Documents/Launch%20Providers/ppod_mk2_icd.pdf (accessed 12 December 2008)

[20] Lan, W. (2007), "Poly Picosatellite Orbital Deployer Mk III ICD", available at: http://cubesat.atl.calpoly.edu/media/P-POD%20Mk%20III%20ICD.pdf (accessed 12 December 2008).

[21] Graziani, F. and Santoni, F. (2002), "The UNISAT program for space education", Keynote Invited Paper, Proceedings of the 23rd International Symposium on Space Technology and Science, Matsue, Japan, May–June, 2002.

[22] Santoni, F. and Graziani, F. (2002), "Space Education using University Microsatellites", Proceedings of the 53rd International Astronautical Congress, Houston, TX, USA, October, 2002.

[23] Piergentili, F. and Graziani, F. (2006), "SIRDARIA: A low-cost autonomous deorbiting system for microsatellites", *Proceedings of the 57th International Astronautical Congress, Valencia, Spain, October, 2006*.

[24] Santoni, F. and Piergentili, F. (2008), "Analysis of the UNISAT-3 Solar Array In-Orbit Performance", *Journal of Spacecraft and Rockets*, Vol. 45, No. 1, pp. 142-148.

[25] Santoni, F. and Piergentili, F. (2006), "In orbit performances of the UNISAT-3 solar arrays", *Proceedings of the 57th International Astronautical Congress, Valencia, Spain, October, 2006*.

[26] Bellini, A. and Güçeri, S. (2003), "Mechanical characterization of parts fabricated using fused deposition modeling", *Rapid Prototyping Journal*, Vol. 9, No. 4, pp. 252-264.

[27] ISC Kosmotras (2001), "DNEPR Space Launch System User's Guide," Issue 2, November 2001, available at: http://snebulos.mit.edu/projects/reference/launch_vehicles/DNEPR/Dnepr_User_Guide.pdf (accessed 12 December 2008).

[28] Graziani, F., Santoni, F., Piergentili, F., Bulgarelli, F., Sgubini, M. and Bernardini, S. (2004), "Manufacturing and launching student-made microsatellites: "hands-on" education at the University of Roma", *Proceedings of the 55th International Astronautical Congress, Vancouver, Canada, October, 2004*.

[29] Stratasys Inc, (2007), "ABS datasheet" available at: http://www.stratasys.com/uploadedFiles/North_America/Materials/ABS07.pdf (accessed 12 December 2008).

[30] Chang, T.C., Wysk, R.A. and Wang, H. P. (1998), *Computer-Aided Manufacturing*, Second Edition, Prentice Hall.

[31] Kim, G.D. and Oh, Y.T. (2007), "A benchmark study on rapid prototyping processes and machines: quantitative comparisons of mechanical properties, accuracy, roughness, speed, and material cost for direct manufacture", *Proceedings of the I MECH E Part B Journal of Engineering Manufacture*, Vol. 222, No. 2, pp. 201-215.

[32] Wendel, B., Rietzel, D., Kühnlein, F., Feulner, R., Hülder, G. and Schmachtenberg, E. (2008), "Additive Processing of Polymers", *Macromolecular Materials and Engineering*, Vol. 293, No. 10, pp. 799-809.

[33] Bal, S. and Samal, S.S. (2007), "Carbon nanotube reinforced polymer composites–A state of the art", *Bulletin of Materials Science*, Vol. 30, No. 4, pp. 379-386.

[34] Mylvaganam, K. and Zhang, L.C. (2007), "Fabrication and Application of Polymer Composites Comprising Carbon Nanotubes", *Recent Patents on Nanotechnology*, Vol. 1, No. 1, pp. 59-65.



CEAS 2011

Airspace Design and Optimization

3rd CEAS Air&Space Conference
21st AIDAA Congress

Operational Feasibility of Sectorless ATM

Marcus Biella, Bettina Birkmeier, Bernd Korn,
Christiane Edinger, Sebastian Tittel, and Dirk Kügler
DLR Institute of Flight Guidance (German Aerospace Center), Germany

Keywords: *sectorless ATM concept; airspace management; separation management; en-route efficiency; operational feasibility*

Abstract

Operational feasibility of a new air traffic management concept is assessed which might increase the efficiency of en-route operations by up to 100%. The concept follows the idea of a sectorless airspace organization in which one controller is responsible for n>3 *aircraft during their entire flight within one big airspace. Hypotheses, validation strategy and the real time simulation campaign are described.*

1 General Introduction

The current practice to organize the airspace and to balance demand and capacity is the partitioning of airspace into sectors. With the increase in air traffic a common practice has been to decrease the size of the sectors in order to limit the workload of the air traffic controllers. However, this approach has its natural limits. With the increasing number of sectors the number of coordination activities will increase as well. In addition, a smaller sector size reduces the possibilities for controllers for tactical and strategic control of aircraft. Two basic approaches are currently discussed to mitigate these problems. On the one hand, dynamic sectorization including an optimized process for partitioning the airspace into sectors according to the main traffic flows is pursuit to reduce the complexity since less crossing traffic can be expected. This can be beneficial especially for sectors feeding terminal areas of major airports. On the other hand, automation is introduced to reduce the workload of the controllers. E.g. for en-route sectors concepts for multi sector planning have been introduced. But it is not obvious yet whether or not this will enable the handling of the envisaged increase in air traffic within the next 15 years.

This contribution reports about a different approach. Within this concept, the entire en-route airspace (e.g. the entire German en-route airspace) is considered as one big airspace. The basic idea is that an aircraft that enters this airspace will be controlled by only one air traffic controller (ATCo) during its entire flight in this airspace (e.g. from TMA exit to TMA entry or from entry point into German airspace until its exit point). Following the statistics for the day in 2008 with the highest traffic load in Germany (23rd of September), in average on every hour of an ATCo on the controller working position, 1.77 hours of flights have been managed. Transferring this to our approach means that an ATCo who is controlling two

aircraft at the same time would be at least as efficient as within today's operation. Controlling more than three aircraft simultaneously would mean a dramatic increase in efficiency. First ideas about such a sectorless ATM-concept have been addressed in [1]. Here, the starting point was, that "instead of having two controllers controlling one sector containing n aircraft, one controller will be responsible for a limited number of m>2 aircraft, from departure to arrival terminal areas (TMA)". In a follow on activity, Riviere [2] proposed "… to provide a very simple route network and to improve it by using optimization techniques". Initial idea of [1] was to combine the sectorless ATM idea with the concept of an autonomous aircraft i.e. a trajectory-based, individual flight control. In the last case a sectorless controller working position (CWP) is required providing features like a Cockpit Display of Traffic Information (CDTI) combined with an Airborne Separation Assistance System (ASAS). Similar to free flight operations the question of how to resolve separation conflicts arises. Especially the coordination of actions of the sectorless controllers was identified to be critical. Initial studies of using VFR rules revealed that these rules seem to be insufficient for a proper conflict management. Therefore, Extended Flight Rules (EFR) should be applied which have been defined (see [3]) for a Free Flight scenario. In terms of the sectorless concept [1] several options have been discussed: (a) Extended Flight Rules, (b) fully automatic coordination of action and (c) a combined solution of procedural and automatic coordination. However, dedicated validation activities to assess the feasibility have not been conducted.

In a joint project call Airspace Management 2020 (LRM2020), DLR and DFS have further assessed the feasibility of this sectorless ATM concept [4], [5]. A deeper analysis of ATCo – a/c relationship [6] revealed that an average of three a/c per ATCo results into an equivalent efficiency as in current operations. A ratio of 1:6 would mean an increase of 100%. In a series of small workshops with controllers and ATM experts the initial concept has been refined and benefits have been anticipated. Special emphasis has been given to the new roles and tasks of the ATCos and the required support tools including their HMI integration. After being more familiar with the concepts, the HMI, and the support tools, controller stated that a ratio of 1 to 6 (meaning controlling six a/c at the same time) could definitely be possible. They as well stated that the aircraft under control do not need to fly in same area. Based on these positive results and feedbacks, next steps have been undertaken to further assess the operational feasibility. A two weeks simulation campaign with eight experienced air traffic controllers (upper area control Karlsruhe) has been carried out in November/December 2010.

2 The Concept of Sectorless ATM

The basic idea behind this new concept is rather simple. As already stated in the introduction, a controller is no longer in charge of managing the entire traffic within a given sector. Instead, he is now responsible for a certain number of aircraft throughout their entire flights within a given airspace (e.g. within the German airspace). This way of traffic control will not change the basic responsibilities given to the ATCos: They will remain responsible for a safe conduction of flights. Once an aircraft enters the airspace, either from a terminal area or via an airspace entry point, it will be assigned to the "next available" air traffic controller. This controller will then be responsible for the entire flight within the airspace until the exit point or until transitioning into a terminal area. The basic task of the controller will remain untouched: he has to ensure a conflict-free flight. This concept offers a lot of advantages over the traditional air traffic control. The most important ones are:

- The traffic load can be easily distributed in a very balanced way over the controllers on duty.
- Airspace capacity is no longer restricted by sector capacity. Instead, we can expect an increase in ATCo efficiency of up to 100%.
- It offers an easy way to implement contingency actions since controllers can

take over aircraft regardless of which center they are currently working for.

- Coordination actions between adjacent sectors are no longer necessary. The amount of voice communication between the aircraft and ATC are reduced. There are no handover and identification communications necessary.
- SESAR (or NextGEN) concept elements like business trajectories can be easily incorporated in this concept since the controller will have in mind the entire flight of his aircraft. He is only supposed to interfere in case of conflicts. The better the individual business trajectories are coordinated with each other the less difficult will be the job of the controller. A more detailed discussion can be found in [7].
- As the airspace structure (sectors, airways, etc.) are no longer necessary, a "Direct-To" based traffic organization can be envisaged.
- Controllers will immediately see what their control actions will mean to the aircraft. A closer relationship between controllers and aircrews will be established. The ATCo could be regarded as an additional (temporary) aircrew member taking care for conflict free routing. The development of a common situation awareness of aircrew and controller will as well contribute to an increase in safety.

Obviously, some technical changes with respect to voice communication between controller and pilot need to be introduced before such a concept can be realized. However, research has already been conducted that can enable decoupling air-to-ground communication cells from ATC sectorization [8].

Although the basic concept seems to be surprisingly simple and attractive at the same time there are some open questions remaining from an operation al point of view:

- What kind (if any) of ATCo-ATCo coordination will be necessary?
    - In case of conflict, the two affected aircraft most likely will be controlled by different ATCos. How can we assure that the necessary actions are defined unambiguously and
    - What kind of flight rules are required that necessary ATCo-ATCo coordination activities will not increase workload to an unacceptable level?
- How many aircraft can be handled by one controller simultaneously? What support tools and safety nets are required?
- How should the assignments of controllers to aircraft be realized?

From a safety standpoint, the first two points are most relevant. According to the main concepts elements, the main tasks of the controller will be in tactical control. Planning activities and most likely the most of the coordination activities will be carried out by the coordination and assignment center. The separation management is divided into two different tasks. Mid term conflict detection with a look ahead of 10 to 15 minutes (or up to 20 minutes) will be done automatically by the support functions. The main task of the controller is to perform separation management within 5-10 minutes look ahead time. This task sharing between controller and system will allow the controller to still generate a picture of the surrounding traffic of each of his aircraft under control. Doing this for all (up to six) aircraft for the full time horizon up to 20 minutes seems to be hardly possible.

3 Experimental Setup

3.1 Validation approach

In order to guarantee a successful validation process the European Operational Concept Validation Methodology (E-OCVM) [9] was followed. It defines different stages of concept maturity in the concept validation life cycle. This model is applied for the purpose of determining what type of validation exercises should be performed with the new concept / services under test. Due to its novelty the sectorless ATM can be allocated to a "V1" stage: in this stage general principles of the

concept are about to be validated, furthermore essential topics like operability, acceptability, and usability. Thus, this study can and will concentrate on operational feasibility only. In these early stages it has to be checked if operational requirements are fulfilled – otherwise the concept is not ready to be transferred to higher levels of maturity.

Operational feasibility trials make baseline test runs redundant as direct effects of a new concept in a pre-mature status cannot be expected. Therefore the users will be confronted with the new system in each test run. Results will focus mainly on the users' feedback. The trials will be characterized as confidence-building measure.

Following hypotheses have been identified to be tested in the subsequent study:

- Hypothesis 1: It is possible to manage six aircraft in one large sector by one ATCo.
- Hypothesis 2: Six radar tiles compiled on one single HMI work sufficiently according to the ATCos; geographical proximity is not necessary to handle the traffic.
- Hypothesis 3: Given the new sectorless concept, an additional conflict detection tool is essential.
- Hypothesis 4: In addition to the conflict detection, resolution advisory tools are desirable.

3.2 Methodology

The **simulations were run on DLR's traffic simulator** (TrafficSim) in the facilities of the DLR-Institute of Flight Guidance in Braunschweig, Germany. The simulator generates a 4-D trajectory for each aircraft and moves the aircraft along the trajectories in real time. As input the simulator uses flight plans, aircraft performance data, aircraft state vector at starting position and meteorological data. For each aircraft it can be operated in FMS mode and in open vector mode, which is used for controller commands. The simulator is designed for flexible use in human-in-the-loop real time experiments. It can also be used for fast time simulations; however, this was not used in our simulations. A more detailed description of the simulator can be found in [4], [5].

There were **two ATCo working positions** with a 4k*2k display each. The HMI design has to take into account that the allocated aircraft can be located in completely different parts of the airspace. In order to be able to provide all necessary information the display was divided into eight information areas, as can be seen in Fig. 1. There were six radar display regions (marked by R), one conflict display area (marked by C) and a situation area (marked by S).

For each aircraft under control there was a radar display with the same symbolism and layout as present radar displays. All displays provide basic flight information like position, altitude, speed, history and flight plan information. The aircraft under control is shown in magenta. The ATCo could zoom in and out of the radar display as needed. Conflict information was provided in case the aircraft was in conflict with another. The radar display also provided flight plan information and a MinSep tool to measure the minimum distance between two aircraft, either vector-based or trajectory-based. Fig. 2 shows the radar display in more detail.

R	**R**	**R**	**R**
R	**C**	**S**	**R**

Fig. 1 Abstract diagram of the display and its eight information areas

There was also an information area containing conflict information, the conflict display. The system identified lateral as wells as vertical conflicts. A conflict horizon of 0 minutes allows a look-ahead of separation conflicts specified with 5NM laterally and 1000ft vertically. The conflict display showed which aircraft were in conflict with each other and which rules applied to this conflict. It was indicated which of the aircraft had to introduce an avoidance maneuver and which of them was going to reach the conflict point first. Fig. 3 shows a picture of the

conflict display. In addition to this medium term conflict detection the controllers were also provided with safety-net measures, such as a short term conflict alert.

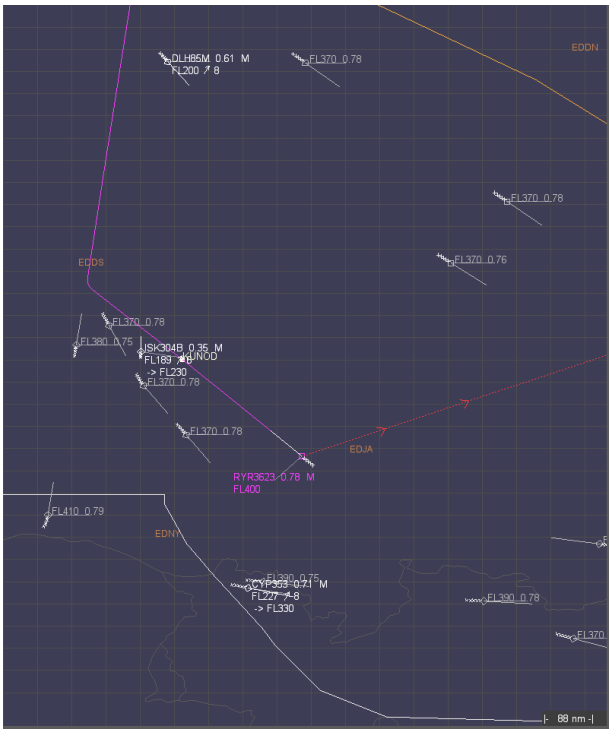


Fig. 2 The radar display



Fig. 3 The conflict display

Another information area showed a **newly developed situation display** containing information on all six situations at a glance. A discussion of this display will not be a part of this paper. The details and results concerning this situation display will be discussed in a separate paper.

The **ATCos communicated with the aircraft via datalink**. This was a purely practical choice to avoid having to introduce voice transmission and simulation pilots. It is expected that LRM2020 will extensively use datalink communication; however, it is not part of the concept to completely dispose of radio communication. It will still be necessary to have a direct voice communication with the pilots in addition to datalink communication. Reducing the communication to datalink only was just a simplification for this simulation.

The scenarios were based on real traffic over Germany, namely radar data from September 10th, 2008, a day with very high traffic load provided by DFS. The traffic was varied in order to be rich of conflicts. The conflicts were not constructed but resulted from the real traffic; however, they were condensed in order to have many conflicts within the 30-minute-span of a scenario. There were five scenarios, which were constructed differently (cf. table 1). The first scenario, which was used for training, contained mainly traffic moving from west to southeast and from southeast to west. The second scenario contained mainly traffic moving from south to north and from west to southeast. The third scenario contained all kinds of traffic movements. The fourth and fifth scenario contained for one ATCo traffic which was close to each other and moved from west to south, whereas for the other ATCo it contained all kinds of traffic movements. For simplification the scenarios did not contain weather information. Also restricted areas and temporarily reserved areas were not part of the simulations.

Tab. 1 Scenario Overview

Scenario	ATCo 1	ATCo 2
1 (training)	West → Southeast Southeast → West	
2	South → North West → Southeast	
3	All directions, no geographical proximity	
4	West → South Close geographical proximity	All directions, no geographical proximity
5	West → South Close geographical proximity	All directions, no geographical proximity

There were **only few independent variables** that were varied on purpose:

In most runs the ATCos were provided with conflict resolutions, i.e. they were given suggestions on how to deal with conflicts and avoid them. These suggestions could be chosen or the ATCos could develop their own strategies and implement them. In some of the runs these resolutions were withheld, such that the ATCos had to develop their own solutions all the time.

The situation display was not provided for all the runs. Since this display will not be discussed here but in a different paper this variation in the scenarios is not of interest here.

Only one run was dedicated to be conducted without the support of the conflict alert. This was only done once and mainly for system development interests. The outcome of this run will be discussed separately.

Appropriate **measurements** had to be identified to test the hypotheses. As ATCos' feedback had been identified as major source of information it was decided to measure users' response by

- extensive debriefing after each test run and by an
- operational feasibility questionnaire.

An additional data source is the

- observation of the test subjects during the test runs by two experienced ATCos acting as subject matter experts (SME). They have been involved in the development of the sectorless ATM concept for two years.

The questionnaire is tailor-made and follows the style of the interview technique which has already been used successfully in the European Project EMMA2 [10]. Participants of the study are confronted with several statements in a paper-pencil questionnaire after the final test run. They have to choose their amount of agreement to these statements with a six point Likert scale (ranging from "1 - strongly disagree" up to "6 - strongly agree"). In addition to that, they are invited to put down comments to each statement as well.

The **questionnaire** consists of the following sub-scales:

- General feedback on the sectorless ATM concept ("L"; 14 items); connected with Hypothesis 1.
- General feedback on the HMI and its presentation of all tools ("G"; 3 items); connected with Hypothesis 1.
- Radar Display ("R"; 23 items); connected with Hypothesis 2
- Conflict Detection and Resolution ("C"; 33 items) connected with Hypothesis 3&4.
- Newly developed Situation Display ("I"; 15 items)
- Preliminary tentative results on users' workload ("W"; 3 items)
- Preliminary tentative results on users' situation awareness ("S"; 3 items)

A thirty minute slot was reserved for a **debriefing** after each test run. Both ATCos, system developers, SMEs and experimenters took part in that session. The session allowed room for the ATCos assessment of workload, the traffic situation in general, the feasibility of the sectorless ATM concept under the conditions of the previous test run.

Training consisted of a two-hour long briefing and two test runs. It can not be compared to a realistic training on a new system that will be used operational by an ANSP in the near future.

Each ATCo worked in front of his own display. They performed four different test runs (cf. scenarios), each of them containing two ATCo working situations, i.e. overall each ATCo was confronted with eight different situations. There was also a training scenario which was not used in the simulations.

Since the experiments were human-in-the-loop experiments each run was expected to develop differently from the other, even if the starting situation was the same. Different controller reactions and reaction times were expected to cause different traffic developments later on in the runs.

The **trials were executed with eight ATCos** between the 29th November and the 9th December 2010. The experiment supervisor defined, organized and conducted the validation

tests; ensuring that all relevant results were recorded.

Participants of the study were eight air traffic controllers from the German air navigation service provider Deutsche Flugsicherung (DFS). Six of them were male, two were female. Their average age was 35.4 years (standard deviation: 8.4 years) and the mean reported experience as ATCo was 11.1 years (standard deviation: 7.7 years). While seven ATCos worked on the upper air space position, one ATCo worked both upper and lower air space.

4 Results and Discussion

The results of this study reflect the very first reaction of ATCos with long lasting experience on small and geographically well known sectors while interacting with one possible concept of tomorrow's ATM world.

Since the sample size is only eight different ATCos per item, the binomial test as a non-parametric statistic was used to prove the results for its statistical significance. The binomial distribution can be seen as a rather conventional and robust statistic, which is supporting the H0 hypothesis, meaning that results become harder significant in the expected direction (H1). The advantage is that results, which become significant, are more likely to be really meaningful.

By use of a binomial test for a single sample size, each item was proven for its statistical significance by following conditions:

- Answers from 1 (disagreement) through 6 (agreement)
- Expected mean value = 3.5
- Test ratio: .50
- N = 8
- $\alpha = 0.05$

A star (*) attached to the p-value means that the agreement with a statement has been significantly unambiguous because the p-value is equal or less than the critical error probability which is 0.05.

A plus (+) attached to the p-value means that the agreement with a statement has at least a significantly unambiguous trend because the p-value is equal or less than 0.10. More tests are needed to clarify if this trend is really unambiguous significant.

Mainly significant results are presented in this paper as any other results are unambiguous and therefore subject for further investigation.

4.1 Feasibility of six aircraft in one large sector

Given the ATCos' response to the statements of the **questionnaire**, there are already **significant trends** that the new concept is easy and intuitive to understand. This can be explained by their feedback that input devices were functionally simple and involving them in a minimum number of input actions. They think that it is conceivable to have even more than six aircraft under control. Nevertheless, rules and procedures are not yet complete according to the ATCos. Furthermore, there was a trend that working with the new concept revealed some problems in the simulations.

Tab. 2 Questionnaire Results: The concept

ID	Statement	Mean	Standard deviation	P-value
L01	The new concept is easy and intuitive to understand.	4.5	0.8	0.07+
L14	In my opinion rules and procedures are complete.	2.1	1.2	0.07+
L 11	In my opinion it is conceivable to have more than six aircraft under control. *(N=7)*	4.0	1.4	0.13
G02	The used terminology or symbology was easy to interpret.	4.9	1.3	0.01*
L06	Working with the new concept revealed some problems in the simulations.	4.6	1.3	0.07+
G03	Input devices were functionally simple - involving me in a minimum number of input actions.	4.9	0.8	0.07+
W05	Workload was higher when vertical conflicts had to be resolved.	4.6	1.2	0.07+

The positive feedback is supported by a **statistically significant result** that the used terminology and symbology was easy to interpret.

Yet first hints on experienced workload reveal that there was a trend for a higher amount of workload when vertical conflicts had to be resolved (cf. table 2).

In addition, ATCos were asked about their opinion up to which flight level they think that the sectorless ATM concept would be feasible. Their answers range from FL330 to FL380.

Further results from the ATCos' **comments** reveal that early implementation of conflict solutions results in more time to monitor the traffic. Without sector borders there were more possibilities to solve conflicts.

As indicated by one ATCo it was almost impossible to solve a conflict whose solution had been delayed too long. This result is supported by the **observation** made in one test run where such a situation resulted in the only operational error of the whole test runs.

Additional results from the observations reveal that a general scepticism towards the concept which was present in the beginning of the test runs, diminished as controllers became more comfortable with the new way of working.

From a technical point of view it can be summarized that the simulators were stable. During the 40 test runs only three other interruptions occurred; two due to the display and one due to the simulator.

4.2 Feasibility of one single HMI with six radar tiles

Given the ATCos' response to the **questionnaire**, there is significance that six radar tiles on one single HMI are feasible: the configuration of the radar display is accepted by them regarding size, contrast, zoom functionality, and update rate.

Furthermore there is a statistical trend that the configuration of the trajectory is accepted as well.

There are significant results that they were provided with a quick and efficient means to assign speeds and direct tos. A statistical trend exists at least for efficient means to assign headings as well.

ATCos tend to agree that assignments were manageable without any problems.

The feasibility was stressed by the agreement about the possibility to correct a mistaken action easily (cf. table 3).

Additional results from the ATCos' **comments** support these results. It was mentioned that the conflict display should be placed in the center of the screen; the radar displays should be as close to each other as possible. The new situation display can be at the top or the edge of the screen.

This request by one of the ATCos was immediately granted by DLR's system developers and the new layout was used and tested by this ATCo in the following runs.

Taking into account the ATCos' feedback it is important to be able to see all the situations at the same time. In their opinion zooming into a big overview display would not be a sufficient alternative to having individual displays.

Tab. 3 Questionnaire Results: Radar Displays

ID	Statement	Mean	Standard deviation	P-value
R01	The configuration of the radar display is fine with me, regarding • Size	5.3	0.5	0.01*
R03	• Trajectory	4.5	0.8	0.07^{+}
R05	• Contrast	5.1	0.8	0.01*
R06	• Zoom functionality	5.3	0.5	0.01*
R08	• Update rate (= it provides a continuous service)	5.5	0.8	0.01*
R11	I was provided with a quick and efficient means to assign • Headings	5.1	1.4	0.07^{+}
R12	• Speeds	5.3	0.7	0.01*
R13	• Direct tos	5.1	0.8	0.01*
R14	Assignments were manageable without any problems.	5.1	1.0	0.07^{+}
R18	The display allowed me to configure the display capabilities (e.g. range scale selection…).	5.4	0.7	0.01*
R21	It was possible to easily correct a mistaken action.	4.9	0.6	0.01*
R22	The response time of the HMI was adequate to allow making inputs without having to wait unduly for the system to process and validate the input.	5.3	1.0	0.07^{+}

The first impression that geographical proximity of the a/c under control is necessary

could not be sustained throughout the trials. In the end controllers admitted that geographical proximity is not as important as thought at first glance.



Fig. 4 ATCo in discussion with DLR staff

4.3 Need for a conflict detection tool

According to the ATCos' feedback to the **questionnaire**, they agree on the need for a conflict detection tool within the concept of a sectorless ATM.

Tab. 4 Questionnaire Results: Conflict Detection

ID	Statement	Mean	Standard deviation	P-value
C08	It was early enough to detect when a conflict partner gave way.	4.9	1.0	0.07^{+}
C13	I appreciate the function to highlight critical a/c automatically.	5.4	0.5	0.01*
C17	Colours are selected properly (e.g. "orange" for conflicts that have been noticed).	5.1	0.4	0.01*
C18	The change of colours (e.g. from red to orange) is ok for me.	5.3	0.5	0.01*
C25	I appreciated the information who was in charge of a conflict.	5.4	0.5	0.01*
C26	I appreciated that the conflict information was displayed continuously while the conflict is present.	5.1	0.8	0.01*
C23	I appreciated the conflict information in general. *(N=7)*	5.1	0.7	0.02*

There are significant results that they appreciated the conflict information in general and the function to highlight critical a/c automatically. They agree that colours and the change of colours are selected properly.

Furthermore they report that an identification of the ATCo who is in charge of a conflict is given as well. They approve that the conflict information was displayed continuously while the conflict is present.

A **significant trend** exists that it was early enough to detect when a conflict partner had to give way (cf. table 4).

As mentioned above only one run was dedicated to be conducted without the support of the conflict alert. This was done only once and mainly for system development interests. As expected this variation resulted in a tense atmosphere compared to the other runs. ATCos worked highly concentrated in order to maintain separation. However the ATCos mentioned after the trials that turning off the conflict detection did not shock them. They report that they also experience extreme situations nowadays.

4.4 Request for a conflict resolution tool

According to the ATCos' feedback to the **questionnaire**, there is no trend that a conflict resolution tool is necessary within the concept of a sectorless ATM.

Yet there is an agreement that they paid attention to the proposed solutions.

Nevertheless, there is a trend that ATCos think, that in some cases it would have been more sensible if the other conflict partner had had to give way. There are as well results from other ASAS related projects that do support a less stringent set of rules and revealed big advantages if resolution responsibility can be shifted in a deterministic way from one party to another, see [11] for more details. Questionnaire results give no hint that workload was higher in scenarios without suggested conflict resolutions. In addition to that ATCos hesitate to state that they missed suggested conflict resolutions when they were not available in selected simulator runs.

However, results of the **observations** show a tendency for a need of a conflict resolution tool. Since the ATCos were neither required to analyse the solutions provided by the tool nor

had to implement them, the provision of conflict resolutions is not disadvantageous.

5 Conclusions

The general outcome of the simulations was promising. Most of the ATCos were positive about the new concept and could imagine working in this manner in the future. Although all of them stressed the importance of further development and research only one of the simulation subjects announced a strong dislike for the concept itself and the way the ATCo would have to make use of the support tools.

The simulation revealed that it is possible for the ATCos to work with six aircraft under control in normal circumstances. It was obvious that introducing more complex situations might reduce the time the ATCo can spend with each of the aircraft, because his focus must stay longer on a certain situation. Also the size of the display and ergonomics do limit the number of aircraft under control. The ATCo must at all times be allowed to monitor the aircraft, i.e. their respective radar displays without movement of the head. This limits the number of radar displays since each of the displays must not be too small.

The simulations clearly showed the importance of a conflict alert tool. Since in a sectorless concept the ATCo works alone, without the help of a colleague (planner), most of the tasks of the planning ATCo must be taken over by system support. In a sectorless concept the ATCo relies on the conflict alert, which gives a 20-minute-ahead warning of conflicts. Since the ATCo has to monitor 6 aircraft at the same time it is impossible to perform such a long look-ahead, on the one hand because of the enormous complexity of such a long look-ahead and on the other hand because of the smaller radar display range around the aircraft under control.

Future research should focus on introducing realistic weather situations and restricted areas to the simulations. It must be analyzed if the sectorless concept works in airspace below FL 350 as well and the developed tools must be refined.

The expected advantages regarding capacity and flexibility keep the sectorless concept a research topic well worth investigating further.

References

[1] V. Duong, G. Gawinowski, J.-P. Nicolaon, and D. Smith, “Sector-less air traffic management”, 4th USA/Europe Air Traffic Management R&D Seminar 2001.

[2] T. Rivière, “Redesign of the European route network for sector-less”, Digital Avionics Systems Conference 2004.

[3] V. Duong, E. Hoffman, J.-P. Nicolaon, L. Floc’hic, K. Zeghal, and A. Bossu, “Extended flight rules (EFR) to apply to the resolution of encounters in autonomous airborne separation”, Eurocontrol Experimental Centre, 1997.

[4] B. Korn, C. Edinger, S. Tittel, D. Kügler, T. Pütz, O. Hassa, and B. Mohrhard, “Sectorless ATM – A concept to increase en-route efficiency”, Digital Avionics Systems Conference 2009.

[5] T. Pütz, O. Hassa, B. Mohrhard, B. Korn, Ch. Edinger, and D. Kügler, “Airspace Management 2020: Flying without sectors. Sectorless ATM — Increasing En-route efficiency by entire control of A/C’s trajectory in upper airspace”, Innovative Research Workshop & Exhibition 2009.

[6] A. Schmitt, C. Edinger, and B. Korn, „Balancing controller workload within a sectorless ATM concept“, Deutscher Luft- und Raumfahrt Kongress Bremen 2011.

[7] B. Birkmeier, B. Korn, and D. Kügler, “Sectorless ATM and advanced SESAR concepts: Complement not Contradiction”, Proceedings of the 29th Digital Avionics Systems Conference 2010.

[8] Eurocontrol, “Future Communications Infrastructure (FCI), http://www.eurocontrol.int/communications/gallery/content/public/documents/FCI_Step%202%20Report_V10.pdf

[9] E-OCVM, http://www.eurocontrol.int/valfor/gallery/content/public/docs/E-OCVM3%20Vol%20I%20WebRelease.pdf.

[10] EMMA2, Prague Test Report, http://www.dlr.de/emma2/maindoc/2-D631_PRG-TR_V1.0.pdf.

[11] Irvine, R., “Comparison of pair-wise priority-based resolution schemes through fast-time simulation”, Innovative Research Workshop & Exhibition 2009.

Algorithms and tools for dynamic airspace configuration based on genetic approach

K. Veresov
GosNIIAS, Russia

Keywords: *ATM, airspace configuration, genetic algorithms*

Abstract

The aim of the research presented in the paper is to obtain dynamic airspace configuration problem solutions applicable on practice and to overcome the shortcomings of existing approaches. For this purpose new method is proposed based on special genetic algorithm. The developed algorithm was tested on real data and showed good results in comparison with real airspace configurations.

1 Introduction

The dynamic airspace configuration problem consists in dynamic adaptation of the managed airspace structure to the air traffic changes.

Despite the fact that the air traffic intensity in the Russian Federation is incommensurably less than European level nowadays, the dynamic airspace configuration problem is relevant. First of all, the air traffic growth rate is rapidly increasing. Secondly, air traffic rate is high in a number of the flight information regions (FIR) because of considerable air traffic irregularity. And thirdly, the airspace structure modernization currently is in progress and includes the FIRs consolidation and airspace optimization.

It is necessary to seek for the optimal time and space partition of the airspace into FIRs and sectors. But the air traffic growth and FIRs consolidation combined with daily and seasonal traffic variations makes it impossible to solve the problem without using optimization algorithms and decision support tools.

There are two main approaches to the formulation of airspace partitioning in space and time problem. The first one considers the problem as a whole, the second – as dividing into two subproblems:

- sectorization – optimal dividing airspace into time-stable elementary volumes (elementary sectors);
- dynamic airspace configuration – creating optimal schedule of elementary sectors grouping into air traffic control (ATC) sectors during the day.

Both ways are used in existing works: for example in [1], [2] and [5] the problem is considered as a whole, and [3], [4] is focused on the dynamic airspace configuration problem.

The problem decomposition makes it impossible to obtain optimal solution in terms of a whole problem. On the other side it can be considered as a price for meeting ergonomic constraint on permissible sectors boundaries changes [6]. The sectorization problem was successfully solved by GosNIIAS specialists [7], this research deals with dynamic airspace configuration problem.

2 Background and related works

Investigations in this field are conducted about 20 years in Europe and USA. During this time, various algorithms for the dynamic airspace configuration have been proposed. Among them are the clustering algorithm [1] [2], the genetic method [3], branch-and-bound principle [4]. There are also combinations of different approaches, for example the algorithm based on Voronoi diagram and genetic method [5].

The dynamic airspace configuration tools are included in several ATM simulation tools designed for airspace usage planning. NEVAC (Network Estimation Visualization of ACC Capacity) developed by EUROCONTROL can be given as an example.

The existing approaches have a number of limitations or weaknesses: simplified airspace structure description; insufficient taking account of practical rules of real air traffic controllers work; using of simple optimization criteria with no account taken of a number of important factors. At present, new approaches to the problem aimed to overcome listed shortcomings are developing [5]. The proposed method, which is a subject of the paper, has similar nature.

3 Problem formulation

This work was aimed at practical solving of the problem (i.e. obtaining applicable in practice solutions). The problem formulation has many features in this context.

The problem is solving within the limits of single ATC region, which is defined by its boundaries, airspace partitioning into elementary sectors and air route structure. Airspace usage practice in Russian Federation allows airspace altitude partitioning into two layers. Airspace partitioning into elementary sectors can be result of computer-aided sectorization problem solving or solving by expert without using special tools.

The problem can be solved on the basis of one of the two assumptions:

- traffic intensity in the region is variable during the day, but the variations schedule is stable from day to day
- the intensity variations schedule considerably changes from day to day.

In the first case the solution corresponds to cyclic schedule of elementary sector grouping that pass through one day to another, in the second the problem is solving for each day separately.

The flight plans for chosen ATC region and time interval are used as an input data.

The time discretization unit is equal to one hour.

The optimization criterion (Eq. (1)) provides for minimization of the sector-hours per day total number and controller workload balancing for the ATC sectors working at the same time.

$$J = \frac{1}{24}\sum_{i=1}^{24} N_{C_i} + \rho \frac{1}{24M}\sum_{i=1}^{24}\sum_{j=1}^{M} Q_{ij} \to \min \qquad (1)$$

where M – number of daily flight plans in set

N_{Ci} – number of ATC sectors working in i-th hour

ρ – weight coefficient;

Q_{ij} – controllers workload equability measure for ATC sectors working at the same time

$$Q_{ij} = \left[\frac{1}{N_{C_i}}\sum_{k=1}^{N_{C_i}} \left(z_{ijk} - Z_{ij}\right)^2\right]^{\frac{1}{2}} \qquad (2)$$

z_{ijk} – nondimensionalized controller workload for k-th sector in i-th hour of j-th flight plan

Z_{ij} – mean z_{ijk} with respect to sectors in i-th hour of j-th flight plan.

The constraints are based on practical rules of controllers work:

- Controller workload must not exceed maximum permissible value every hour for each sector.
- Airspace configuration should change not more frequently than minimum fixed value.
- To meet previous constraint the sector grouping schedule should be closed (cyclic) in case of stable traffic intensity variations schedule. Otherwise it should take into account the sector grouping schedule used in the previous day.

The formulated optimization problem has the following features:

- discrete nature of the optimization parameters;
- high dimensionality of the optimization parameters;
- multiextremal criterion function;
- search space nonconvexity, disconnected feasible regions
- discontinuity of the functional of factor dependence on the optimized variables
- expediency of obtaining the spectrum (set) of solutions.

These specific features of the mathematical synthesis problem, on the one hand, make impossible the application of nonlinear programming, and on the other hand, do not allow constructing efficient numerical procedures using complete or limited enumeration. Realistic are algorithms based on ordered and directed random search.

4 Proposed solution method

Genetic optimization algorithms based on evolutionary approaches for finding an optimal solution possess the important advantages which provide successful solution in the conditions described above. These algorithms are related to zero order methods (that do not require derivative calculation) and consequently, do not impose conditions of continuity and differentiability of the criterion function and constraint functionals. They are robust with respect to the form of the optimized function and provide finding the global extremum. In these algorithms, the optimal solution is sought by direct "manipulation" of the sets of several admissible solutions (to which inadmissible but attractive from some point solutions can be added) forming the population. The search is concentrated on the most promising of them with account of the search prehistory. However, complexity of the formulated problem didn't allow using simple genetic algorithm with binary coding for its solution. Therefore the special genetic algorithm has been developed.



Fig. 1 Coding

Fig. 1 shows coding of the schedule of elementary sectors grouping/degrouping as string of integers (chromosome in terms of genetic algorithms). As one hour is chosen as the discretization unit, the string contains of 24 string cells (genes). Each of them contains number (A, B, C) which corresponds to the unique scheme of elementary sectors (fig. 2) used in this hour. Thus the chromosome is divided into time intervals, and different sector configuration is used in each of them.

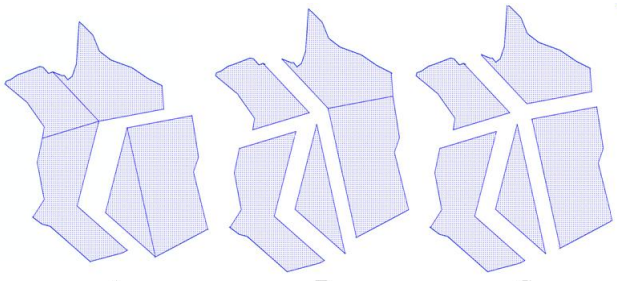


Fig. 2 Grouping/degrouping of elementary sectors during the day

Solutions (individuals) for the first initial set (zero population) are randomly generated and ranged by criterion in ascending order. Then the search of the best solution is

accomplished by iterative improvement of the solutions set (population).

As we can see in fig. 3, the full population is formed on the each iteration. It includes all persons from previous population and new persons, obtained by applying crossover and mutation operators. The operators are applied in parallel to different randomly selected «parent» individuals. Their selection is based on a probability line so the best individuals have a better chance to be chosen. Then the full population is ranged by criterion in ascending order and the best individuals is selected into the next population. Thereby the algorithm scheme is close to the standard one. The main features of the developed algorithm lie in realization of specific genetic procedures. It is caused by specificity of the solution coding.

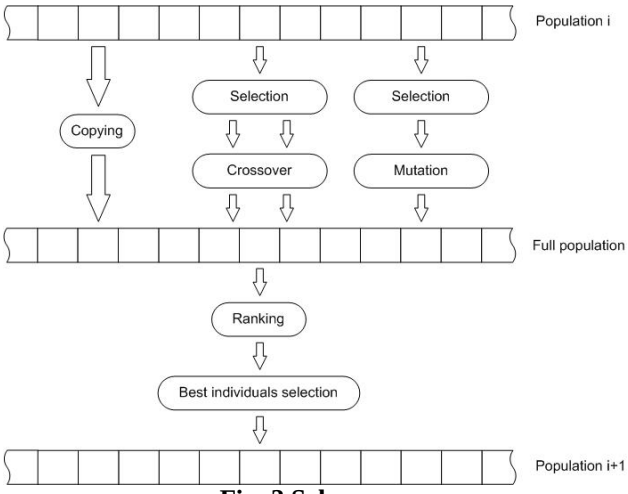


Fig. 3 Scheme

It should be noted that the developed algorithm includes two special methods to prevent premature convergence to local minimum and to increase the diversity of solutions.

1) Too similar solutions are eliminated from the new population, for which purpose the concept of affinity based on configurations geometrical likeness is used. The worst of two related individuals is discarded.

2) Individuals are divided into two categories: solutions that meet the constraint on controller workload and solutions that failed to meet it. Part of the population is allocated for individuals from the second category.

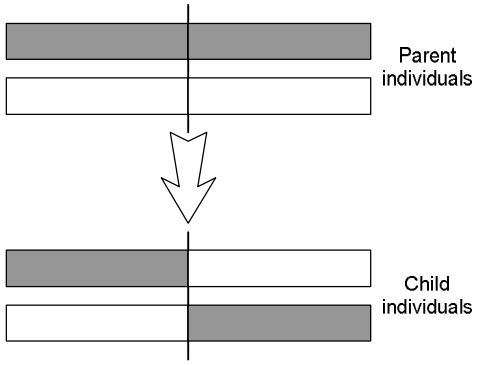


Fig. 4 Crossover

The crossover operator is constructed under the one-point scheme. After selecting two parents the crossover point is randomly chosen. Children chromosomes are formed of two parts of parents chromosomes (fig. 4). By default, the crossover point becomes boundary between two intervals. However child individual may not meet the constraint of configuration change frequency (i.e. intervals length) as a result of such division. Therefore the «cure» operator is applied to each child individual that has not passed verification. It is using two methods: merging short intervals and slight shifting of intervals boundaries. If neither one of them didn't give positive results, the individual is considered unsuccessful and is not included into the full population.

The mutation operator includes three different mutation types, which based on alteration of interval structure of the chromosome:

1) focused change of sectors configuration on the randomly selected interval;
2) shift of the boundary between two intervals;
3) merging of two intervals/division of the interval into two.

Thus the first mutation type designed to optimize the configuration on the interval. Focused configuration change means merging of underloaded sectors and division of overloaded compound sectors.

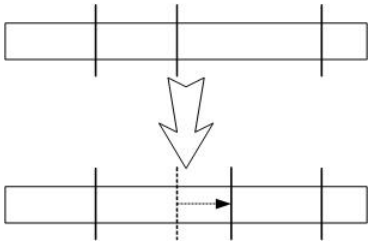


Рис. 5 Shift mutation

The second and third mutation types intended for optimization of intervals bound (for example fig. 5 shows the scheme of shift mutation type).

All prerequisite calculations (such as workload computation) are carried out into preparatory stage to increase algorithm performance itself.

5 Tools

The described algorithm is realized in dynamic airspace configuration software tool. This tool is included into ATM simulation tools «KIM OrVD», which is designed for the ATM researches and planning by GosNIIAS [8].

The dynamic airspace configuration software tool allows to estimate airspace structure both in terms of numerical characteristics (controller workload, traffic capacity), and from the geometrical point of view. The 3D visualization tool is developed for convenience of sector configuration visual analysis (fig. 6).

6 Results

Conducting of dynamic airspace configuration by an expert is possible without using special tools. Schedules received by this way has some improvements and meet the constraints, however such solutions are not optimal in most cases. Also it takes too much time and requires good knowledge of the region airspace to obtain schedule even for small ATC region.

The effectiveness of computer-aided dynamic airspace configuration based on the developed algorithm in comparison with the results received by an expert can be shown on the example of St. Petersburg FIR consisting of five elementary sectors (Fig. 6). At the present moment schedule of this FIR is divided into two intervals: from 18:00 to 7:00 two elementary sectors are merged, the rest of the time elementary sectors are separated. Thereby this schedule saves 12 sector-hours (total cost is 108 sector-hours instead of initial 120). Airspace structure and sectors grouping schedule change 1-2 times a year in St. Petersburg FIR, such scheme can't adequately react to traffic intensity variations.

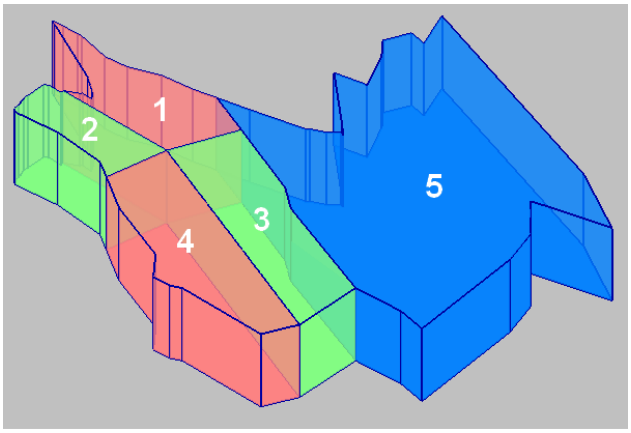


Fig. 6 St. Petersburg FIR

Let's take several flight plans for April 2011 as input data and conduct dynamic airspace reconfiguration for the described FIR. For the results shown below (table 1) the configuration change frequency parameter was set to 8 hours.

Night interval	Morning interval	Day interval
17:00 - 24:00, 0:00 - 1:00	1:00 - 9:00	9:00 - 17:00
1+3+5	1+2+3+4+5	1+3
2+4		2
		4
		5

Table 1

As we can see from fig. 7, significant reduction of the ATC sectors number is achieved: 2 sectors are used at night interval, 1 at morning interval and 4 at day interval, in total 56 sector-hours are spent. Fig. 8 shows the difference between maximum and minimum workload of controllers working at

the same time. The workload became more balanced as a result of dynamic airspace configuration. At the day interval difference is slightly reduced or remained the same. At the night interval significant reduction is achieved, values fell from 0.2 – 0.4 below 0.1 level. At the morning interval all elementary sectors are merged in one ATC sector so the difference on the chart fell to zero.

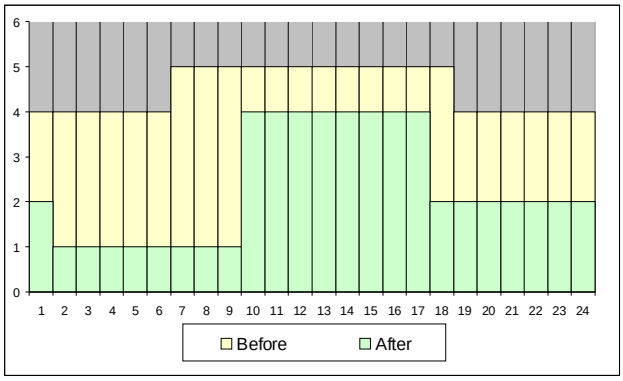


Рис. 7 Sector-hours minimization

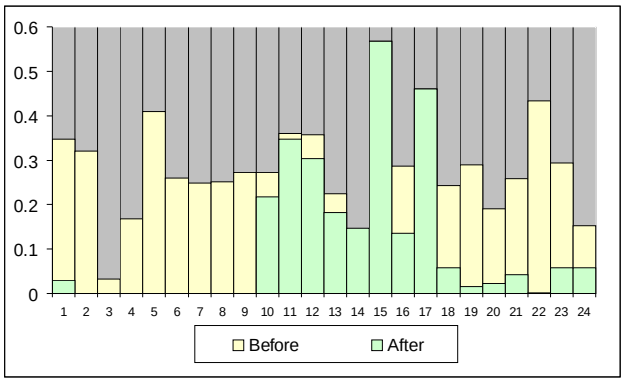


Fig. 8 Controllers workload balancing

The efficiency analysis of the developed algorithm was carried out not only for small ATC regions but for large ATC centers as well (Moscow FIR, 19 sectors). After splitting twenty-four hours into three 8-hour intervals a number of spent sector-hours is reduced by 15% on average in comparison to the structure used in practice. The algorithm showed good performance. The optimal solution can be reached in 3-5 iterations (≈3 seconds) for small ATC regions. Finding near optimal solution for large regions takes about 15-20 minutes.

7 Further work

As noted above, FIRs trend toward consolidation at the development of Russian Federation ATM system. As a consequence the dynamic airspace problem becomes more complex. Therefore the further work main goal is to optimize algorithm for more effective solutions search in complex cases. The parallel genetic algorithm is one of the ways to achieve it. In such algorithms optimization process is conducted in several populations at the same time and solutions interchange between populations (migration) is used.

8 Conclusion

The new approach to the dynamic airspace configuration problem is proposed. The software tool based on it makes it possible to effectively manage throughput of the flight information region in dynamically changing traffic conditions. This work is executed with financial support by RFBR (project 11-08-00841-a).

References

[1] A. Yousefi, G. Donohue. “Optimum Airspace Partitioning for Center/Sector Boundary Design”, *1st International Conference on Research in Air Transportation* - ICRAT 2004, November 22-24 2004, Zilina, Slovakia.

[2] S. Martinez, G. Chatterji, D. Sun, A. Bayen. “A Weighted-Graph Approach for Dynamic Airspace Configuration”, *AIAA Guidance, Navigation and Control Conference and Exhibit*, Hilton Head, South Carolina, AIAA 2007-6448, 20 - 23 August 2007.

[3] D. Delahaye, J.M.Alliot, M. Schoenauer, J.L.Farges. “Genetic algorithms for automatic regroupement of air traffic control sectors”, September 1994.

[4] D. Gianazza. “Forecasting workload and airspace configuration with neural networks and tree search methods”, *Artificial Intelligence* No. 174, 2010, pp. 530–549.

[5] Min Xue. “Airspace Sector Redesign Based on Voronoi Diagrams”, *Journal of aerospace computing, information, and communication*, Vol. 6, December 2009, pp. 624–634.

[6] O. V. Degtyarev, V. N. Minaenko, and M. O.

Orekhov “Solution of Sectorization Problems for an Air Traffic Control Area. I. Basic Principles and Questions of Airspace Sectorization and its Formalization as an Optimization Problem”, *Journal of Computer and Systems Sciences International*, Vol. 48, No. 3, 2009, pp. 384–400.

[7] O. V. Degtyarev, V. N. Minaenko, and M. O. Orekhov “Solution of Sectorization Problems for an Air Traffic Management Area. II. Development of Sectorization Algorithms”, *Journal of Computer and Systems Sciences International*, Vol. 49, No. 4, 2010, pp. 624–642.

[8] L. Vishnjakova, O. Degtyarev, A. Kan, E. Filenkova “Algorithms, procedures and tools for FUA by ATFCM based on modern information technologies”, *Concepts and technologies of air traffic management*, 19-21 august, 2009.

Universal algorithm and program tools of 4D flight plan synthesis under conditions of dynamic special use airspace

E. Filenkova
GosNIIAS, Russia

Keywords: *ATM, flight plan synthesis, graph theory, free flight*

Abstract

Algorithm and software tools for optimal route synthesis problem are considered. The 4D route is built, different criteria are examined: total flight time, route length, fuel consumption, flight cost. The proposed solution provides the route planning by various criteria, under different restrictions, for different rules, with a various detail level aircraft performance characteristics. Route building is implemented in two stages: an optimal 2D route construction on a graph, then - the vertical profile and time synthesis. Algorithm accounts for dynamically changing spatial and temporal restrictions of different nature

1 Introduction

The 4D flight plan synthesis problem is one of the basic, most important and difficult tasks for air traffic management (ATM) and air traffic control (ATC).

This problem is solved at different stages of flow management by various participants of air traffic:

- airline's planning department for flight's requests;
- the flight crew during flight-plan creation and on-board route calculation and correction (if flight conditions changed);
- planning services of ATM at all stages of traffic flows planning (for flights request adaptation and for traffic flow management).

In this paper the problem of 4D flight plan creation is considered. The term "4-dimensional flight plan" (that's the same 4D route) means number of plan data of 2-dimensional route (2D route), altitudes and time profile of the flight. Now this problem becomes more actual because of the development control ideology named as "business trajectory" management.

The flight plan synthesis problem is to select the best route from the set of possible routes for the selected criterion, subject to maintaining all rules and restrictions on the airspace use.

The full 4D route synthesis is a non-trivial task. It is necessary to take into account numerous spatial-temporal airspace restrictions that can have different mathematical description.

There are many applications related to the construction of routes that have their own characteristics. But the procedure should be one for each problem.

In this paper we propose a two-step and iterative scheme for 4D route synthesis under conditions of dynamic airspace use constraints.

The first step is the 2D route synthesis and the second step is building time, altitude and speed flight profiles for this route. 2D route synthesis is executed as follows. Optimal for given criterion route is searched on the basis of current criterion components values for each route segment.

For the formed route effective (with minimum fuel consumption) flight profiles are determined in all phases of flight (climb, cruise, descent).

So, next problem solving is provided: to find optimal route for given criterion from the set of routes executed on effective regime of flight and satisfying all constraints

Iterative search for the optimal solution has the following reasons. First, it is often impossible to determine a priori, what the spatial-temporal airspace restrictions will be active for the synthesized route. Secondly, it is impossible to determine a priori in what kind of conditions (altitude, time and weight) the aircraft will fly the specific segments. Part of an integral criterion corresponding to this segment depends on these characteristics (the exception is a special case of optimizing the length of the route).

An additional advantage of this approach is that algorithms for solving the second step can be unified for both en-route, and for "free" flights. In this paper we consider solving the problem of a route synthesis, within the concept of airways flights.

2 Statement of route synthesis problem

Algorithm and software tool must be universal and is intended for various users. So the technical problem statement can be described as follows.

1. It is necessary to build optimal 4-D route for selected criterion. This criterion can be considered as one of the following indicators:
   - total flight time;
   - total route length;
   - total fuel consumption;
   - flight cost - the total cost of its implementation.

2. Generated route must avoid violations of airspace use constraints. It can be:
   - overloaded airspace elements;
   - restricted usage of sectors;
   - regions, areas and points;
   - special use airspace;
   - specific waypoints to fly using it.

3. The final route must meet additional requirements, for example, the similarity of the given route.

4. Following specific requirements describes universality of the proposed route creation procedure:
   - route can be created: between two airports, between the airport and the point, between the point and the airport, between two points;
   - route can be created for any criterion, selected by user.

5. It must be possible to create route with the user specific rules to use en-route structure (international and domestic routes, the civil and military flights).

6. Program tools must provide a good performance (to have a low computational complexity) for operational use in difficult conditions (for example, air traffic flows rerouting, or for the iterative search for air traffic flow management solutions).

Formal statement of the challenge is as follows:
   - The route starting and end points (waypoints or airports) are defined.
   - The type of route is defined.
   - The aeronautical structure is defined, it is included:
     - segments and waypoints for defined airspace;
     - FIR and air traffic control sectors.
   - Airports and SID/STAR for it.

   It is identified additional conditions:
   - air traffic control sectors list, possible times of special use for it;
   - list of non-active (or overloaded) FIR;
   - segments list from special use airspace, with limitation times and heights.

Types of aircraft and aircraft performance characteristics are defined.

The date and time of departure are defined.

The weather data (as equivalent winds for all segments) is defined.

Criterion for route creation is defined.

It is necessary to generate route between the two selected points. It must meet all conditions, rules and airspace use restrictions for minimum value or criterion.

Route generation is the optimization task with constraints. The solution is found for airway's network. Criterion can be considered as an additive - flight time is sum of the flight times for all route segments, total route length is sum of the lengths for these segments, the total cost - the sum of the cost for all flight segments. The task ideology (on the en-route structure) is similar to the tasks in graph theory. Each route segment can be characterized by a certain price, and search for the optimal solution is a set of these segments, corresponding to a number of conditions. The main practical problem of using these algorithms is:

- "cost of use" for each route segment is a function not only of its length but also of flight conditions (altitude and speed);
- it is often impossible to know a priori, a possibility to use specific segment and the cost of it due to spatial and temporal constraints.

3 Algorithm of 4D route synthesis and its features

The proposed algorithm is implemented as an iterative process of route creating (Figure 1 shows the integrated scheme for search of the optimal decision).

The airways structure data are prepared a priori. It depends on the given route options and rules. They are stored in accordance with a graph theory in the "conformity array".

Preliminarily the conformity array correction and tuning are executed: unconditionally inadmissible segments for use are excluded (accounting static airspace constraints). After this route initial approximation is constructed.

Then the full 4D route is synthesized; specific algorithm depends on various detail level of aircraft performance characteristics and the flight rules taken into consideration.

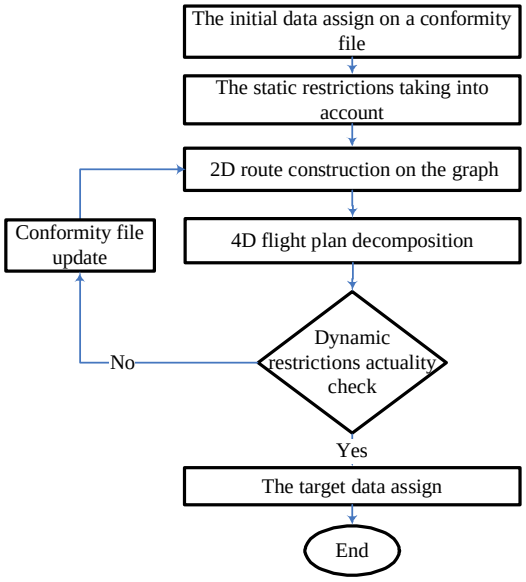


Fig. 1 Scheme for search of the optimal decision

The conformity array correction is made again on the basis of the received results.

First, this full 4D flight plan is checked on satisfaction of constraints, non-stationary or existing on the part of altitude range. If the constraints are violated on the route segments, unacceptably great values are appropriated to conditional lengths in the conformity array. Secondly, for all other allowable route segments, the conditional lengths are calculated as actual components of the criterion which are resulting from the conditions of overflying these segments.

After the conformity array correction the new 2D route is synthesized. The cycle is repeated before the decision coincides exactly with the result of previous step, that is the fully admissible decision is obtained.

In this paper we present a number of original algorithmic decisions, what are caused by:

- using methods of graph theory in a 4D problem of conditional extremum;
- the need to optimize the criteria, which not allows building the conformity matrix a priori;
- the need for high processing speed.

The basic decision is shown below.

3.1 Particular features of the proposed algorithm and their realization

1) The 2D flight plan synthesis problem is solved with algorithm of the shortest way search on the graph, which is based on Dijkstra's method.

In general, this algorithm is based on assigning temporary marks to the graph nodes, and the *j*-th graph vertex mark is a high bound of the length path from the starting vertex *s* to this *j*-th vertex. One of the marks becomes permanent at each iteration step. This indicates that the mark is no high bound and it gives the exact length of the shortest path from graph vertex *s* to vertex that corresponds to this mark [3].

In this task, the graph vertexes are the airways points, and the graph arcs are the airways segments. The arc cost (the cost of transition between two neighboring points that are connected with a segment) is the segment length or the conditional length. Path length (cost) μ between two given points is the length of the route (or conditional length - for example, flight time). The algorithm application is based on preliminary formation of conformity matrix *C*. Values of the segment length that connects *i*-th and *j*-th points are assigned to *ij*-th and *ji*-th matrix elements. If segment between two points not exists, then the *ij*-th and *ji*-th matrix element is appropriated unacceptably big size. If the motion is allowed only from the *i*-th point to the *j*-th, then *ji*-th matrix element is appropriated unacceptably big size.

The large dimension of the conformity matrix that is determined by the number of vertices is one of the practical algorithm application bottlenecks. This dimension is the total number of airports and waypoints, which may participate in the route construction. The number of waypoints is many thousands for Russian Federation, the number of airports - a several thousand. As a result, the total number of matrix elements is very large. When the long transit route is created, this number is increased many times. At the same time a number of airways segments in airspace of Russian Federation is much less, it is about 10000. This is the total number of graph arcs. The number of significant matrix elements is small - hence, the matrix is sparseness.

Here the matrix dimension problem is solved as follows. First, arcs costs data are formed as an array instead of a matrix, the so called “conformity array”. Secondly, only two airports - the departure and destination- SID/STAR data are includes in the array instead of all airports data.

The conformity array *MS* dimension is equal to *4*Nm*. The number of columns *Nm* is determined by the total number of possible outcomes of all the graph vertices and is calculated in the process of the array creation. *MS* is ordered: graph vertex №1 is departure airport, graph vertex №2 - destination airport, graph vertices from №3 to *Nm* - all airspace points (the total vertices number in this graph is *Nm*, the number of points – (*Nm* -2)). If the first route point is the airways point, not airport, in the first column of *MS* zero length is recorded, similarly - for the last point. The route is always created from the graph vertex №1 to the graph vertex №2. To change the end route points it is sufficient to change only the first 2*Ns* array columns or their part. It is believed that the route has one intermediate point at least. The *Ns* value is determined by the maximum number of SID/STAR routes for the airport.

2) The route creating task is relatively simple for minimal route length criterion. But the task becomes much more complicated at use of the cost criterion and its special cases. In these cases it is offered to use the conditional cost of the segment overflying as the arc length.

Let’s examine this method by the example of the flight time minimization task. Ideally, the conditional cost can be represented as:

$$Cs = l / [Vkr] \qquad (1)$$

where V_{kr} - cruising speed for given aircraft type, l – segment length.

But this is correctly under the certain conditions:

- Cruising airspeed on the segment is a constant;
- The value of this speed is known in advance;

- The wind can be ignored.

However, this simple case is too rough and does not consider essential things:

- The speed at the climb and descent route segments is variable and can be different from the cruising speed. At the same time it is impossible to determine in advance (until route creating) what phase of the flight is – descent, climb or cruise;
- The effect of a wind is not considered. But the minimum time criterion is used, when the wind effects. The value of equivalent wind depends on the segment of route, altitude and time of its overflying. Because *MS* is becoming also dependent on the aircraft type, it has to be generated again when the new route is created.

It is offered the following solution.

Estimation of a minimum time of flight over any segment of the length l by any aircraft is:

$$t_u^{min} = l/(V_{kr}^{max} + V_{eq}^{max}) \qquad (2)$$

V_{kr}^{max} – the maximum cruising speed for a given aircraft type,

V_{eq}^{max} - the maximum value of the associated equivalent wind.

The idea of the suggested solution is: at the initial generation of the conformity array the segments lengths cost is taken as t_u^{min}. *MS* differs from the array corresponded to the minimum length criterion in a constant:

$$k = 1/(V_{kr}^{max} + V_{eq}^{max}) \qquad (3)$$

Therefore, a minimum-length route will be construct at the end of the first step. The moments of overflight of the segments, which will be calculated under actual conditions, can be in general case higher than the initial ones in *MS*. Array is corrected at each step of the iterative procedure: corresponding elements of *MS* are changed to the calculated moments of overflying. Then route synthesis is repeated. Therefore at the second and following steps the algorithm is trying to find the best route that can be different from the shortest. Iterative process of route creating with the correcting of the *MS* will continue until the increase its length becomes ineffective and at the next step modifying of array elements does not occur.

In order to avoid an initial MS adaptation to each aircraft type, it is expedient to consider "conditional length" of segment as the elements of *MS*. This "conditional length" of segment differs from the overflying time in a coefficient $(V_{kr}^{max} + V_{eq}^{max})$. Then, independently of an aircraft type, the single array of segments lengths will be used as an initial approximation. The synthesis of the 2D route allows to calculate an actual flight profile and moments of overflying t_u of the segments and to correct the corresponding *MS* elements. The segment "cost" is taken as:

$$C_s = (V_{kr}^{max} + V_{eq}^{max})t_u \qquad (4)$$

Thus, by the time when the solution will be found, the elements of the *MS*, related to segments that were a part of the route at least at one step, characterize actual moment of overflying, and the rest elements- the minimum possible time.

Similarly, the *MS* can be formed and the fuel consumption criterion optimization can be executed. In this case, the lower limit of fuel consumption on segment is:

$$W_y = q^{min}[V^{min}/(V^{min} + V_{eq}^{max})]*l \qquad (5)$$

where q^{min} - minimum fuel consumption per kilometer of a path,

V^{min} – the minimum airspeed.

The arc cost is taken as a fuel consumption q_u on segment, with the coefficient:

$$k = (V^{min} + V_{eq}^{max})/(q^{min}V^{min}) \qquad (6)$$

The initial value of any *MS* element is the arc cost value, which corresponds to the lower limit of fuel consumption at the segment (that is the actual segment length).

Solution of the optimization problem in the case of the most general integral criterion can be organized in accordance with the same scheme.

In this case the cost of the segment overflying can be represented by the next value:

$$CT_u^{min} = \{q^{min}[V^{min}/(V^{min} + V_{eq}^{max})] + CI/(V_{kr}^{max} + V_{eq}^{max})\}*l \quad (7)$$

where *CI* – index cost.

As the corrected conformity array *MS* element value the cost of segment overflying is taken with the coefficient:

$$k = \{q^{min}[V^{min}/(V^{min} + V_{eq}^{max})] + CI/(V_{kr}^{max} + V_{eq}^{max})\}^{-1} \quad (8)$$

and the initial *MS* value is taken as the matrix of segments lengths.

So the algorithm to find the optimal solution for all versions of the criterion becomes single and universal. It is an iterative search for the optimal path on a directed graph with subsequent correction of conformity array *MS* elements. This correction is related with an account of "true cost' of flying over route segments and of space-time constraints.

3) The algorithm for calculating the third (altitude) and fourth (time) 4D route components is determined by aircraft performance characteristics that were accepted. So, in accordance with the data structure and data volume of BADA the following flight characteristics are selected in each current time moment:

- at the climb phase: the vertical speed and fuel consumption as function of mass and flight altitude; airspeed as function of flight altitude:
- at the cruise phase: airspeed as function of flight altitude, fuel consumption - as function of mass and flight altitude. Flight level is chosen closest to the maximum possible altitude for the cruise (subject to climbing capacity restrictions). The choice is made with the account of general rules of airspace use and specific conditions of segments use;
- at the descent phase: the vertical speed and hourly fuel consumption as function of mass and flight height; airspeed as function of flight altitude.

It should be noted that aircraft mass must be given in initial data, this causes certain difficulties during planning air traffic flows by ATM planning services.

4) One of the particular algorithm features is the scheme of taking into account the real airspace restrictions. Various restrictions can be reduced to segments use restrictions, consequently, their influence can be taken into account by the corresponding change of the cost of segment overflying and appropriate values of conformity matrix elements. If the segment can not be used for some reason in the synthesizing route then unacceptably big size is assigned to the matrix elements corresponding this segment.

Route synthesis on the current algorithm iteration allows to determine with a high accuracy the time and the altitude of flight over concrete segment, therefore it is possible to check time-space constraints. Theoretically there may be cases that the segment which was forbidden at previous iterations could be used in the synthesized route (although at the limit). Nevertheless, some loss of strict optimality is justified because it allows to fulfill the requirements with some margin.

4 Experience of using

The developed algorithms and the program tools are realized in the ATM simulation tools «KIM OrVD», which is developed in the State Scientific Centre of Russian Federation «GosNIIAS». These tools are maintained in "State ATM Corporation" of the Russian Federation for a number of years. 4D flight plan synthesis algorithms are used for decision of following problems:

- air space usage planning;
- the air traffic flows management (flight plan correction in reply to flight conditions changes; air traffic flows rerouting at flight conditions change (special use airspace, an airspace elements overload, weather restrictions [2]));
- proposals development and measurement estimation on airspace structure perfection.

In particular, the route synthesis procedure was successfully used at route structure modernization for a number of aggregated FIRs (St.-Petersburg, Rostov, Moscow and others FIRs).

The example demonstrates the algorithm route synthesis result for the flight from UUEE (Moscow) to UHHH (Khabarovsk), aircraft type is B744, departure date and time - 03/17/11 08:00 (Figure 2). Three restrictions were determined on a given direction: 1 - the sector USHH/SECTOR 2 (Khanty-Mansiysk) is forbidden for flights because of sector capacity overload during two hours interval; 2 - the segment GINUR-RULAS is closed for flights, 3 – airspace area is closed for flights during interval from 08:00 to 12:00 3/17/11. Optimal route (without restrictions) is shown red for criterion of flight time. This route violates all 3 restrictions. The new route avoids the restrictions, it is shown by lilac color. It satisfies to restrictions 1 and 2. It seems that the trajectory violates area 3, but it isn't so: figure 3 shows the part of new route profile. Aircraft cross bound constraints (segment LEMGU - LEBNA) at 13:06, when the restriction has finished the action.

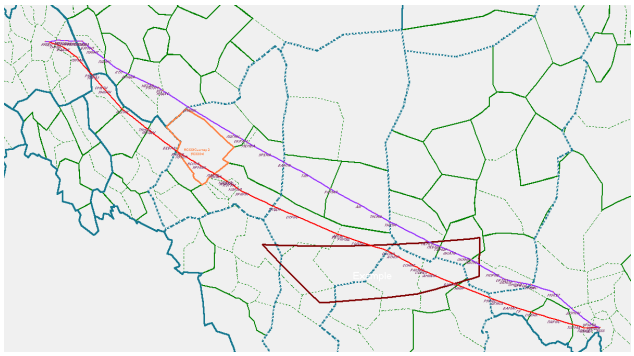


Fig. 2 The route synthesis result

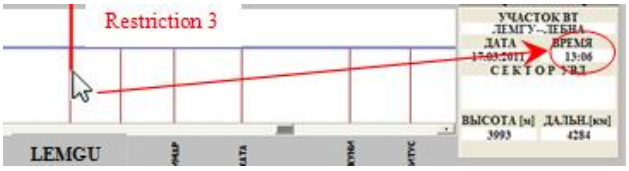


Fig. 3 New route profile

5 Other applications

Free flight areas are used more widely from year to year. The each user has greater freedom to create a preferred route in this area. There are the tasks: full route synthesis in this area, and the route synthesis in the free flight area as a part of full route with other airways parts. Such area describe as pseudo-en-route structure. This approach is used by a number of experts and implemented in some software tools. For example, flight planning tool across the Pacific Ocean (it uses a pseudo-fan-route structure), route planning for general aviation. In this paper, the structure is complicated and modernized to ensure minimum deviation of new routes from optimal trajectory (route length deflection does not exceed 0.1-0.2%). Figure 4 shows pseudo-en-route structure. Connection of this structure with airways structure outside free-flight area is provided by pseudo-airways segments that connect airways points near free-flight area boundary with the nearest nodes of pseudo-airways net. This structure allows to overfly the area in any direction. Linear network size is chosen for practical reasons and accuracy requirements of 50 - 150 km.

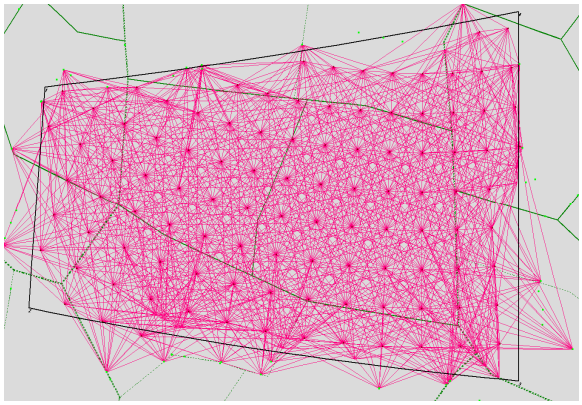


Fig. 4 Pseudo-en-route structure of free flight area

Figure 5 shows optimal route synthesis result from USTR (Tyumen) to UHMM (Magadan).

This route passes through the conditional area (area boundaries are shown in black). The optimal route is shown in red, the route with part of free flight – in purple.

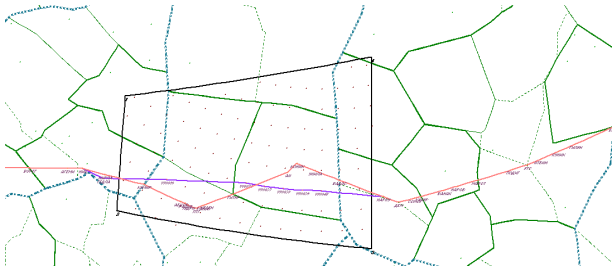


Fig. 5 The route synthesis result within free flight area

The flight time was reduced to 7 minutes, the route length was reduced to 106 km.

One of the promising on-board applications (based on new features and functions of the observation and increased situational awareness on board) the rerouting problem is considered on board in response to dynamic constraints. At present the algorithm is adapted and implemented to rerouting on board for flying around closed areas or bad weather areas. Figure 6 shows the route synthesis result on the on-board monitor screen image. The current route

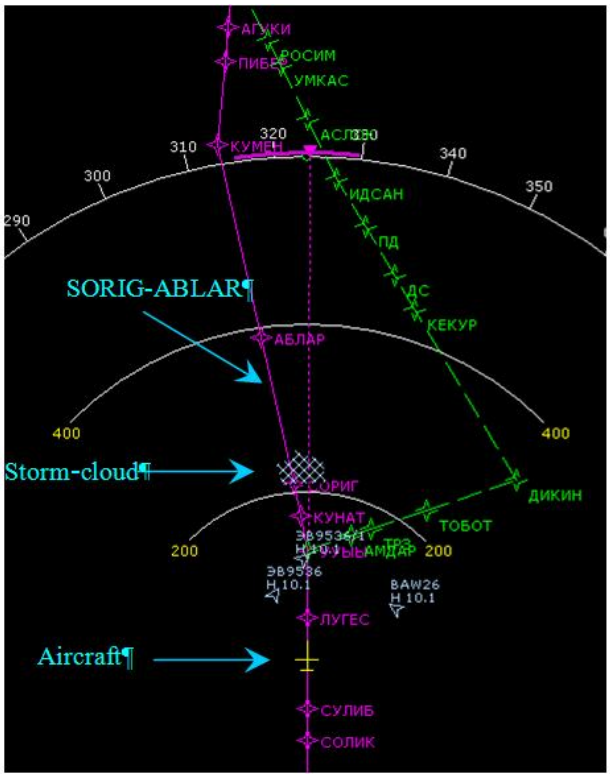


Fig.6 Route synthesis result on board

(purple color) crosses storm-cloud at the segment "SORIG-ABLAR". The new route (green color) avoids the storm-cloud using the available airways structure and returns to the planned route.

6 Conclusion

The proposed algorithm and program tools show high speed, universality, performance and efficiency, this is verified by results of its wide application to the various problems solving. The algorithm has several particular features:

- it is universal and provides the route planning by various criteria, under different restrictions, for different rules, with a various detail level aircraft performance characteristics;
- the combined scheme of the route construction is implemented in two stages: an optimal 2D route construction on a graph, then - the vertical profile and time synthesis;
- the possibility of accounting for dynamically changing spatial and temporal restrictions of different nature is implemented;
- designed initially to route building on a airways structure it is extended to the routes, including free flight areas.

The effectiveness of software tools is confirmed by results of its use during many years in planning tasks solved in the interest of services of the Russian Federation ATM. Currently, the algorithm is generalized to the case of time constraints account in the form of slot allocation for waypoints overflying or slots of arrival.

This work is executed with financial support by RFBR (project 11-08-00841-a).

References

[1] O.V. Degtjarev. “Methodical and algorithmic problems 4D routes construction”, “News of the Russian Academy of Sciences. The theory and control systems.”, 2006, №5.

[2] O.V. Degtjarev, E.V. Filenkova. “Decision support rerouting tool for air traffic flows”, VI International Scientific-technical conference “Chkalov readings”, Speech abstracts, Egor’evsk, 2007.

3. N. Christofides. Graph Theory. An algorithmic approach. Moscow, "Mir", 1978

CEAS 2011

Fatigue and Damage Tolerance

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Numerical Evaluation of the Compressive Residual Strength in Impact-Damaged Composite Laminates

A. Baldi, A. Airoldi, M. Crespi, P. G. Nettuno, G. Sala
Dipartimento di Ingegneria Aerospaziale – Politecnico di Milano –
Campus Bovisa, Via La Masa 34, 20156 Milano, Italy

Keywords: *delamination, cohesive elements, damage tolerance, compression-after-impact*

Abstract

The paper applies a numerical approach based on a cohesive zone model to investigate the compression behaviour of impact-damaged composite laminates. The initial impact-induced interlaminar damage configuration is introduced in a finite element scheme, which models all the interlaminar layers of the laminate. A modelling technique is applied to analyse the growth of interlaminar damage by means of explicit analyses. A series of numerical simulations are performed and compared to experimental evidences, in order to verify the capability of the proposed cohesive modelling approach to qualitatively predict the compressive failure mode of the impact-damaged laminates as well as their residual compressive strength.

1 Introduction

The advantages of fibre reinforced composites in aircraft manufacturing and their increasing diffusion motivate the efforts to develop effective numerical approaches to analyse and study the structural behaviour of composite parts. In particular, the prediction of damage development and of damage effects in composite represents one of the fundamental issues to fully exploit the potential of such materials. The weakest point of conventional composite laminates is their low out-of-plane strength and the resulting susceptibility to delamination due to a series of circumstances, such as manufacturing defects, low energy impacts [1, 2] and denting load conditions. Delamination in carbon composites can lead to significant strength reduction of up to 60% under compressive loading conditions, even if damage cannot be detected by visual inspection [3, 4, 5]. For such a reason, the concept of barely visible damage and the correspondent degradation of the compressive strength of impact-damaged composite laminates play a fundamental role in the design of damage tolerant structures. In fact, the most diffused design philosophies consider the presence of a certain level of damage in the structure and consequently degrade the admissible stress and strain levels by means of experimental tests on damaged specimens and components. The large amount of required experimental data motivates the development of analytical and numerical approaches for the evaluation of damage onset, propagation and for the prediction of residual load carrying capabilities of damaged composite elements.

The paper presents a peculiar numerical approach based on cohesive zone models, which is particularly suitable to analyses carried out by

means of explicit time integration scheme [6]. Such approach is applied to analyse the compression-after-impact response of unidirectional carbon fibre composite laminates. The experimental activity performed on compression-after-impact tests is first presented and the numerical approach is then described and preliminarily assessed considering conventional tests for interlaminar toughness characterization. A procedure was developed to generate configurations of multiple delaminations which can realistically represent impact-induced delaminations patterns, given a projected-delamination area. The propagation of such initial impact-induced damages at multiple ply interfaces is modelled by means of the proposed cohesive approach. The performed compression-after-impact analyses are presented and compared to experimental evidences.

2 Experimental Procedure

2.1 Impact Tests

The standard test method to evaluate the residual post-impact compressive strength of composite laminates [7] was applied to a set of nine specimens. The test requires specimens with 100 x 150 mm^2 in-plane dimensions, which were manufactured by means of a standard autoclave process. The adopted lay-up is a quasi-isotropic $[45/0/-45/90]_{4S}$ sequence, consisting of 32 unidirectional carbon epoxy pre-preg plies (IM7/8552). The resulting average thickness of the produced laminates was 4.02 mm.

In order to evaluate the effects of the presence of delaminated areas and their dimensions on the compressive residual strength of composite laminates, six of the nine specimens were subjected to low energy impact tests. In particular, three specimens were impacted, in accordance with the drop-weight configuration suggested by the standard [8], at a nominal energy level of 11.5 J and the other three ones, at a nominal energy level of 17 J. The remaining three specimens were directly subjected to the compression test, without the production of any delaminated areas (*pristine specimens*).

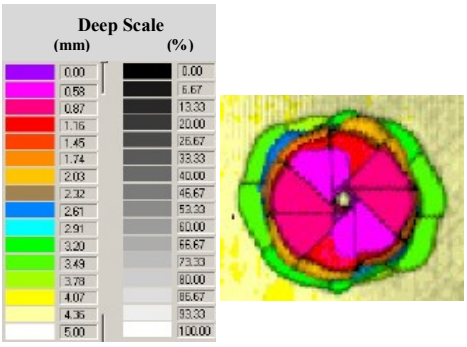


Fig. 1 Ultrasonic C-Scan map of the interlaminar damage for the *Spec*#2 impacted at *Ep* = 11.5 J

The extension of the impact damaged areas was determined by ultrasonic C-Scan. All the six C-Scan images exhibit a typical roughly circular delamination area. Damage morphology, as presented in Fig. 1, agrees with the typical damage pattern of thick composite laminates impacted at low energy level. Delaminated areas increase from the impacted face towards the rear side of the specimens ("*pine tree*" shape) and have a typical oblong or peanut shapes, which is oriented in the direction of the fibres of the lower ply at that interface [2]. The C-scan images allowed the calculation of the projected-delamination areas (A_{Del}), which are summarized in Table 1. The three specimens impacted at E_p = 17 J are characterized by higher average values (A_{Del} = 1251 mm^2) of projected-delamination areas with respect to the three specimens impacted at E_p = 11.5 J (A_{Del} = 1066 mm^2). The results impact tests are presented with more details in [9].

2.2 CAI Tests Setup

The specimens were compressed until failure using the test support fixture shown in Fig. 2, which was designed in accordance with the indications provided by the standard [7].

Lateral anti-buckling guides prevented the global instability of the laminate providing simply supported conditions for the longitudinal edges (*z* direction in Fig. 2). Each specimen was placed leaving a clearance of 0.8 mm ÷ 1.5 mm between its lateral edges and the side walls of

the support fixture, in order to permit the lateral expansion of the panel during the compression phase, due to the effect of Poisson's coefficients.

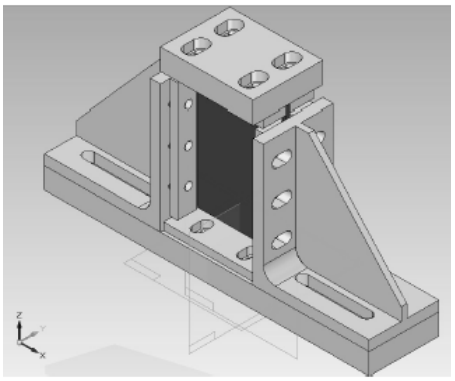


Fig. 2 Sketch of the support fixture for CAI tests

Two couples of gages were disposed in a back-to-back configuration, in agreement with the standard [7]. The strain gage locations are indicated in Fig. 3.

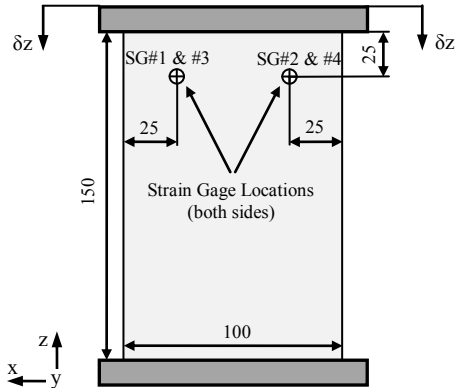


Fig. 3 CAI test setup (dimensions in mm)

The compression tests were performed adopting a 250 kN MTS/858 hydraulic system. The load was applied to the upper lid at a cross head constant speed of 1.0 mm/min until the failure of the laminate. Displacement was recorded by means of the MTS/858 LVDT transducer. A pre-loading phase, up to a load level of about 450 N, was performed for all the tests in order to ensure the contact of all the loading surfaces and the correct alignment of specimen, by checking the signals provided by the back-to-back strain gages.

2.3 CAI Tests Results

All the compression tests performed on the impact-damaged specimens showed a typical failure mode, characterized by the formation of a blister in correspondence of the delamination zone due to the tendency of the unsupported laminas to locally buckle. Figure 4, which is referred to *Spec*#1, reports the failure mode.

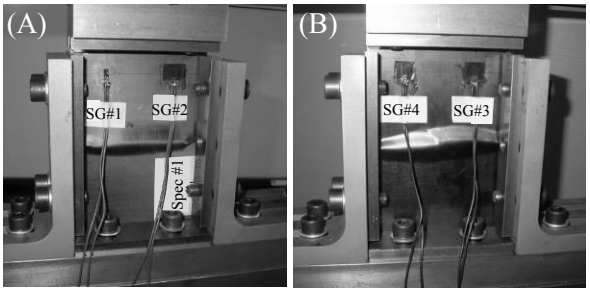


Fig. 4 Front (A) and rear (B) view of the local buckling failure mode of *Spec*#1

The failure by local delamination-buckling is confirmed by the evolution of the bending coefficients calculated from the strain gage signals. Such coefficients represent the percentage of global bending strain at gage locations and were evaluated by means of the expressions reported in Eq. 1. The evolution of the coefficients against the overall displacement is presented in Fig. 5 for the specimens impact-damaged at E_p = 11.5 J.

$$B_{yI} = \frac{\varepsilon_{SG\#1} - \varepsilon_{SG\#3}}{\varepsilon_{SG\#1} + \varepsilon_{SG\#3}} \cdot 100$$
$$B_{yII} = \frac{\varepsilon_{SG\#2} - \varepsilon_{SG\#4}}{\varepsilon_{SG\#2} + \varepsilon_{SG\#4}} \cdot 100 \qquad (1)$$

It can be seen that all the curves are included in the limit range of ± 10% until the final failure of the laminate, which is the limit imposed by ASTM standard for test validity [7]. Indeed, the coefficients remain close to zero, until a sudden deviation occurs, which corresponds to the failure of the specimens.

On the contrary, a global buckling phenomenon has affected the pristine specimens during the compression tests leading to the collapse of the laminate. The failure mode is

reported in Fig. 6-(A) and it is consistent with a global buckling of the plate.

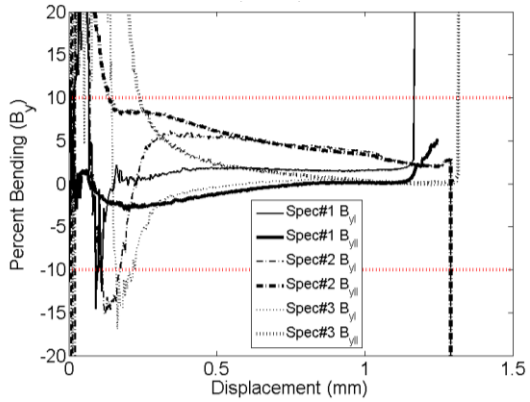


Fig. 5 Percent bending (B_y) for the specimens impact damage at E_p = 11.5 J

Such hypothesis is confirmed by the trend of the percent bending coefficients, which progressively increase and exceed the limit range of ± 10%, before the collapse of the laminate, as presented in Fig. 6-(B).

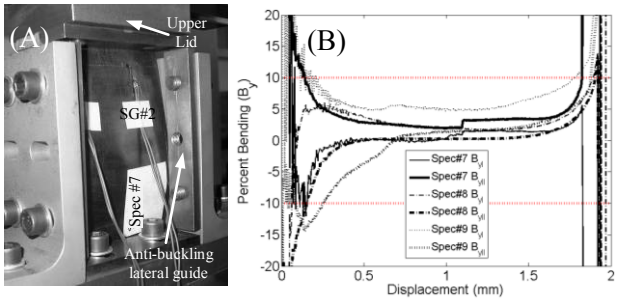


Fig. 6 Global buckling failure mode of *Spec*#7 (A) and percent bending (B_y) for the pristine specimens (B)

Figure 7-(A) and 7-(B) compare the two different failure modes at the lateral edges of *Spec*#1 and *Spec*#9, respectively. In particular, the final damage configuration, shown in Fig. 7-(A), is referred to an impact-damaged laminate and it is characterized by the presence of interlaminar damages and fibres failure in a very localized zone. Only the external lamina seems to be delaminated, whereas failure in the core resembles a compressive failure. The failure mode in Fig. 7-(B) is typical for all the pristine specimens. It is characterized by a global curvature of the laminate. The failure, which can be reasonably attributed to bending, occurred in correspondence of the unsupported end section between the lateral anti-buckling guides and the upper lid.

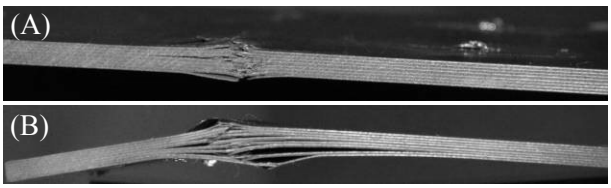


Fig. 7 Particular of the local (A) and global (B) buckling failure mode of *Spec*#1 and of *Spec*#9

Figure 8 presents the force vs. displacement responses for all the tested specimens. The results point out that the global stiffness is not affected by the presence of interlaminar damage as the specimens exhibited a linear response until the failure with the same slope. As it was previously observed, the final failure of the laminate occurred by a sudden load drops, and before this point no deviation from the linear trend was recorded.

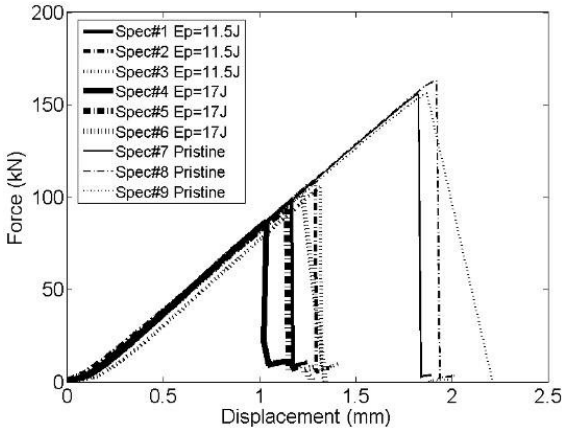


Fig. 8 Experimental force vs. displacement responses of the tested specimens

Table 1 Residual compressive strength of impact-damaged specimens

Spec. #	E_p (J)	A_{Del} (mm^2)	F_r (kN)	σ_r (MPa)
1	11.5	1187	97.9	244.8
2	11.5	959	109.0	272.5
3	11.5	1051	105.1	262.8
4	17	1477	86.9	217.3
5	17	1290	93.5	233.8
6	17	987	102.4	256

The pristine specimen failed at load levels higher than 150 kN, whereas the strength data of impact-damaged specimens are summarized in Table 1. As expected the strength decrease with the increment of projected-delamination area. The minimum recorded strengths are about 60% of the level referred to the pristine specimens.

3 Modelling Technique

An approach based on cohesive zone model was used to analyse the behaviour of the impact-damaged laminates subjected to compressive action in the CAI tests. The numerical model is required to represent the response of a laminate in presence of several pre-damaged interlaminar layers and all of such damages can potentially propagate under the action of compressive loads. The prediction of the sudden failure indicates that the transition from a no-growth or stable growth condition to a dynamic unstable propagation mode must be correctly captured and this motivates the choice of an explicit time integration scheme for the analyses.

Cohesive elements can be applied to model delamination, but such approach is complicated by several factors, as the need of very high penalty stiffness levels that eventually lead to unacceptable computational cost and to numerical instabilities.

For such a reason, an alternative modelling technique was applied to develop a robust and reliable numerical approach. In the proposed modelling technique, composite laminates are described as a collection of sub-laminates that are represented by bi-dimensional elements, such as membrane or shell elements. The area of each sub-laminate cross section is considered lumped at its mid-plane.

The kinematic variables used to describe the fracture process in mode I, II and III are the relative displacements at the mid-planes of such sub-laminates (Fig. 9-(A)). The components of the relative displacement vector ($\Delta=U^{+}-U^{-}$) can be associated to the three possible modes of fracture openings, as indicated in Eq. 2.

$$\Delta_I = \begin{cases} \Delta_z & if \quad \Delta_z > 0 \\ 0 & if \quad \Delta_z \le 0 \end{cases} ; \quad \Delta_{II} = \Delta_x \; ; \; \Delta_{III} = \Delta_y \qquad (2)$$

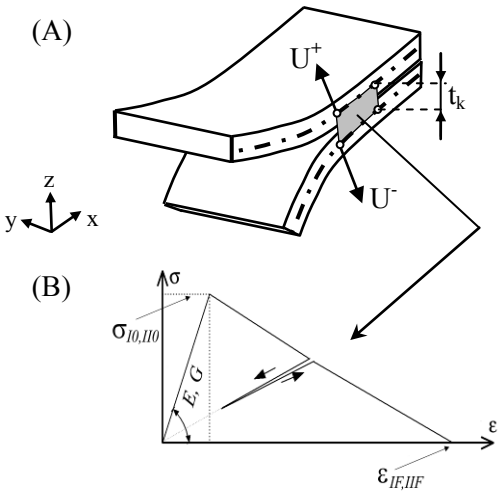


Fig. 9 Description of the interlaminar fracture process (A) and bi-linear constitutive law (B) in the proposed modeling approach

Within a small strain assumption, the vector Δ can also be related to the average value of the out-of-plane strain components acting in the material volume between the sub-laminates with thickness t_k (Eq. 3):

$$\boldsymbol{\varepsilon} = \left\{ \varepsilon_{zz} \quad \gamma_{xz} \quad \gamma_{yz} \right\}^{T} = \boldsymbol{\Delta} / t_k \qquad (3)$$

Conventional solid elements are used to connect the bi-dimensional elements and to represent both the interlaminar layer between adjacent sub-laminates and the adhesive interface. The response of the connection element can be modelled as in Eq. 4, where a scalar damage variable was introduced to model the fracture process.

The resulting finite element scheme, made of bi-dimensional elements which represent the composite plies connected by solid elements with null in-plane response, guarantees the translational equilibrium of each sub-laminate included in the laminate lay-up, as presented in [6, 10]. The capability of the technique to model the bending response of composite laminates was assessed by several numerical benchmarks and by the numerical-experimental correlations reported in [6, 10].

$$\begin{Bmatrix}\sigma_{zz}\\ \tau_{xz}\\ \tau_{yz}\end{Bmatrix}=\begin{bmatrix}E_{zz} & 0 & 0\\ 0 & G_{xz} & 0\\ 0 & 0 & G_{yz}\end{bmatrix}(1-d)\begin{Bmatrix}\varepsilon_{zz}\\ \gamma_{xz}\\ \gamma_{yz}\end{Bmatrix} \qquad (4)$$

The onset and the evolution of the damage variable introduced in Eq. 3 can be set to model the strength and the toughness of the interlaminar layers. To accomplish such objective, the proposed approach exploits the links between the critical energy release rates and the energy dissipated in the damage process [11, 12]. Taking into consideration Eq. 3, the link between the strain-stress response and the energy release rates for fractures opening in mode I and II can be expressed as in Eq. 5.

Equations 3 and 5 allow modeling the strength and the toughness of the cohesive interface by attributing a bi-linear stress vs. strain response to the solid element, as the one presented in Fig. 9-(B).

$$\begin{aligned}\int_0^{\infty}\sigma_{zz}d\Delta_I &= t_k\int_0^{\infty}\sigma_{zz}d\varepsilon_{zz}=G_{Ic}\\ \int_0^{\infty}\tau_{xz}d\Delta_{II} &= t_k\int_0^{\infty}\tau_{xz}d\gamma_{zx}=G_{IIc}\end{aligned} \qquad (5)$$

In the constitutive model, properties in mode II and mode III are considered identical. Damage onset in processes evolving in pure mode I or mode II occurs beyond the peak stress σ_{I0} and σ_{II0}, which represent the axial and shear strength of the interlaminar layer. Mixed mode processes are addressed by introducing the mixed-mode criterion proposed by Benzeggagh and Kenane [13] (*B-K criterion*). The criterion is based on the expression reported in Eq. 6, where η represents an experimental parameter, that was set equal to 1.45 (η = 1.45) [11, 12]. In such form, the cohesive law was implemented in a *Fortran Vumat* subroutine to be linked to the *3DS/Abaqus Explicit Code*.

Globally, it should be observed that the proposed approach does not introduce free surfaces that do not exist at the beginning of the computation and does not require the traditional cohesive elements with infinitesimal or zero thickness. As a consequence, all the material characteristics can be directly determined on the basis of physical considerations without introducing non-physical penalty stiffness, which can severely affect computational costs and the numerical performances.

$$G_C = G_{IC}+\left(G_{IIC}-G_{IC}\right)\left(\frac{G_{II}}{G_T}\right)^{\eta} \qquad (6)$$
$$\text{with: } G_T=\left(G_I+G_{II}\right)$$

Accordingly, the approach can be considered well suitable to model laminates with several interfaces, all of them representing a potential location of interlaminar damage and of their subsequent propagation.

4 Preliminary Assessment of the Modelling Approach

This section presents the application of the modelling technique to a standard interlaminar fracture test, in order to compare the accuracy and the computational costs of the approach with the ones of a model based on conventional zero-thickness cohesive elements.

The considered test is a Double Cantilever Beam test (DCB) [14], which was performed to evaluate the interlaminar toughness properties of a plane weave fabric pre-preg, with AS4 carbon fibres and a Hexcel 8552 epoxy-resin (AS4/8552). The DCB specimen was modelled by applying the proposed numerical technique considering a 150 mm long laminate, which was represented as an assembly of 12 $[0]_2$ sub-laminates. The sub-laminates were connected by 11 interlaminar layers having a thickness of 0.56 mm each. Conventional shell elements (elements *S4R* [15]) was used to model the in-plane behaviour of each sub-laminate, while conventional solid elements (elements *C3D8R* [15]) were adopted to model the out-of-plane response of each interlaminar layer. Shell elements were characterized by a purely elastic orthotropic material model with the properties of the fabric pre-preg reported in Table 2. The previously described constitutive model, based on the cohesive law, was attributed to solid elements by using the following physical out-of-

plane moduli: E_{33} = 7.8 GPa, G_{23} = G_{13} = 3.0 GPa and the following experimental toughness levels: G_{Ic} = 1064 J/m^2, G_{IIc} = 1750 J/m^2 [9, 10]. Quasi-static analyses were performed using an explicit FE code (*3DS/Abaqus Explicit, v. 6.8*) by smoothly increasing the velocity attributed to the loading tip of the specimen in order to prevent the dynamic excitation of the laminate. No mass scaling factor was adopted in the analyses.

The reference DCB model developed according to a more conventional numerical approach, based on the cohesive elements available in the *3DS/Abaqus* code, was obtained by means of an uniform grid, that is made of 1.25 x 1.25 x 0.56 mm^3 solid elements (elements *C3D8R* [15]). Each solid element represents two fabric plies, so that 12 solids elements through the laminate thickness are used. A single layer of conventional cohesive elements (elements *COH3D8* [15]) with a thickness of 0.02 mm was set in the mid-plane of the laminate. Cohesive elements were characterized with the following penalty stiffness values: K_{nn} = 120 GPa/mm K_{ss} = K_{tt} = 45 GPa/mm, which were defined increasing opportunely the elastic moduli related to a hypothetic resin rich layer (E = 7.8 GPa; G = 3 GPa) in rigorous accordance with the guide lines provided in [16] for an effective calibration of penalty stiffness in cohesive layers. The adopted material properties and boundary conditions were identical to the ones applied to the non-conventional model. A density of 1200 kg/m^3 was assigned to the cohesive elements, in agreement with the physical value of a resin rich layer.

Figure 10 reports the qualitative comparison between the two modelling techniques in terms of stable interlaminar damage evolution at the end of the analysis performed using the same interlaminar strength values of σ_{I0} = 20 MPa and σ_{II0} = 50 MPa. The global responses, in terms of load vs. displacement curves, indicate a good correlation with the experimental values.

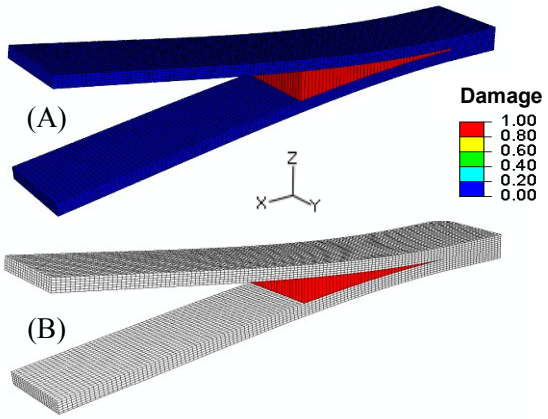


Fig. 10 Interlaminar crack propagation for proposed (A) and conventional (B) cohesive DCB model

In fact, both the approaches are able to identify the load level at the onset of the crack propagation and also to correctly follow the reduction of the load carrying capability of the specimen during the stable propagation of the defect [9, 10].

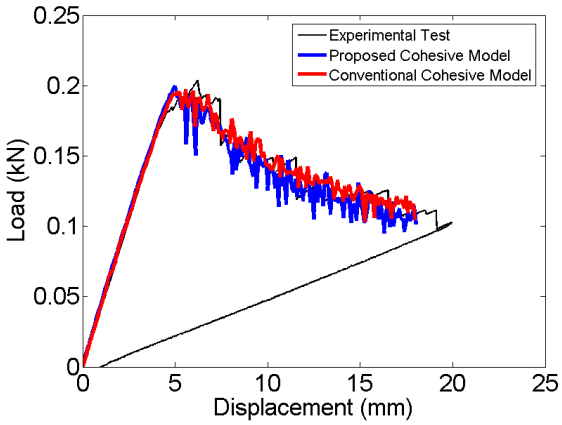


Fig. 11 Load vs. displacement comparison between proposed and conventional cohesive DCB model

Table 2 Elastic properties of carbon fabric (AS4/8552) and unidirectional (IM7/8552) pre-preg

	E_{11} (GPa)	E_{22} (GPa)	E_{33} (GPa)	ν_{12}	ν_{23}	ν_{31}	G_{12} (GPa)	G_{23} (GPa)	G_{13} (GPa)
FB (AS4/8552)	61.0	61.0	7.8	0.05	0.3	0.3	4.4	3.0	3.0
UD (IM7/8552)	155.0	10.0	10.0	0.3	0.3	0.3	5.1	3.8	5.1

However, the penalty stiffness levels which were introduced in the conventional cohesive approach led to significant negative effects on the computational time cost, as confirmed by the data reported in Table 3. As it was previously indicated, both analyses were performed without the adoption of any mass scaling factor [9]. Though the density of the cohesive layers could be artificially increased to enhance the stable time step, the required mass scaling factor, to match the computational performance of the unconventional technique, should be more than 100, thus affecting the overall mass of the specimen and inducing undesired oscillations, as proved by a series of sensitivity studies. Moreover, it should be remarked that only one interlaminar layer is included in the used conventional model. The introduction of several layers would require an increment of the adopted penalty stiffness as the thickness of the connected sub-laminates decreases, as indicated by the guidelines reported in [16]. Overall, the presented result point out the advantages of the proposed modelling technique both in terms of modelling efforts and of computational performances.

Table 3 Computational cost comparison between proposed and conventional cohesive DCB model

Numerical Approach	d.o.f.	Stable Time Step	Tot. CPU time
Proposed Cohesive Model	182864	1.356 E-7	10 h
Conventional Cohesive Model	108876	7.684 E-9	100 h

5 Numerical Simulation of CAI Tests

The proposed cohesive modelling approach was adopted to simulate the behaviour of impact-damaged laminates in the CAI tests using *3DS/Abaqus Explicit, v. 6.8*. The finite element scheme was developed at the ply level, in order to correctly represent the interlaminar damage pattern configuration, adopting a regular in-plane mesh grid with a 2 x 2 mm^2 typical dimension, which is shown in Fig. 13-(A). The resulting 31 interlaminar layers were modelled by conventional solid elements (elements *C3D8R* [15]) endowed with the bi-linear law described in section 3. Shell elements were characterized by a purely elastic orthotropic material model with the properties of unidirectional pre-preg reported in Table 2. Interlaminar solid elements were completely defined with the following physical out-of-plane moduli: E_{33} = 10.0 GPa, G_{23} = G_{13} = 4.5 GPa and with the following toughness properties, which were evaluated in a previously experimental campaign: G_{Ic} = 344 J/m^2, G_{IIc} = 583 J/m^2 [9, 10, 17, 18, 19]. The interlaminar strength were set to σ_{I0} =35 MPa and σ_{II0} = 70 MPa. Such levels had been previously identified and validated in several applications, which are reported in [9, 10, 17, 18, 19], involving the identical unidirectional material.

An analytical approach was developed and adopted in order to reproduce, in the CAI model, a realistic configuration of the pre-damaged area, taking into consideration the experimental evidences provided by C-scan identification of the impact damage. In fact, the correct simulation of the compression behaviour of an impacted laminate requires the correct reproduction of the impact-damaged zone both in terms of shape and size [20]. Moving from the theory proposed by Liu [21], which is based on the assumption that delaminations are driven only by the bending stiffness mismatching between two adjacent plies, a multiple delamination configuration with peanut shape damaged zone was generated.

The procedure, which is described in detail in [9], was implemented in a Matlab™ script and it allows the identification of the shape of impact-induced delaminations in the CAI specimen, once that the size of the projected-delamination area is provided by experimental evidences (ultrasonic C-Scan map). Figure 13-(B) shows the in-plane projection of the interlaminar damage introduced in the initial state of the specimen model, whereas a qualitative sketch of the delamination configuration through the laminate thickness in presented in Fig. 13-(C).

Two pre-damaged models were developed, in order to analyse the compressive behaviour of the impact-damaged laminate at the two different energy levels considered in the low

energy impact tests. The impact-damaged model related to the energy of 11.5 J was characterized by a pre-damaged area of about 1100 mm^2, whereas the one relate at the energy level of 17 J was characterized by a pre-damaged area of about 1300 mm^2, in agreement with the average values provided by the experimental evidences.

The lateral anti-buckling guides as well as the top and bottom guides of the CAI model were modelled by means of rigid elements (elements *R3D4* [15]), shown in Fig. 13-(A)-(B). All the surfaces elements of the specimen were set in contact with the surface of the guides. A penalty contact algorithm [15] with an exponential pressure over-closure relationship was adopted for all of the contacts defined in the model. No friction was introduced. The rigid body representing the moveable upper lid of the test fixture was moved in the downward direction. Quasi-static analyses were performed, by smoothly increasing the velocity ($\dot{\delta}_y$) attributed to the upper guide. A mass scaling factor was applied in order to reduce the total computational time cost of the final model, which consists of 240450 elements.

5.1 CAI Tests Numerical - Experimental Correlation

The numerical results of the models representing impacted configurations were compared with the experimental evidences presented in section 2. Both the numerical models, referred to the two impact energy levels of 11.5 J and 17 J, showed the same failure mode in good agreement with the experimental evidences. The local delamination-buckling failure was correctly modelled in correspondence of the pre-damaged zone.

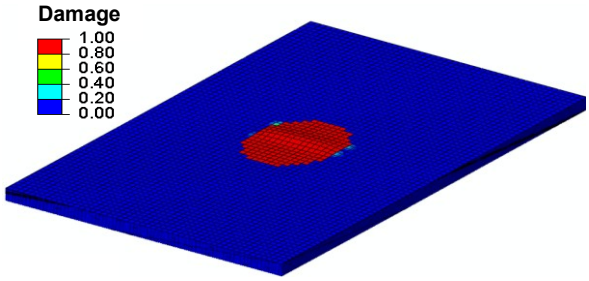


Fig. 12 Blister formation at 94 ÷ 95 % of the maximum compressive load level

Figure 13 presents the development of a blister in correspondence of this site at 94 % ÷ 95 % of the maximum load level. The blister showed the tendency to propagate from the central zone of the impact delaminated specimens towards the lateral edges, in good agreement with the experimental evidences. The propagation of delamination is represented in the sequence of Fig. 14-(A)-(B)-(C). Such figures are referred to 96%, 98% and to 100% of the maximum compressive load of the laminate, respectively. No global instability phenomena were exhibited by the model up to this limit load. The numerical load suddenly drop as the damage propagate laterally. In particular, Fig. 14-(D) presents the interlaminar

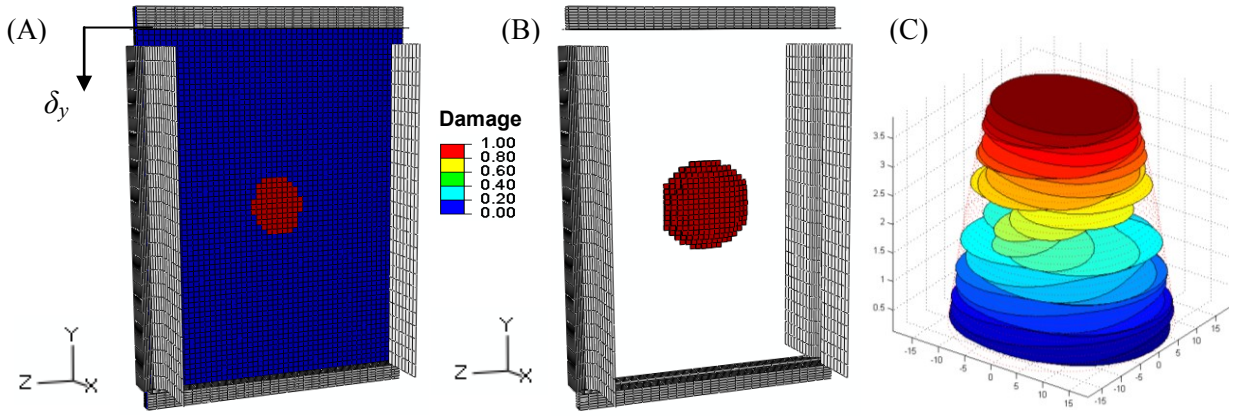


Fig. 13 CAI model (A) particular of impact-damaged area (B) and idealised delamination configuration (C)

damage configuration after the numerical maximum load level was reached.

A quantitative correlation of specimen stiffness and strength was carried out on the basis of the load vs. strain trend, as reported in Fig. 15-(A)-(B). The trends shown in Fig. 15-(A) as well as the ones presented in Fig. 15-(B) confirm a good level of correlation between numerical and experimental evidences. In fact, the global stiffness of the impact-damaged laminates as well as their residual compressive strength are correctly captured by the proposed cohesive modelling approach. The numerical stiffness levels are not affected by the size of delaminated areas, in agreement with the results of the CAI tests campaign.

On the other hand, the maximum compressive load and consequently the compressive residual strength are significantly affected by the initial delamination. The numerical results predict a maximum compressive load of about 106 kN for the specimen impact-damaged at E_p = 11.5 J and a maximum compressive load value of about 92 kN for the one impact-damaged at E_p = 17 J. These values, as confirmed by the force vs. strain curves reported in Fig. 15-(A) and in Fig. 15- (B), are in good agreement with the average values of the experimental tests of 104 kN and 94 kN, referred to the two impact energy levels.

6 Conclusions

In the present study, the compression-after-impact response of unidirectional carbon fibre composite laminates was investigated and numerically modelled. Quasi-static numerical simulations of the tests were performed by means of explicit finite element analyses. The impact damage configuration of the laminates was introduced in FE scheme, which is characterized by the presence of all interlaminar layers, by means of an idealised multiple-delamination configuration, by considering the projected-delamination area evaluated by C-Scan measurements. A peculiar numerical approach, based on a cohesive zone model, was adopted to model the propagation of delamination and the specimen collapse. The advantages of the modelling approach were discussed and assessed considering a standard interlaminar fracture test. The numerical models of the CAI specimens were validated against the experimental evidences in terms of failure mode as well as of load vs. strain response. Two different levels of projected-delamination area, derived by different impact energy levels, were considered. The global stiffness of impact-damaged specimens, the failure mode and the residual compressive strength can be correctly identified by the proposed technique. The appreciable correlation between numerical and experimental results confirms the capability of the proposed approach to correctly capture the propagation of delamination, which is considered as the critical mechanism that drives the collapse of an impacted laminate under compressive load conditions.

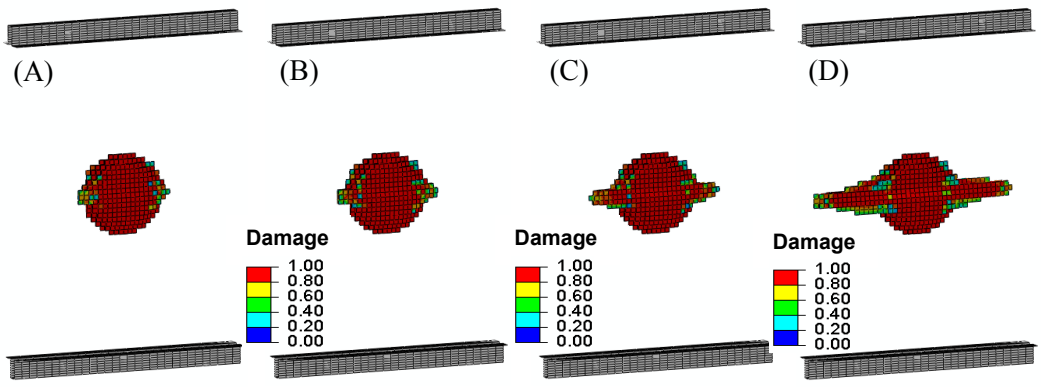


Fig. 14 Interlaminar damage evolution: 96% (A); 98% (B); 100% (C) of the maximum compressive load and configuration beyond the limit load (D)

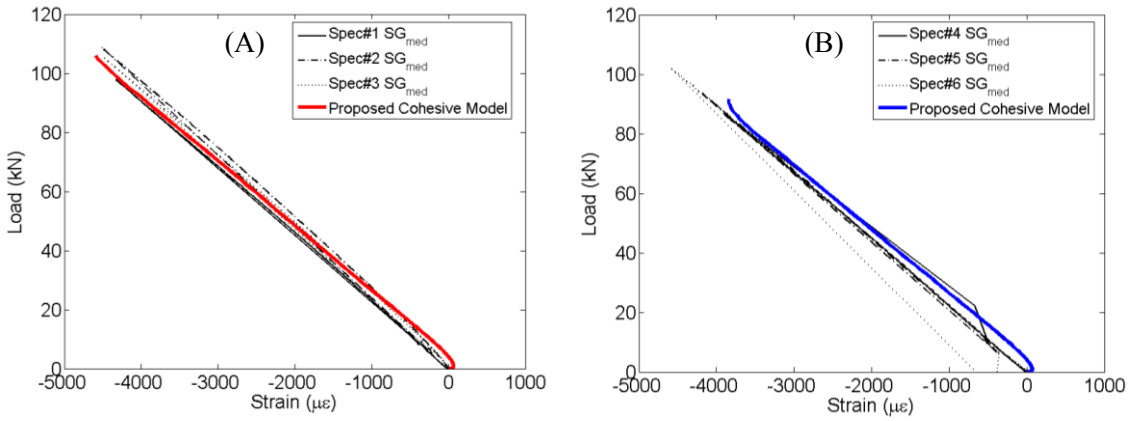


Fig. 15 Load vs. strain numerical-experimental correlation for the CAI model impact-damaged at E_p = 11.5J (A) and E_p = 17J (B)

References

[1] Sala, G., "Impact behaviour of heat-resistant toughened composites", *Composites, Part B*, Vol. 31, 2000, pp. 161-173.

[2] Abrate S., *Impact on composite structures*, Cambridge University Press, Cambridge (England), 1998.

[3] Davies G.A.O., Olsson R., "Impact on composite structures", *Aeronautical Journal*, Vol. 108, No. 1089, 2004, pp. 541–63.

[4] Sjögren A., Krasnikovs A., Varna J., "Experimental determination of elastic properties of impact damaged in carbon fibre/epoxy laminates", *Composite, Part A*, Vol. 32, No. 9, 2001, pp. 1237–42.

[5] Cartié D.D.R., Irving P.E., "Effect of resin and fibre properties on impact and compression after impact performance of CFRP", *Composite, Part A*, Vol. 33, 2002, pp. 483-93.

[6] Airoldi A., Sala G., Bettini P., "Evaluation of a numerical approach for the development of interlaminar damage in composite laminates", *Composites Science and Technology*, In Press, DOI: 10.1016/j.compscitech.2009.10.011.

[7] ASTM D7137/D7137M-05, "Standard test method for compressive residual strength properties of damaged polymer matrix composite plies", American Society for Testing and Materials (ASTM), West Conshohocken (PA, USA), 2005.

[8] ASTM D7136/D7136M-05, "Standard test method for measuring the damage resistance of a fibre reinforced polymer matrix composite to a drop-weight impact event", American Society for Testing and Materials (ASTM), West Conshohocken (PA, USA), 2005.

[9] Baldi A., "A numerical approach to estimate damage tolerance in composite laminates", Ph.D. Dissertation, Aerospace Engineering Dept., Politecnico di Milano, Italy, 2010.

[10] Airoldi A., Baldi A., Daleffe M., Sala G., Basaglia M., "Analyses of delamination in composite laminates in standard tests and low energy impacts", Proceedings of the 20th AIDAA Congress, Milano, June 29-July 3, 2009, pp. 1-20.

[11] Dávila, C.G., Camanho P.P. and de Moura M.F.S.F., "Mixed-mode decohesion elements for analyses of progressive delamination", Proceedings of the 42nd AIAA/ASME/ASCII/ AHS/ ASC Structures, Structural Dynamics, and Material Conference and Exhibit, Seattle, Washington, April 16-19, 2001.

[12] Camanho P.P., Dávila C.G. and de Moura M.S.F.S., "Numerical simulation of mixed-mode progressive delamination in Composite Materials", *Journal of Composite Materials*, (2003), Vol. 37, No. 16: 1415-1438.

[13] Benzeggagh, M.L. and Kenane M., "Measurement of Mixed-Mode Delamination Fracture Toughness of Unidirectional Glass/Epoxy Composites With Mixed-Mode Bending Apparatus". *Composites Science and Technology*, Vol. 56, 1996, pp. 439-49.

[14] ASTM D5528, "Standard Test Method for Mode I Interlaminar Fracture Toughness of Unidirectional Fiber Reinforced Polymer Matrix Composite", American Society for Testing and Materials (ASTM), West Conshohocken (PA, USA), 2001, (Reapproved 2007).

[15] Abaqus®, "Theory and User's Manuals" Hibbit, Karlsson & Sorensen, Pawtucket, USA, 2004.

[16] Turon A., Davila C.G., Camanho P.P., Costa J., "An engineering solution for mesh size effect in the simulation of delamination using cohesive zone models". *Engineering Fracture Mechanics*, Vol. 74, 2007, pp. 1665-82.
[17] Baldi A., Airoldi A., Domenichini P., Crespi M., Sala G., "Analysis of progressive failure of composite T-Joints", Proceedings of the 14th European Conference on Composite Materials (ECCM-14), Budapest, June 7-10, 2010, pp. 1-12.
[18] Baldi A., Airoldi A., Crespi M., Iavarone P. and Bettini P., "Modelling competitive delamination and debonding phenomena in composite T-Joints", *Procedia Engineering*, Vol. 10, 2011, pp. 3483-3489.
[19] Baldi A., Airoldi A., Daleffe M., Sala G., Basaglia M., "A numerical approach for low energy impacts on composite laminates using FE explicit codes", Proceedings of the 17th International Conference on Composite Materials (ICCM-17), Edinburgh, July 27-31, 2009, pp. 1-12.
[20] Craven R., Iannucci L., Olsson R., "Delamination buckling: A finite element study with realistic delamination shapes, multiple delaminations and fibre fracture cracks", *Composites: Part A*, 2010, doi:10.1016/ j.compositesa. 2010. 01. 019.
[21] Liu D., "Impact-induced delamination – a view of bending stiffness mismatch", *Journal of Composite Materials*, Vol. 22, No. 7, 1988, pp. 675–92.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Design Methodology of Damage Tolerant Stiffened Composite Panels with Minimum Cost

M. Corvino✉, L. Iuspa, A. Riccio, F. Scaramuzzino
Dipartimento di Ingegneria Aerospaziale e Meccanica
Seconda Università degli Studi di Napoli, Via Roma 29, 81031 – Aversa (CE), Italy
✉maria.corvino@unina2.it

Keywords: *multi-objective optimization, damage tolerance, damage resistance, evaluation cost, genetic algorithms.*

Abstract

In this paper a novel design strategy for composite structures is presented. The multi-objective methodology reaches the optimal topological configurations with minimum weight and the maximum global buckling load of both damage tolerant and damage resistant, stiffened composite panels. Once performed the optimization stage, the resulting Pareto optimal solution are subsequently analyzed by a specific off-line ranking procedure, specifically developed for cost evaluation of the in-service life .

1 Introduction

In the recent years, an increasingly demand for composite materials in the aircraft industry has been observed. The use of laminate fiber composite structures with Carbon Fiber Reinforced Polymers (CFRP) is very popular for lightweight constructions, due to their superior specific modules (strength-to-density and stiffness-to-density ratios). Nevertheless, to fully exploit the weight saving potential of these materials, an appropriate tailoring with related lay-up definition must be performed. At the same time, the observance of assigned structural constraints requires an almost mandatory use of advanced structural optimization tools.

A common problem in composite aerospace structures is the development of damage tolerance and damage resistance design criteria that guarantee certifiable levels of safety in typical impact scenarios as bird strike, the dropping of tools during maintenance work or the impact of runway debris [1-3].

Damage in composites is difficult to detect because delamination and fiber breakage can occur inside the material, while the damage signs are often invisible at the external surface. This insidious condition may cause a serious decrease in material strength and critical load when instability phenomena happen, because the local buckling of a sub-laminate can take place before the global buckling becomes appreciable.
Furthermore, the successful application of composite materials to aircraft structures is somewhat limited by manufacturing, maintenance and repair costs.

For these reasons, besides weight and performance requirements, the adopted structural solutions should also take in account the cost evaluation. In particular, a damage tolerant design approach can be adopted to reduce the maintenance costs during the in-

service life. The impact damage tolerance deals with the ability to sustain a given level of damage with a limited reduction of the structural performance.

According to these considerations, in this paper a novel design optimization strategy for composite structures is presented. Starting by the design procedures previously developed in [4], the suggested methodology reaches the optimal topological configurations with the minimum weight and the maximum buckling load of damage tolerant, stiffened composite panels [5]. The multi-objective optimization process is based on an in-house genetic algorithm code [6], in conjunction with the Ansys™ FEM engine. In order to determine the optimal configuration of damage tolerant stiffened panels, a parametric FEM model has been implemented in the Ansys environment [7-8].

The search is focused on panels both buckling compliant and damage tolerant. Here, damage tolerance is considered as the capability of the panel to develop a delamination growth initiation load higher than the global buckling load.

The paper is organized as follows: in section 2, the methodologies employed for the multi-objective optimization procedure and a schematic overview of the adopted genetic algorithm are shown.

In section 3, the general description of the structural test-case adopted in numerical application and related modeling is presented.

Furthermore, in section 4, the results of the optimization of the stiffened composite panel under the assigned constraints are given. Finally, in the last paragraph the conclusive considerations are presented.

2 Theoretical analysis of multi-objective optimization procedure

In this paper a multi-objective optimization procedure, based on the adoption of genetic algorithms, is presented. Specifically, the Pareto-optimal configurations with minimum weight and maximum global buckling load of a damage-tolerant and damage-resistant, stiffened composite panel have been found. The numerical procedure is based on an in-house optimization code, *ProGenie*, in conjunction with the *Ansys*™ FEM code.

2.1 Multi-objective optimization

Multi-objective problems arise in many reliability-based and risk-informed activities of system design, operation, maintenance and regulation.

The solutions developed in these activities have to satisfy several goals and specifications, which generally are conflicting. As consequence, there is not a unique, optimal solution satisfying all objectives; rather, a set of suitable solutions can be identified, none being the best for all objectives.

A multi-objective solution consists of a collection of alternative solutions of equivalent quality in the absence of further information regarding the relative importance of each objective functions. This set of optimal solutions in the space of the decision variables is called Pareto Set.

The region in the space of the objective functions defined by all Pareto set points is called the Pareto Front.

Pareto-optimal solutions consist of *non-dominated solutions*, this statement meaning that each objective function can not be further improved without simultaneously decreasing the quality of at least another objective in the set.

2.2 Genetic Algorithm

To perform the optimization task a bit-masking oriented genetic algorithm (BMOGA) has been used. In this special version of GA, classical procedural genetic operators are replaced by a unique synthetic generator able to emulate all the traditional crossover and mutation schemes, plus many others suggested by specific requests, simply varying the arguments of a special boolean function [9].

In this context the features offered by BMOGA are mainly used to define a multi-level

crossover operating in parallel with different methods on partitions of the chromosomal strings.

In order to increase the global efficiency during initial generations of the genetic evolution, discrete and real variables are handled in parallel with a bit-by-bit crossover and one-cut crossover, respectively.

This choice allows a fast reshuffling of short sub-strings related to integer variables without destroying useful schemata correlated to real variables. For the remaining generations, a classical one-cut crossover is used on the whole genetic string.

In Figure 1 a schematic view of the multi-level crossover arrangement is shown. Binary strings of the individuals are ideally split in two partitions (also disjointed if this is the case) and each partition is tuned up with the proper crossover type.

For the optimization task a in-house program, *ProGenie*, was used in conjunction with the *Ansys*™ FEM code.

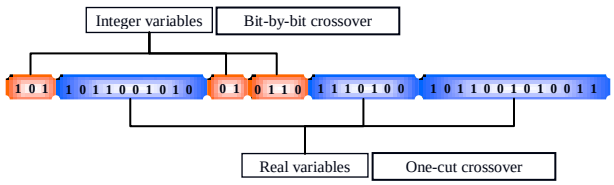


Fig. 1: multi-level crossover layout for integer and real design variables

ProGenie uses a fully bit-masking based genetic engine and offers some very powerful tools for the final user to implement complex genetic operators.

2.3 Linear buckling analysis of the composite structure

The finite element method is used for determining the buckling load factor of the laminated composite structure. A linearised buckling analysis has been performed, *i.e.* the structure is assumed with no geometrical imperfections, and the buckling load found will be an upper limit for the real value [10].

A linear buckling analysis can estimate the maximum load that can be sustained prior to a structural instability or collapse happens.

In structures characterized by stiffness K, the buckling condition can be written as:

$$det({}^{t}\underline{K})=0 \qquad (1)$$

The main hypothesis in linearised buckling analysis is that the stiffness changes proportionally with additional load increments:

$${}^{t}\underline{K}={}^{t-\Delta t}\underline{K}+\lambda({}^{t}\underline{K}-{}^{t-\Delta t}\underline{K})={}^{t-\Delta t}\underline{K}+\lambda\Delta\underline{K} \qquad (2)$$

where ${}^{t}\underline{K}$ and ${}^{t-\Delta t}\underline{K}$ are the stiffness at two different steps and λ is the normalized load factor. From eqs (1) and (2):

$${}^{t-\Delta t}\underline{K}\,\underline{\Phi}=\Delta\underline{K}\,\underline{\Phi} \qquad (3)$$

where λ are the eigenvalues representing the buckling load, while Φ are eigenvectors representing the buckling modes.

2.4 Damage Resistant

In order to evaluate the impact damage resistance of a stiffened composite panel, the Delamination Threshold Force Method has been applied [11-13]. This technique can be split in three different phases:

- Application of Zhang & Davis formula for determination of the delamination threshold force:

$$P_c=\sqrt{\frac{8\cdot\pi\cdot E\cdot t^3\cdot G_{IIc}}{9\cdot(1-\upsilon^2)}} \qquad (4)$$

  where E_r is the mean flexural modulus of elasticity of the laminate, and P_c is the critical impact force (Zhang & Davis theory).

- Linear FEM analysis to evaluate the load-displacement curve when the impact force is applied statically.
- Energy Balance application to determine the minimum impact energy needed to start the delamination.

$$W_c = \frac{k}{n+1} \cdot \left[\frac{P_c}{k} \right]^{\frac{n+1}{n}} + \int_0^{\delta_{max}} P \cdot d\delta \qquad (5)$$

where *Wc* is the threshold impact energy which onsets delamination, *k* is contact stiffness, *n* is a constant, $\frac{k}{n+1} \cdot \left[\frac{P_c}{k} \right]^{\frac{n+1}{n}}$ is the energy absorbed by local contact deformation and $\int_0^{\delta_{max}} P \cdot d\delta$ is the energy developed by the global deformation of structure.

2.5 Damage Tolerance

Delamination is one of the predominant forms of failure in laminated composites, due to a lack of reinforcement in the thickness or a manufacturing defect, that can lead to a significant reduction of the compressive load the structure is able to carry on.

Under a compressive load, composite structures and delaminated stiffened panels can exhibit localized buckling phenomena.

More precisely, for bay delaminations of appropriate size and placed at a relatively small depth, two main scenarios can occur:

I. The thinner sub-laminate buckles locally while the thicker sub-laminate remain unaffected, as shown in Figure 2A;
II. Both sub-laminates buckle, exhibiting a sort of mixed buckling mode (see Figure 2B).

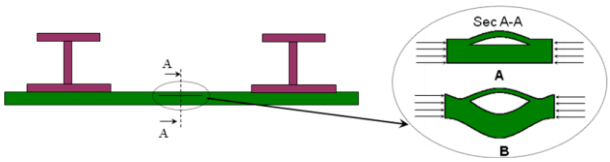


Fig. 2A-B Scenarios of localised buckling phenomena

Delamination can be handled like cracks in brittle materials; hence propagation initiation of a delamination at a generic location along the delamination front can be considered governed by the following expression:

$$G = -\frac{\Delta E}{\Delta A} = G_c \qquad (6)$$

defining the *Energy Release Rate (ERR)* as the variation of the potential energy *ΔE* associated to an increase *ΔA* of the delaminated area *A*.

For the first delamination scenario, when the thickness of the buckled sub-laminate is much smaller than the total laminate thickness (as shown in Figure 3), the contributions of fracture modes II and III can be neglected.

The ERR distribution can be assumed mostly influenced by the out-of-plane displacement along the delamination front at the first delamination buckling.

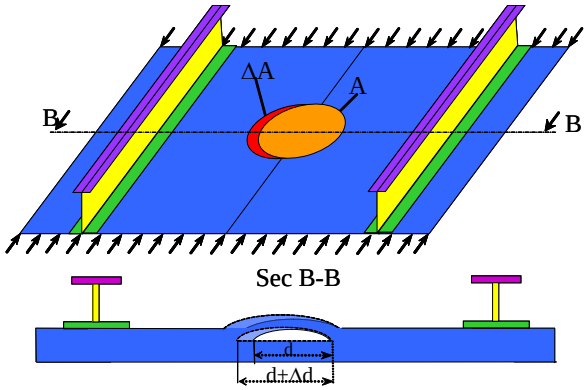


Fig. 3 Propagation initiation of a delamination

The methodology used for the preliminary design of damage tolerant composite stiffened panels, formulated by Riccio and Zarrelli [14], can be schematized in the following steps:

I. calculation of the local buckled shape (mode) of the thinner sub-laminate (delamination buckling) by means of a linearised buckling analysis;
II. determination of the location along the delamination front with maximum ERR, according to the out-of-the-plane buckling displacements;
III. calculation of the energy released by the structure at delamination growth initiation by means of a linearised

buckling analysis, plus two linear static analyses;

IV. evaluation of the delamination growth initiation load;

V. evaluation of the ERR distribution along the delamination front at growth initiation, according to the out-of-the-plane buckling displacements.

2.6 Cost Evaluation

Composite materials are characterized by high performance such as specific strength and stiffness, becoming some of most important aeronautic materials. Nevertheless, an extensive application of composite materials is actually restricted, besides airworthiness regulations limitations, by high manufacturing and repair costs.

Cost estimation before the producing process is mandatory in order to reduce the manufacturing incidence, and it is performed evaluating several trade-off scenarios related to product design, material choice, process efficiency, and investment requirements. Therefore, both a suitable cost evaluation method and a reliable cost information are needed to assess the potential of composite materials.

To preliminarily determine the total Life Cycle Cost (LCC) for stiffened composite panels, the following relation can be used:

$$LCC = MC + \Pr R \cdot RC \qquad (7)$$

where *MC* is the manufacturing cost and *PrR*RC* the repair cost, being *PrR* the probability of a repair during the service life, and *RC* the cost associated with a single repair [15-18].

The total manufacturing costs depend on material cost, labour cost and non-recurring costs (such as tooling) evaluated over the actual number of components:

$$MC = mtc + lbrc + nonrecc \qquad (8)$$

Both repair costs and the probability of a repair during the service life depend on the subpart involved (bay, stringer) and the impact location onto the panel. Consequently, expression (7) for *LCC* has to be slightly adapted (*n* being the number of impact locations considered in the analysis):

$$LCC = MC + \sum_{i=1}^{n} \Pr R_i \cdot RC_i \qquad (9)$$

The probability of a repair of a certain location in the panel is a function of the probability of an impact of given energy at the same location and the associated damage resistance:

$$\Pr R_i = f(\Pr I_i, DR_i) \qquad (10)$$

3 Application

The numerical model of the stiffened panel herein analysed has fixed overall dimension and characterized by several parameters which control stiffeners topology, geometric properties and lamination stack sequence.

The stiffened panels are optimized for minimum weight and maximum global buckling load, and subject to several constraints to meet structural and functional requirements, as the damage tolerant capability.

In Figures 4A-B-C some topological configurations of the analysed stiffened panel are shown.

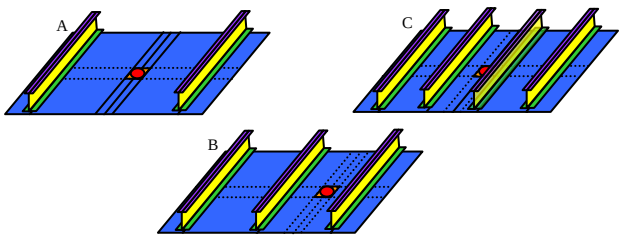


Fig. 4A-B-C: stiffened panel stringer layout: two-stringer (A); three-stringer (B); four-stinger (C)

The fixed parameters are:

i) Panel width (360 mm);
ii) Panel length (300 mm);
iii) Stringer height (30 mm)
iv) Stringer type (I-shaped);
v) Cap length (30 mm);
vi) Foot length (55 mm).

In this optimization process, the stiffened panel is characterized by a fixed delamination with a diameter of 50 mm, located in the center bay at 15% of skin depth. In Figures 5 and 6, the stringer details and the delamination geometry are shown.

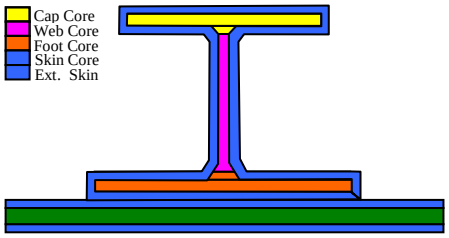


Fig. 5: stringer type

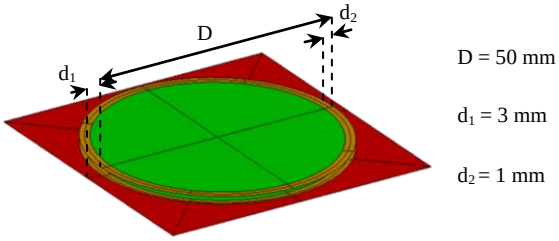


Fig. 6: delamination geometry

In order to obtain a symmetrical and balanced composite laminate, the panel is created assembling a thick core surrounded by two thin external skins. Each external skin consist of two eight symmetrical and balanced layers [19,20].

The FEM model needs to be accurate, especially near at the delamination front, to correctly predict the local buckling modes of the delaminated part of the structure.

The delaminated stiffened panel is modelled by means of eight-node layered shell elements, rigid links and interface elements based on multipoint constraints [21].

Quadratic layered elements are used to model skin, stringers and delaminated sub-laminates; CERIGs (rigid constraints) are used to connect the stringer feet to the skin.

Along the delamination front, rigid elements are also used to connect sub-laminates to the rest of the skin. Contact elements are adopted to connect the delaminated area to the remaining part of the panel without transition mesh.

In Figure 7 the FEM model and the applied natural boundary conditions for a simply supported compressed panel are shown.

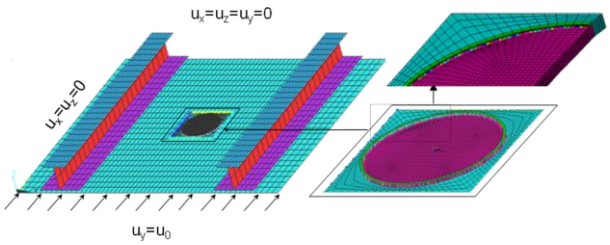


Fig. 7: Panel boundary conditions and delamination fem model

4 Results

A multi-objective optimization procedure has been defined to minimize the structural mass and maximize the global buckling load of the panel.

For this analysis the Pareto optimality criterion has been applied.

In Tables 1 and 2, design parameters with related ranges and resolutions are summarized.

In Table 3 main GA static parameters (individuals, couples, generations) and the envelope of some dynamic parameters (selection and crossover type, mutation rate, scaling fitness) are also reported.

The resulting discrete Pareto set of the multi-objective optimization is shown in Figure 8.

Design Parameters	
n_{stiff}	Number of stringers
SLQ	External skin staking sequence
$N_{ply\text{-}skin}$	Number of couples in the core skin
SCS	Core skin staking sequence
$N_{ply\text{-}web}$	Number of couples in core web
SCW	Core web staking sequence
$N_{ply\text{-}cap}$	Number of couples in core cap
SCC	Core cap staking sequence
$N_{ply\text{-}foot}$	Number of couples in core foot
SCF	Core foot staking sequence

Table 1: design parameters denomination

DV	*Type*	*Min*	*Max*	*Resolution (bit)*				
n_{stiff}	integer	1	4	2	2	2	2	2
SLQ	integer	1	8	3	3	3	3	3
$N_{ply\text{-}skin}$	integer	1	8	3	3	3	3	3
SCS	real	0	1	10	10	12	12	12
$N_{ply\text{-}web}$	integer	1	2	1	1	1	1	1
SCW	real	0	1	10	10	12	12	12
$N_{ply\text{-}cap}$	integer	1	2	1	1	1	1	1
SCC	real	0	1	10	10	12	12	12
$N_{ply\text{-}foot}$	Integer	1	2	1	1	1	1	1
SCF	real	0	1	10	10	12	12	12

Table 2: Design parameters, ranges, type and dynamic resolution

Individuals=200; couples=100; generations=70					
Envelope	0	1	2	3	4
Selection	Roulette-wheel				
Mutation	0.04	0.04	0.04	0.06	0.06
Crossover*	1/2	1/2	1	1	1

*½ One-cut crossover/ Bit-by-bit crossover

Table 3: optimization parameters dynamic envelope for multi-objective task

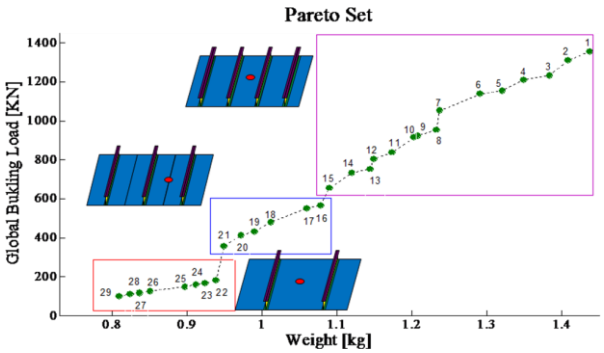


Fig. 8: topologically partitioned Pareto-front for minimum weight and maximum global buckling load

The main features of optimal designs are summarized in Table 4, 5 and 6.

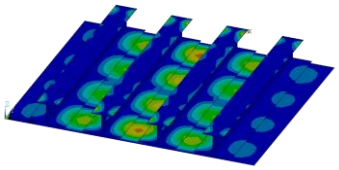
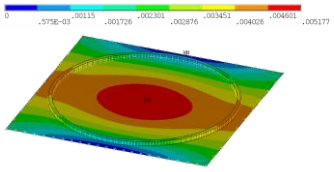
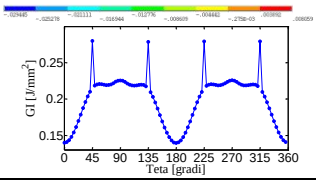
Design	**1**
Weight [kg]	1.43
Global Buckling Load [kN]	1354
Global Buckling Displacement	
Local Buckling Displacement	
G mode I contribution	

Table 4: optimal design 1 main features (four-stinger configuration)

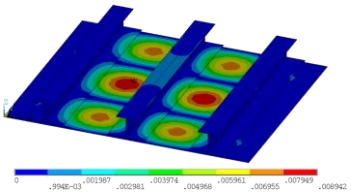
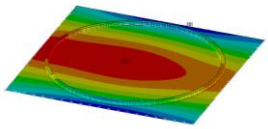
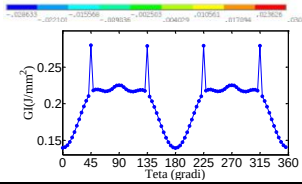
Design	**16**
Weight [kg]	1.07
Global Buckling Load [kN]	564
Global Buckling Displacement	
Local Buckling Displacement	
G mode I contribution	

Table 5: optimal design 16 main features (three-stinger configuration)

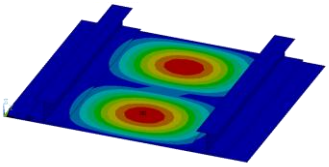
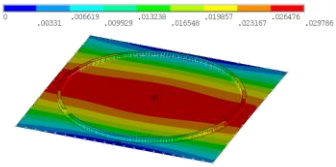
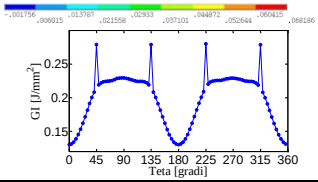
Design	**22**
Weight [kg]	0.93
Global Buckling Load [kN]	180
Global Buckling Displacement	
Local Buckling Displacement	
G mode I contribution	

Table 5: optimal design 22 main features (two-stinger configuration)

By comparing the different Pareto solutions from Figure 7, one may notice that the different optimum configurations can be essentially traced back to a minimum weight or a maximum buckling performance, plus others compromise solutions. These equally ranked solutions give the designer some choice to select a suitable panel configuration for specific proposes.

5 Cost evaluation

The life cycle cost are evaluated considering all the main contributions (raw material, labour and repair). The repair costs are determined taking into account the probability of impact, and are considered dependent on the impact location.

Finally, a specific off-line tag procedure has been applied to arrange the numerical simulation results in a suitable graphical form with evaluation costs highlighted, as shown in Figures 9 and 10.

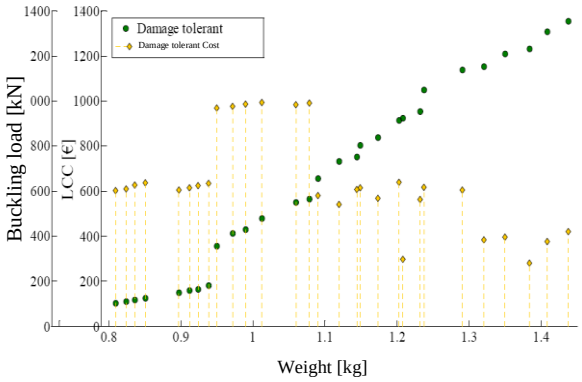


Fig. 9: re-arranged Pareto front with evaluation costs

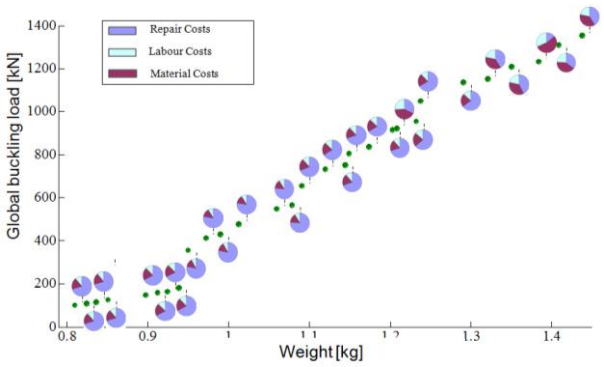


Fig. 10 Evaluation costs

Analyzing the cost-tagged Pareto-front, configurations with minimum weight and minimum global buckling load are less damage tolerant than configurations with maximum weigh and maximum global buckling load. As consequence, the repair cost of the first type of panel is greater than the other one.

6 Conclusion

In this work, a multi-objective optimization procedure, able to find the minimum weight and the maximum global buckling load for composite stiffened panels with impact damage tolerance constraints has been shown.

The specific parametric model has proved to be very flexible and effective while the damage tolerant methodology resulted very fast and simply to implement.

The use of a multi-level crossover allowed by BMOGA has further increased the overall performance, showing an higher convergence speed of the numeric procedure than standard GAs.

The multi-objective optimization has been able to find several Pareto-optimal configuration of the stiffened panel, all damage tolerant.

Finally, the off-line tag procedure has proved to be an useful tools for the evaluation of the costs, giving to the designer an appreciable help in the choice of the preliminary configuration.

References

[1] W. J. Cantwell and J. Morton, *The impact resistance of composite materials – a review*, Composites, 22(5), (1991), 347-62.

[2] Aymerich F., Bucchioni A. and Priolo P., *Impact Behaviour of quasi-isotropic graphite-PEEK laminates*, Key Engineering Materials, 144, (1998), 63-74.

[3] Yang Y., Im, K-H., and Park, J-W., *On impact characteristic evaluation in CF/Epoxy composites under falling-weight impact loadings*, Materials Science Forum, v. 449-452, proceedings of the 3rd International Symposium on Designing, Processing and Properties of Advanced Engineering Materials, 2004, p. 79-87.

[4] M. Corvino, L. Iuspa, A. Riccio, F. Scaramuzzino; *Weight and cost oriented multi-objective optimization of impact damage resistant stiffened composite panels*, Computers and Structures 87 (2009) 1033-1042.

[5] Budiansky B., *On the Minimum Weight of Compression Structures*, International Journal of solids and Structures, Vol. 36, 1999, pp. 3677-3708.

[6] L. Iuspa and F. Scaramuzziono, *A Bit-Masking Oriented Data Structure for Evolutionary Operators Implementation in Genetic Algorithms*; Soft Computering, Vol. 5, 2001, pp.55-68.

[7] David E Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Pubblishing Company, Inc, 1989.

[8] LeRiche R. and Hafka R. T., *Optimization of Laminate Stacking Sequence for Buckling Load Maximization by Genetic Algorithm*, AIAA Journal, Vol. 31, 1993, pp.951-956.

[9] Iuspa L., Scaramuzzino F. and Petrenga P., *Optimal Design of an Aircraft Engine Mount via it-Masking Oriented Genetic Algorithms*, Advances in Engineering Software, Vol. 34, 2003, pp. 707-720.

[10] O.F. Hunghes, B. Ghosh, Y. Chen, *Improved prediction of simultaneous local and overall buckling of stiffened panels*, Tjin-Walled Structures, 2004.

[11] Davies G. A. O., Zhang X., *Impact damage prediction in carbon composite structures*, International Journal of Impact Engineering, Vol. 16, 1995, pp. 149-170.

[12] G. A. O. Davies, X. Zhang, G. Zhou and S. Watson, *Numerical modeling of impact damage*, Composites, Vol. 25, 5, 1994, pp. 342-50.

[13] G.A.O. Davies, D. Hitchings and J. Ankersen, *Predicting Delamination and Debonding in Modern Aerospace Composite Structures*. Composite Science and Technology 2006; 66:846-854.

[14] A. Riccio, M. Gigliotti. *A novel Numerical Delamination Growth Approach For The Preliminary Design of Damage Tolerant Composite Structures.* Journal of Composite Materials 2007; 41(16): 1939-1960.

[15] Creemers R. J. C., *Phase 1 Milestone Report – The panel optimization and manufacture – DAMOCLES 2 deliverable for work package 1.11*, NLR-CR-2004-054, National Aerospace Laboratory NLR, Amsterdam, 2004.

[16] Creemers R. J. C., *Numerical design and optimization of CFRP panels in a modular wingbox – DAMOCLES 2 deliverable for work package 1.12*, NLR-CR-2003-442, National Aerospace Laboratory NLR, Amsterdam, 2003.

[17] Creemers R. J. C., *Numerical design and optimization of an all-composite wing-DAMOCLES 2 deliverable for work package 2.1*, NLR-CR-2004-429, National Aerospace Laboratory NLR, Amsterdam, 2003.

[18] Creemers R.J.C. *DAMOCLES 2 Deliverable for Work Package 1.11*, NLR- CR 2004-054, NLR, Amsterdam, 2004.

[19] Fukunaga, H and Vanderplaats G. N., *Stiffness Optimization of Orthotropic Laminated Composites Lamination Parameters*, AIAA Journal, Vol.29, 1991, pp. 641-646.

[20] Nagendra, S., Haftka, R.T. and Gurdal, Z., *Stacking Sequence Optimization of Simply Supported Laminates with Stability and Strain Constraints*, AIAA Journal, Vol. 30, 1992, pp. 2132-2137.

[21] Venkataraman S., Lamberti L., Hafka R. T. and Johnson T. F., *Challenges in Comparing Numerical Solution for Optimum eight of Stiffened shell*, Journal of Spacecraft and Rockets, Vol. 40, 2003, pp.183-192.

Fatigue Behavior of Composite Stiffened Components Working in Post-buckling

Duo Zou and Chiara Bisagni
Politecnico di Milano, Dipartimento di Ingegneria Aerospaziale
Via La Masa 34, 20156 Milano, Italy
E-mails: zou@aero.polimi.it, chiara.bisagni@polimi.it

Keywords: *Fatigue, Delamination, Buckling, Tests.*

Abstract

The concept of damage tolerant design has gained increasing acceptance and could result in better competitiveness in the field of aerospace. Such design considers fatigue as a critical phenomenon. Indeed, fatigue loading can induce delamination initiation and propagation. Currently a large research effort is being spent on fatigue damage of composite structures. However, there are few investigations focusing on how delamination propagates under cyclic buckling.

The present work considers a study of fatigue behavior of composite structures, which characterizes mainly the onset and growth of delamination. The structures here investigated are composite specimens with single L-shaped stiffeners. The research begins with the design, manufacturing and finite element analysis under static loading. Then the testing methodology under cyclically compressive loading is proposed. During the future tests, the influence of repeated buckling will be investigated in terms of global structural responses, and of delamination growth per fatigue cycle with respect to the strain energy release rate.

1 Introduction

Motivated by the increasing use of composites in primary structural components, lots of research is dedicated to fatigue damage behavior [1-3] of fiber-reinforced composite materials when they operate under static or cyclic buckling load [4, 5]. As a result of the anisotropic and heterogeneous material characteristics, the behavior of composite materials is complicated and different from that one of homogeneous and isotropic materials [6].

In metals, the stable stage of gradual invisible deterioration spans nearly the complete life time and the stiffness of a metal remains almost unaffected during the fatigue loading.

In fiber-reinforced plastic composites made of reinforcing fibers embedded in a tough matrix, the damage can initiate very early and different types of damage may appear (e.g. fiber fracture, fiber breakage, matrix cracking, fiber buckling, fiber-matrix interface failure and delamination…). The gradual deterioration of a fiber-reinforced composite in the damage zones can bring to a significant loss of stiffness and lead to a continuous redistribution of stress.

During the damage process of laminated composites, delamination is regarded as the most common failure mechanism [7]. Hence, the ability to predict delamination behavior is important for establishing damage tolerance criteria [8].

In general, there are three possibilities [9] to study the fatigue phenomena in composite materials: fatigue life models, which use the information from the actual degradation mechanisms to derive S-N curves or Goodman-type diagrams [10]; fracture mechanics models, which basically study the crack growth per load cycle versus calculated energy release rate as included in Paris law [11, 12]; and damage mechanics models, which use one or more damage variables related to measurable manifestations of damage [13].

The second scenario based on Paris law is taken into consideration in this study. The fatigue behavior of single-stiffener specimens with a co-cured L-shaped stiffener is investigated under cyclically compressive loading. The specimens studied here represent a level of complexity between the coupon elements and the structural components, in a so called building block approach [14]. The specimens are designed, analyzed and manufactured, and will be tested until high number of cycles [13]. During the tests, plots of measured delamination progression rate with respect to the calculated energy release rate will be obtained.

The finite element model in static analysis is performed using ABAQUS [15] and the predicted buckling load provides some guidance for the range of the cycling loading during the tests.

2 Description of specimens

Four specimens have been manufactured and are ready for the fatigue tests.

2.1 Geometry of the specimens

The specimens investigated in this study present an L-shaped stiffener co-cured with the skin panel. The width of the specimens is stiffener pitch representative, and is taken equal to 120 mm. The specimens have a total length of 300 mm, and both the stiffener web and the stiffener flange are equal to 30 mm. Moreover, a Teflon-insert with a length equal to 20 mm is introduced in the specimen in order to simulate an initial delamination between the skin and the stiffener.

The skin panels consist of a quasi-isotropic laminate with eight layers in stacking sequence [0/45/-45/90]s for a total thickness of 1 mm, while the stiffeners are composed of sixteen layers with a quasi-isotropic layup [0/45/-45/90/90/-45/45/0]s which results in a total nominal thickness of 2 mm. Both the skin and the stiffeners are made from HexPly®8552 unidirectional IM7 carbon fiber material.

A schematic of the specimens is given in Fig. (1).

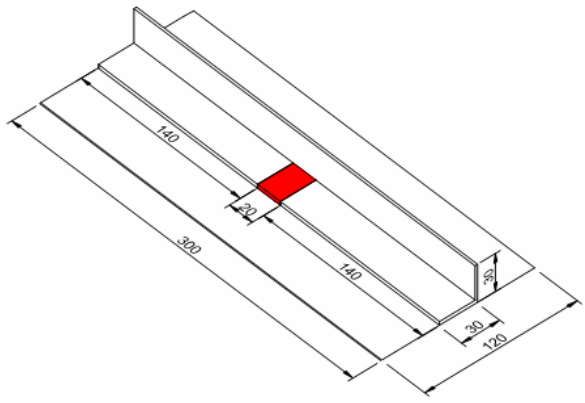


Fig. 1: Drawing of the specimen

The mechanical properties of the composite material are reported in Table 1.

Table 1: Material properties - IM7/8552

E_{11}(GPa)	E_{22}(GPa)	G_{12}(GPa)	ν
150	9.08	5.29	0.32

2.2 Manufacturing

The components were manufactured at the Dipartimento di Ingegneria Aerospaziale of Politecnico di Milano. The laminates describing the fiber direction 0, 45, and 90 degrees are shown in Fig. (2).

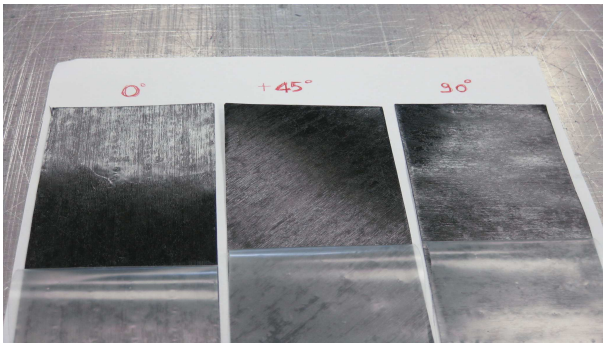


Fig. 2: Orientation of laminates

The normal steps were used during the manufacturing, such as laminating by placing each prepreg layer on top of the previous layer in the pre-defined orientation as shown in Fig. (3). Then the specimens are enclosed in a sealed vacuum bag and all the air is removed from the bag as illustrated in Fig. (4). At the end, the specimens were cured in an autoclave for six hours according to the material specifications. The autoclave is shown in Fig. (5).

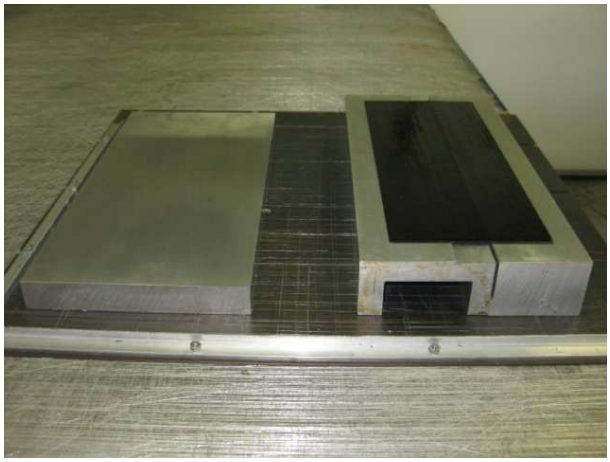


Fig. 3: Skin-stiffener laminating on the aluminum mold



Fig. 4: Vacuum bag in the auto-clave



Fig. 5: Autoclave and control panel

2.3 Measurement

The measurements taken on the four specimens are reported in Table 2. In particular the Table reports the average measured thickness on the skin, the stiffener web and the stiffener foot, as well as the mass.

Table 2: Measurements of the specimens

No	Skin (mm)	Web (mm)	Foot (mm)	Width (mm)	Length (mm)	Mass (g)
1	1.06	1.85	3.30	121	301	113.3
2	0.90	2.00	2.90	123	299	110.6
3	0.99	1.98	3.00	120	298	109.9
4	1.05	1.85	3.00	120	299	110.6

The back view of a specimen with 20 mm Teflon-insert in the middle is shown in Fig. (6). Resin tabs are then added to both ends of the specimens for the application of the load during the tests.

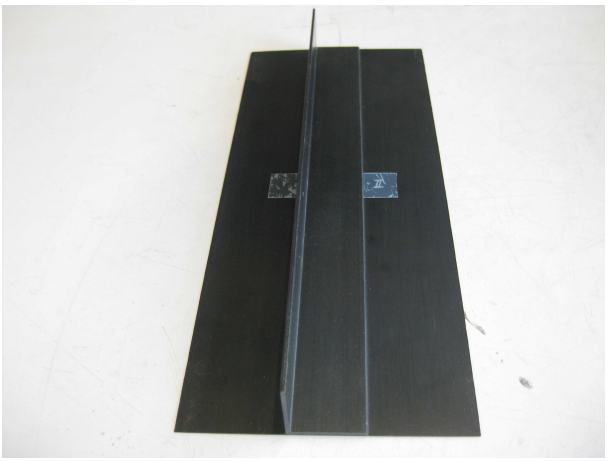


Fig. 6: Specimen back view

3 Finite element analysis

The finite element model is developed to determine the buckling load, the collapse load and the damage in the postbuckling regime.

3.1 Model

Four-node shells elements S4R with six degrees of freedom at each node are used to model the stiffened panel. The finite element mesh size is 1.0 mm by 1.0 mm and presents 49800 elements and 43862 nodes. The finite element mesh view is illustrated in Fig. (7).

The boundary conditions applied during the analysis step constrain one of the two ends of the specimen, while allow free shortening in the longitudinal direction on the other end.

The two major damage modes are taken into account by the model: skin-stiffener separation and intralaminar damage. The skin-stiffener separation failure mode is modeled by placing a layer of cohesive elements COH3D8 between the skin and the stiffener along the flanges of the stiffener. These elements can represent mixed-mode delamination growth with a damage evolution law specified as a traction-separation law [16]. Regarding intralaminar damage it is considered over the model, studying both the initiation and the evolution. The damage activation is based on Hashin failure criteria [17]. According to the failure criteria, there are four different failure modes investigated in the finite element analysis: fiber tension, fiber compression, matrix tension and matrix compression.

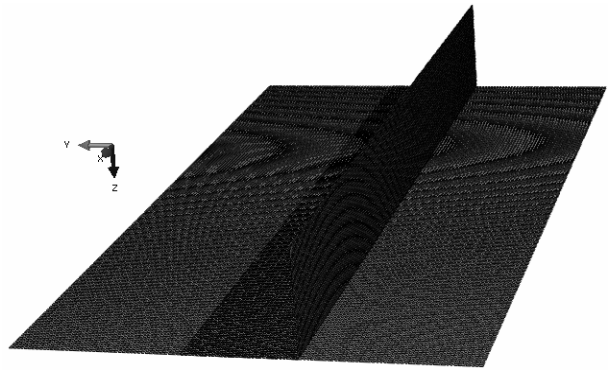


Fig. 7: Finite element mesh view

3.2 Dynamic analysis in Abaqus/Standard

The analysis is performed as a dynamic analysis using general direct-integration method provided in Abaqus/Standard, called the Hilber-Hughes-Taylor operator [15]. In the analysis, it is shown that the convergence difficulties can be avoided by introducing a small fictitious viscosity with the value of 1.0e-4 in the cohesive zone law that characterizes the interface. All the analyses are based on an implicit dynamic solution.

A careful characterization of the material properties [18] is important for the analysis. The interlaminar properties and ply fracture toughness used in the damage model are shown in Tables 3-5.

Table 3: Ply strength - IM7/8552 (MPa)

X^T	X^C	Y^T	Y^C	S^L	S^T
2323	1200	101.4	199.8	107	75.3

Table 4: Ply fracture toughness - IM7/8552 (N/mm)

Gft	Gfc	Gmt	Gmc
81.5	106.3	0.277	0.788

Table 5: Interlaminar properties - IM7/8552

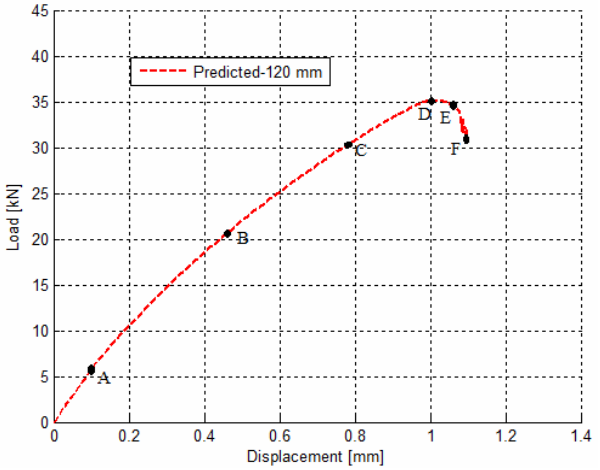
$G_{IC}(N/mm)$	$G_{IIC}(N/mm)$	$S_{zz}(MPa)$	$S_{xz}(MPa)$	$B-K$
0.277	0.788	40	80	1.6

3.3 Out-of-plane response

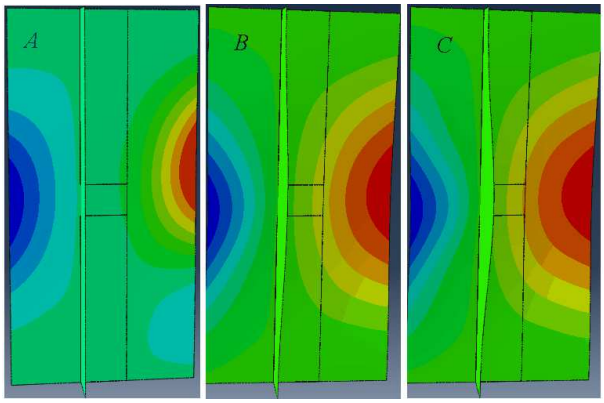
The load-shortening curve and the out-of-plane response are illustrated in Fig. (8). Point A corresponds to the buckling load of the skin and point B results from the buckling of the

stiffener. Point C corresponds to the displacement of 0.76 mm where the eighth ply has matrix tensile damage initiation. Point D indicates the separation between the skin and the stiffener. Point E shows two single half-waves existing on one of the free edges, which makes the separation between the skin and the stiffener larger. The analysis is not able to predict the full separation of the skin and the stiffener because of the convergence difficulties.

a)



b)



c)

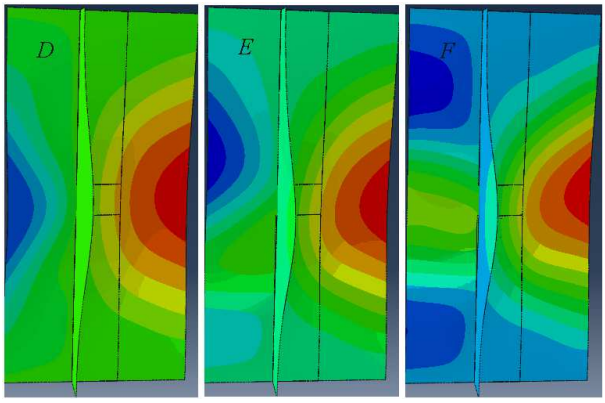
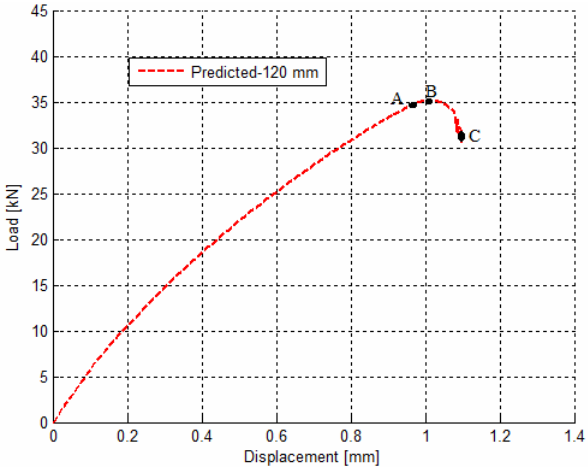


Fig. 8: Analysis results: a) load-shortening curve; b) c) out-of-plane response.

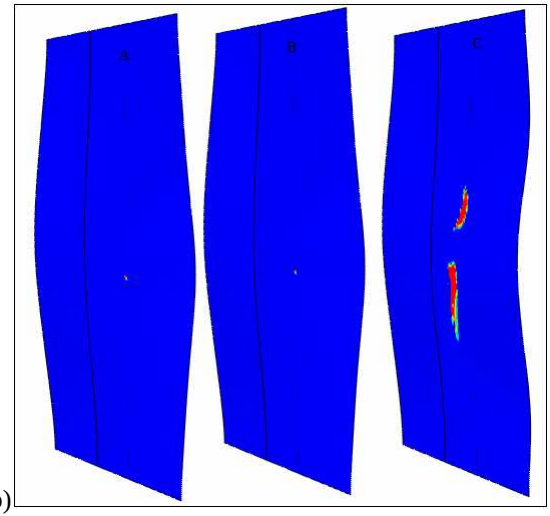
3.4 Intralaminar damage analysis

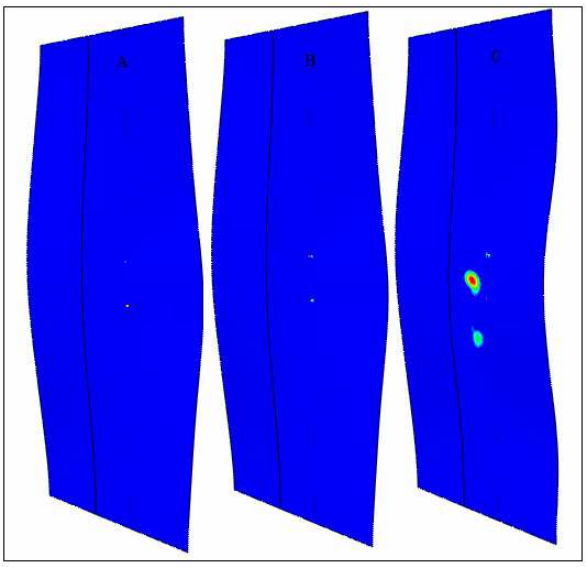
The location of the damage variables related to fiber and matrix in compression at three load points are illustrated in Fig. (9). The matrix tensile damage in the eighth ply is complete at a load of 34.8 kN, while the fiber compressive damage just initiates at the point A. At the peak load represented by point B, the skin-stiffener starts to separate. At point C, both the matrix damage and the fiber damage extend to a larger area along the interface.

a)



b)



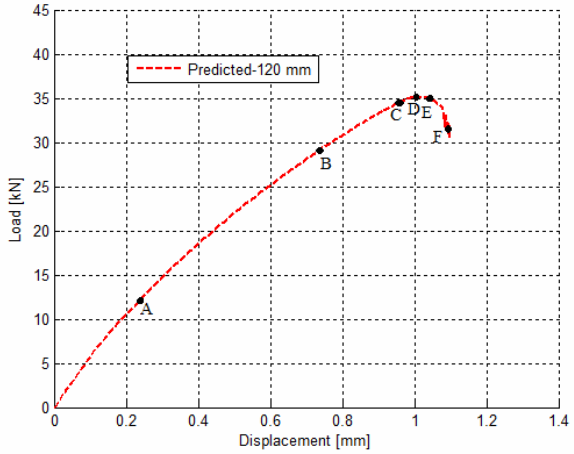


c)

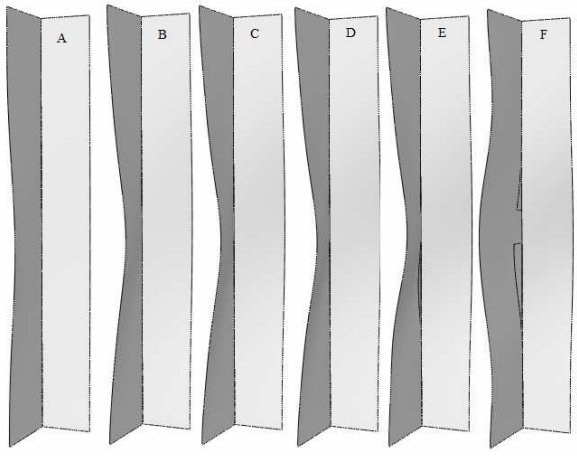
Fig. 9: Intralaminar analysis: a) load-shortening curve; b) matrix tensile damage area; c) fiber compressive damage area.

3.5 Interlaminar analysis using cohesive elements

The delamination between the skin and the stiffener is investigated by introducing COH3D8 elements into the interface between the skin and the stiffener foot. The cohesive elements are not placed in the middle of the specimen stiffener where there is the Teflon insert. The load-shortening curve and the delamination between the skin and stiffener are shown in Fig. (10). Point A corresponds to the initiation of the damage on the cohesive elements, while the delamination onset corresponds to the peak load given by point D, and its propagation occurs during the collapse of the specimen.



a)



b)

Fig. 10: Analysis of delamination: a) load-shortening curve; b) skin-stiffener separation.

4 Test methodology

The tests methodology has been decided on the basis of the numerical analyses.

4.1 Fatigue of composite materials

In compression-loaded L-shaped composite structures, delaminations are likely to be the most critical type of fatigue damage. In any case, since the specimens are laminated with multiple plies, the fracture behavior will be characterized by interacted damage modes.

Many experimental fatigue tests have been carried out to study the fatigue crack growth in composites. One of the most popular fatigue damage models is Paris law model [11, 12], as shown in Eq. (1), which has been used to describe the fatigue crack propagation behavior:

$$\frac{da}{dN} = C\left(\frac{\Delta G}{G_c}\right)^r \qquad (1)$$

In this model, delamination growth rate $\frac{da}{dN}$ has been related to the cyclic strain energy release rate $\frac{\Delta G}{G_c}$ using a power law, where C

and r are material constants. For composites, the exponent r, which relates propagation rate to strain energy release rate, has shown high values [19]. With a large exponent, small scatter in the applied loads would result in a large scatter in the predicted delamination growth rate. For this reason, the derived power law is not reliable enough for design criteria. Hence, it is important to ensure that the obtained strain energy release rate threshold value G_c corresponds to no delamination growth by experiments. Through the tests, two constants C and r in the Paris law, together with the threshold value G_c could be obtained and may be useful for developing cyclic delamination characterization.

It is so important to introduce the expression for compliance used to calculate the strain energy release rate. If the specimen is subjected to a load P when the crack has length a, the specimen is less compliant than when the crack has length $a+\delta a$.

The compliance C_p of the L-shaped specimen is the displacement per unit load and it can be obtained from load-shortening curve. In general it is written

$$C_p = \frac{\delta}{P} \qquad (2)$$

Where δ is the displacement of the specimen at the point of load application. The compliance is geometry constant, dependent on crack length and dimensions of the body. Through the experimental data, it could be possible to plot values of compliance C_p versus crack length a. From the curve, the delamination strain energy release rate can be expressed as [20]

$$G = \frac{P^2}{2b}\frac{dC_p}{da} \qquad (3)$$

where b is the Teflon insert width in the interface between the skin and the stiffener.

4.2 Test technique and procedures

The specimen edges are coated with water-based, brittle typewriter correction fluid, and marks are made at 1 mm intervals from the initial delamination tip. All fatigue tests will be conducted in an MTS servo-hydraulic test machine at a frequency of 5 Hz and an R (load ratio) value of 0.1.

For the initiation of the delamination, Teflon insert is used in the tests to provide a straight delamination front. To make the delamination more visible during the tests, light source and high-resolution camera are used to enhance observation of the delamination growth. Crack tip positions are read at intervals, temporarily stopping the load cycling.

According to the previous finite element analysis under static loading, the buckling load of the specimen is equal to 6.3 kN, while the collapse load is equal to 35.1 kN. Three sequences are considered, always with maximal and minimal loads according to the load ratio equal to 0.1. Firstly, the cycling is performed from 2 kN to 20 kN. During the cycling, it is important to check the difference of structural response between static and fatigue loading. At the second stage, a larger maximal load equal to 28 kN together with the corresponding minimal load of 2.8 kN is introduced during the cycling. With the increasing maximal load, the number of cycles performed during the test decreases. Finally, the cycling is performed from 1.2 kN to 12 kN.

4.3 Proposed measurements

Different measurements are taken to calculate the compliance, the fatigue threshold value and the strain energy release rate. In particular, the following data are expected from the tests:

1) Load-shortening curve that will be used to obtain the compliance C_p.
2) Number of cycles where delamination onset appears, so to be able to measure the strain energy release rate threshold, G_c.
3) Crack length a at different numbers of cycles.
4) Delamination growth per fatigue cycle, $\frac{da}{dN}$, related to the strain energy release rate G.

4.4 Expected results

The proposed experimental tests are used to characterize the cyclic delamination in the stiffened specimens. There are some results that are expected from the tests:

1) Two crack tips will probably develop because of the centered Teflon insert.
2) Delamination growth would proceed in a slow stable extension way instead of unstable jumps.
3) Compliance of the specimen should always increases as the crack grows, in order to derive the compliance C as a function of crack length.
4) Maximal strain energy release rate would decrease with the increased crack length. The energy release rate could be related to compliance C_p by using $G = \frac{p^2}{2b}\frac{dC_p}{da}$.
5) Crack growth rate $\frac{da}{dN}$ will be drawn with respect to the strain energy release rate in a power law.
6) Fatigue life constants C and r will be determined as shown in Paris law $\frac{da}{dN} = C(\frac{\Delta G}{G_c})^r$.

4.5 Correlation between the static loading and fatigue loading

It will be useful to compare the structural response of the L-shaped specimens under both static and fatigue loading. Moreover, the fatigue tests could provide some design criteria for similar structures.

5 Conclusions

Single stiffener specimens are proposed to investigate experimentally the fatigue behavior of composite stiffened structures that work in post-buckling conditions. Numerical analysis under static loading have been performed and used to provide some guidance for the cycling load range, such as the buckling load, the collapse load and the pre-buckling stiffness. In addition, the finite element analyses are important to investigate both intralaminar and interlaminar damage mechanisms. The cohesive zone modeling approach has been used to characterize the interface behavior between the skin and the stiffener.

The application of elastic fracture mechanics criteria has been proposed as a possible tool for predicting delamination growth under compression fatigue loading. A power law relating the delamination growth to the cyclic strain energy release rate is investigated.

References

[1] Revuelta D., Miravete A., “Fatigue damage in composite materials”, *International Applied Mechanics*, Vol. 38, No. 2, 2002, pp. 121-134.

[2] Turon A., Costa J., Camanho P. P., Davila C. G., “Simulation of delamination in composites under high-cycle fatigue”, *Composites Part A: applied science and manufacturing*, Vol. 38, 2007, pp. 2270-2282.

[3] Shokrieh M. M., “Progressive fatigue damage modeling of composite materials, part II: material characterization and model verification”,

Journal of Composite Materials, Vol. 34, No. 13, 2000, pp. 1841-1852.

[4] Melin L. G., Schon J., "Buckling behavior and delamination growth in impacted composite specimens under fatigue load: an experimental study", *Composites Science and Technology*, Vol. 61, 2001, pp. 1841-1852.

[5] Cordisco P., Bisagni C., "Cyclic buckling tests under combined compression and shear on composite stiffened panels", *AIAA Journal*, Vol. 47, No. 12, 2009, pp. 2879-2893.

[6] Paepegem V. W., Degrieck J., Baets P. D., "Finite element approach for modeling fatigue damage in fibre-reinforced composite materials", *Composites Part B*, Vol. 32, No. 7, 2001, pp. 575-588.

[7] Bisagni C., "Progressive delamination analysis of stiffened composite panels in post-buckling", *Proceedings of the 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Material Conference*, Newport, Rhode Island, USA, May 1-4, 2006, AIAA 2006-2178.

[8] Kan H. P., Whitehead R.S., Kautz E., Damage tolerance certification methodology for composite structures. *8th DOD/NASA/FAA Conference on Fibrous Composites in Structural Design*, NASA Langley Research Center, 1990, pp. 479-498.

[9] Degrieck J., Paepegem W. V., "Fatigue damage modeling of fibre-reinforced composite materials: Review", *Applied Mechanics Reviews*, Vol. 54, No. 4, 2001, pp. 279-300.

[10] Reifsnider K. L., Fatigue of composite materials, Elsevier, Amsterdam, 1991.

[11] Paris P., Gomez M., Anderson W., "A rational analytical theory of fatigue", *Trend in Engineering*, Vol. 13, 1961, pp. 9-14.

[12] Paris P., Erdogan F., "Critical analysis of propagation laws", *Journal of Basic Engineering*, Vol. 85, 1963, pp. 528-534.

[13] Lemaitre J., Sermage J., Desmorat R., "A two scale damage concept applied to fatigue", *International Journal of Fracture*, Vol. 97, 1999, pp.67-81.

[14] Composite Materials Handbook, "*Polymer matrix composite materials usage, design, and analysis*", Department of Defense, USA, Vol. 3, 2002, Chap. 4.

[15] ABAQUS Manuals, Version 6.9.1., Dassault Systemes, Providence, RI, USA, 2008.

[16] Camanho P. P., Davila C. G., Moura M. D., "Numerical simulation of mixed-mode progressive delamination in composite materials", *Journal of Composite Materials*, Vol. 37, 2003, pp.1415-1438.

[17] Hashin Z., "Failure criteria for unidirectional fiber composites", *Journal of Applied Mechanics*, Vol. 47, 1980, pp. 329–334.

[18] Camanho P. P., Maimi P., Davila C. G., "Prediction of size effects in notched laminates using continuum damage mechanics", *Composites Science and Technology*, Vol. 67, 2007, pp. 2715-2727.

[19] Mall S., Yun K. T., Kochhar N. K., "Characterization of matrix toughness effect on cyclic delamination growth in graphite fiber composites", In ASTM STP 1012, American Society for testing and materials, Philadelphia, PA, 1989, pp. 296-310.

[20] Anderson T. L., "Fracture mechanics: fundamentals and applications", Taylor & Francis, 2005, Chap. 1.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A three-dimensional damage model to simulate the progressive failure of composite structures

D. Fanteria and G. Longo
Department of Aerospace Engineering, University of Pisa, Italy

Keywords: *Composite structures, low-velocity impacts, continuum damage mechanics, cohesive zone model, implicit finite element code.*

Abstract

A three-dimensional damage model has been developed to simulate the progressive failure of thin composite structures. Both intralaminar and interlaminar damage mechanisms have been considered and the proposed constitutive model has been developed on the basis of thermomechanics of the nonlinear irreversible physical process. The intralaminar damage modes has been analysed in the context of the Continuum Damage Mechanics, whereas the interlaminar damage mode has been analysed by means of the Cohesive Zone Model. The objectivity of the numerical discretization has been assured using the smeared crack formulation. The proposed damage model has been benchmarked by means of an experimental low-velocity impact test taken from the Literature, reproducing the main damage mechanisms, i.e. fibre breakage, fibre kinking, matrix cracking, matrix crushing and delaminations.

1 Introduction

Since decades, in aerospace industry fibre-reinforced polymers have become attractive because of their light-weight, high-strength and high-modulus properties. However, composite structures are still more expensive than their metal counterpart due to the high costs of production, quality assurance and experimental testing. Furthermore, in the context of the damage tolerance design philosophy, low-velocity impacts represent an issue of great concern because they can induce serious damages which may propagate up to the collapse of the structure.

This paper presents a three-dimensional energy-based damage model for composite structures subjected to low-velocity impacts. The intralaminar damage model has been developed, on the basis of thermomechanics of the nonlinear irreversible physical processes, in the context of the Continuum Damage Mechanics (CDM). Moreover, the interlaminar damage mechanism has been analysed by means of the Cohesive Zone Model (CZM) and the delamination propagation under mixed-mode loading has been modelled using special-purpose cohesive elements.

The proposed three-dimensional damage model has been exploited by means of the implicit nonlinear finite element code Abaqus/Standard using a direct-integration dynamic analysis, which has been preferred to the commonly used explicit central-difference

time integration rule. The final goal is to enable the prediction of low-velocity impact damages in thin carbon/epoxy composite laminates, considering the extension of the damaged areas in terms of fibre breakage, fibre kinking, matrix cracking, matrix crushing and delaminations.

2 Progressive failure analysis of composite structures

The World Wide Failure Exercise (WWFE) [1,2] in the late 90's provided a good assessment of the status of the most used theoretical methods for predicting material failure in fibre reinforced polymer composites. It was clearly emphasized that, even for simple laminates, predictions may differ significantly from the experimental results.

The current intralaminar constitutive damage models consider both damage onset and damage propagation, assuming that the response of a single lamina is linearly elastic up to failure, with the exception of in-plane and out-of-plane shear which exhibits typically a nonlinear behaviour [3]. For damage onset, criteria can be either non-interactive, e.g. maximum stress, maximum strain, Tsai-Hill, Tsai-Wu, Hoffmann criteria, or interactive which distinguish between the different damage modes, e.g. Hashin's criterion, Puck's criterion, LaRC03 criterion for 2D models [4] and LaRC04 criterion for 3D models [5].

Damage progression strategies that provide material degradation can be usually categorized into two main groups: heuristic models based on a ply-discounting material degradation approach [6] and models based on the CDM approach using internal state variables. The former models are characterized by the degradation of the elastic stiffness coefficients by a value β which diminishing the elastic moduli directly after damage initiation is detected. The degradation can be instantaneous (β ≈ 0) or recursive, mitigating the convergence problems.

Material softening can be more rigorously analysed by means of the CDM. The material damage state is described using appropriate internal state variables (the *damage variables*) which directly modify the entries of the constitutive matrix. Some methodologies have been proposed in order to solve the mesh dependency concern of the progressive failure analyses. Bažant [7] has proposed the smeared crack formulation introducing the element's characteristic length when evaluating the strain at failure by means of the energy dissipated per unit of volume.

As far as the interlaminar damage modes are concerned, many fracture methods are available in the Literature. Elder et al. [8] presents an interesting review of numerical methods and among which the Linear Elastic Fracture Mechanics (LEFM) and the CZM are the most notable. The LEFM method is widely used by means of the Virtual Crack Closure Technique. The CZM uses special-purpose elements with zero or quasi-zero thickness in order to simulate inter-ply delamination onset and propagation. The commonly used traction-separation approach assumes a linear elastic behaviour at small loads, followed by a linear material softening after reaching the interlaminar allowable strengths. The mixed-mode delamination propagation is then simulated by means of several approaches, such as the Benzeggagh-Kenane law or the Power law.

3 The intralaminar damage model

The intralaminar three-dimensional damage model is based on the definition of a damaged stiffness tensor Q_d which relates Cauchy stress tensor $\boldsymbol{T}$ and the strain tensor $\boldsymbol{E}$:

$$\boldsymbol{T} = \mathrm{Q}_d \boldsymbol{E} . \qquad (1)$$

The damaged stiffness tensor can be defined by means of the damage tensor D, as follows:

$$\mathrm{Q}_d = \mathrm{D}^{-1}\mathrm{Q}, \qquad (2)$$

where the damage tensor is defined as:

$$[\mathrm{D}]_{ij} = \frac{1}{1-d_i} \cdot \delta_{ij}, \qquad (3)$$

with d_i being the damage variables and δ_{ij} indicating the Kronecker delta. The damage variables are defined in the range $[0,1]$, that is

$d_i = 0$ represents the initial undamaged material, whereas $d_i = 1$ represents the complete loss of load-carrying capability. The variable d_1 describes the in-plane longitudinal damage mode, the variable d_2 describes the in-plane transverse damage mode, whereas the variable d_4 describes the in-plane shear damage mode. Thus, a plane stress behaviour is considered, which is a good assumption for thin laminae. No damage has been considered for the through-the-thickness direction and for the out-of-plane components of shear, $d_3 = d_5 = d_6 = 0$. Moreover, the damage variables d_1 and d_2 take into account the sign of the strain/stress state. Considering as example the longitudinal damage variable d_1, other two damage variables are introduced in order to consider the tensile and compressive effects, that is d_{1T} and d_{1C}. This permits also to consider the stiffness recovery due to the closure of pre-existing cracks under load reversal:

$$\begin{cases} d_1 = \max\left(d_{1T} \dfrac{\langle \varepsilon_{11} \rangle}{|\varepsilon_{11}|}, d_{1C} \right) \\ d_2 = \max\left(d_{2T} \dfrac{\langle \tilde{\sigma}_{22} \rangle}{|\tilde{\sigma}_{22}|}, d_{2C} \right) \end{cases}, \qquad (4)$$

where $\langle \bullet \rangle$ is the Macaulay brackets operator. The damage variables are activated by a damage initiation criterion defined through a damage activation function Φ_i. Damage is activated whenever the function exceeds one, then the material begins to degrade by means of a suitable softening law.

3.1 Fibre Failure (FF) due to tensile loading

Usually, fibre fracture marks ultimate failure. In carbon fibre-reinforced polymers shear stresses have negligible influence on the tensile Fibre Fracture limit. In this work, the damage initiation due to the tensile loading is predicted using a non-interacting strain-based failure criterion, i.e. the longitudinal tensile damage activation function is given by:

$$\Phi_{1T} = \left(\frac{E_1}{X_T} \varepsilon_{11} \right)^2, \quad \text{for } \varepsilon_{11} \geq 0, \qquad (5)$$

where E_1 is the longitudinal elastic modulus and X_T is the longitudinal tensile allowable strength. The material softening is defined by a linear law in terms of a suitable driving strain component. For damage variable d_{1T} the softening law is expressed by the following formula:

$$d_{1T} = \max_t \left\{ 0, \min_t \left(1, \frac{\hat{\varepsilon}_1^f}{\hat{\varepsilon}_1^f - \hat{\varepsilon}_1^0} \left(1 - \frac{\hat{\varepsilon}_1^0}{\hat{\varepsilon}_1} \right) \right) \right\}, \qquad (6)$$

with the driving strain component defined as $\hat{\varepsilon}_1 = \varepsilon_{11}$. Fig. 1 shows the stress-strain relation characterized by a linear softening. $\hat{\varepsilon}^0$ represents the driving strain at damage onset, whereas $\hat{\varepsilon}^f$ is the driving strain at complete failure.

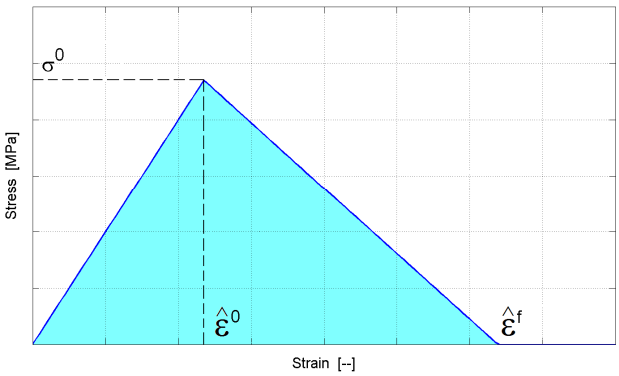


Fig. 1 Stress-strain curve.

3.2 Fibre Failure (FF) due to compressive loading

The fibre failure due to compressive loading is characterized by kinking of the fibres which start to buckle and then reach final failure. In this work a strain-based damage initiation criterion is chosen, whose damage activation function has the same form of the tensile damage case:

$$\Phi_{1C} = \left(\frac{E_1}{X_C} \varepsilon_{11} \right)^2, \quad \text{for } \varepsilon_{11} < 0, \qquad (7)$$

where X_C is the longitudinal compressive allowable strength. Similarly to the longitudinal tensile case, the material softening is defined by a linear law as follows:

$$d_{1C} = \max_t \left\{ 0, \min \left(1, \frac{\hat{\varepsilon}_1^f}{\hat{\varepsilon}_1^f - \hat{\varepsilon}_1^0} \left(1 - \frac{\hat{\varepsilon}_1^0}{\hat{\varepsilon}_1} \right) \right) \right\}, \qquad (8)$$

with the driving strain component defined as $\hat{\varepsilon}_1 = |\varepsilon_{11}|$.

3.3 Inter Fibre Failure (IFF) due to tensile loading

The effects of transverse tension loads, such as the transverse compression loads, are taken principally by the matrix. In the case of transverse tension loads, damage occurs as a through the layer crack runs parallel to the fibres, which may be initiated by a stress concentration in the matrix or a local matrix-fibre debonding. In this work, a transverse tensile damage activation function is used in the following form:

$$\Phi_{2T} = \left(\frac{\tilde{\sigma}_{22}}{Y_T} \right)^2 + \left(\frac{\tilde{\tau}_{12}}{S_{12}} \right)^2 + \left(\frac{\tilde{\tau}_{23}}{S_{23}} \right)^2, \quad \text{for } \tilde{\sigma}_{22} \geq 0, \qquad (9)$$

where Y_T is the transverse tensile allowable strength, whereas S_{12} and S_{23} are the in-plane shear strength and transverse shear strength. Following the bilinear damage formula written in paragraphs 3.1 and 3.2, the material softening is defined by a linear law as follows:

$$d_{2T} = \max_t \left\{ 0, \min \left(1, \frac{\hat{\varepsilon}_2^f}{\hat{\varepsilon}_2^f - \hat{\varepsilon}_2^0} \left(1 - \frac{\hat{\varepsilon}_2^0}{\hat{\varepsilon}_2} \right) \right) \right\}, \qquad (10)$$

with the driving strain component defined as $\hat{\varepsilon}_2 = \sqrt{\langle \varepsilon_{22} \rangle^2 + \gamma_{12}^2 + \gamma_{23}^2}$.

3.4 Inter Fibre Failure (IFF) due to compressive loading

Generally, failure in transverse compression is due to the crushing of the matrix. Experimental evidence suggests that the fracture is created by the stresses which act on a typical fracture plane inclined of $\theta_{fp} = 53°$ with respect to the through-the-thickness direction [3]. These stresses are the normal stress σ_n, and two shear stresses τ_{nl} and τ_{nt} as shown in Fig. 2. σ_n represents the stress perpendicular to the fracture plane, τ_{nl} is the transverse-longitudinal shear stress and τ_{nt} is the transverse-transverse shear stress.

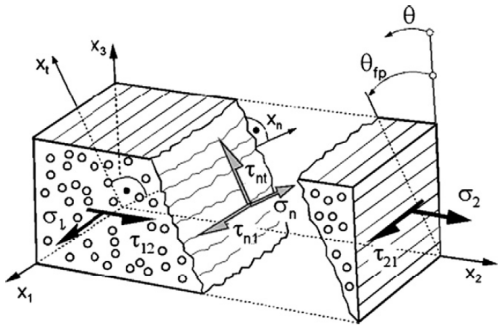


Fig. 2 Fracture plane for Inter Fibre Fracture (IFF) due to compressive loading, [3].

Given the effective stress tensor $\tilde{\boldsymbol{T}}$ in the lamina coordinate system, using a transformation matrix the stresses acting on the inclined fracture plane are obtained:

$$\begin{cases} \sigma_n = \tilde{\sigma}_{22} \cos^2 \theta_{fp} + \tilde{\sigma}_{33} \sin^2 \theta_{fp} + 2\tilde{\tau}_{23} \cos \theta_{fp} \sin \theta_{fp} \\ \tau_{nl} = \tilde{\tau}_{12} \cos \theta_{fp} + \tilde{\tau}_{13} \sin \theta_{fp} \\ \tau_{nt} = (\tilde{\sigma}_{33} - \tilde{\sigma}_{22}) \cos \theta_{fp} \sin \theta_{fp} + \tilde{\tau}_{23} (\cos^2 \theta_{fp} - \sin^2 \theta_{fp}) \end{cases} \qquad (11)$$

Considering that a tensile stress $\sigma_n \geq 0$ causes crack opening and propagation whereas a compressive stress $\sigma_n < 0$ impedes shear fractures [3], for $\sigma_n < 0$ the shear stresses have to cause fracture against an additional resistance, which increases with increasing $|\sigma_n|$ like an internal friction. Following these consideration, a stress-based damage activation function is used:

$$\Phi_{2C}=\left(\frac{\tau_{nt}}{Y_C-\mu_T\sigma_n}\right)^2+\left(\frac{\tau_{nl}}{Y_C/\left(2\tan\theta_{fp}\right)-\mu_L\sigma_n}\right)^2 \quad (12)$$
$$\text{for } \tilde{\sigma}_{22}<0 \text{ and } \sigma_n<0,$$

where Y_C is the transverse compressive allowable strength and the friction coefficients are given by the following equations:

$$\mu_T=\frac{1}{\tan\left(2\theta_{fp}\right)},\ \mu_L=\frac{\mu_T}{2Y_C\tan\theta_{fp}}. \quad (13)$$

When Φ_{2C} reaches a value equal to unity, then the damage begins to grow and the material begins to degrade and the damage variable d_{2C} is expressed by the following formula:

$$d_{2C}=\max_t\left\{0,\min_t\left(1,\frac{\hat{\varepsilon}_2^f}{\hat{\varepsilon}_2^f-\hat{\varepsilon}_2^0}\left(1-\frac{\hat{\varepsilon}_2^0}{\hat{\varepsilon}_2}\right)\right)\right\}. \quad (14)$$

The driving strain component which defines the damage evolution law is given by $\hat{\varepsilon}_2=\sqrt{\gamma_{nl}^2+\gamma_{nt}^2}$. Thus, only the shear strain components in the fracture plane are considered for the transverse compressive damage mode. The shear strain components can be evaluated as:

$$\begin{cases}\gamma_{nl}=\gamma_{12}\cos\theta_{fp}+\gamma_{13}\sin\theta_{fp}\\ \gamma_{nt}=\left(\varepsilon_{33}-\varepsilon_{22}\right)\cos\theta_{fp}\sin\theta_{fp}+\gamma_{23}\left(\cos^2\theta_{fp}-\sin^2\theta_{fp}\right)\end{cases} \quad (15)$$

3.5 In-plane shear failure

In this work, the in-plane shear damage mechanism has been considered as due to damage mechanisms in longitudinal and transverse directions. This means that shear damage is activated by fibre breakages and matrix cracking in the following manner:

$$d_4=1-\left(1-d_1\right)\left(1-d_2\right). \quad (16)$$

3.6 Smeared crack formulation

Progressive damage in composites materials is known to be affected by the strain localization phenomenon which occurs whenever failure is preceded by the emergence of narrow and highly strained zones [9].

Strain localization during the fracture process typically causes FEM based progressive failure analyses incorporating material degradation into the constitutive model to suffer from mesh dependency.

A smeared crack formulation, whose theoretical background can be found in the Literature [7,9], can be used to avoid mesh dependency. Using this formulation, the fracture energy is distributed (smeared) over the full volume of the element. A length parameter (characteristic length), related to the element dimension, is introduced into the constitutive law in order to achieve a constant energy release per unit area of crack, regardless of the element dimension. With the introduction of the smeared crack formulation, the strain $\hat{\varepsilon}^f$ is properly scaled on the basis of the characteristic length of the finite element l, so that the area under the stress-strain curve (see Fig. 1) is equal to the energy dissipated per unit of volume g, that is:

$$g=\frac{G}{l}, \quad (17)$$

where G is the intralaminar fracture toughness of the material. Thus, the less the characteristic length, the bigger the strain at complete failure $\hat{\varepsilon}^f$.

The smeared crack formulation, albeit being quite crude, is the simplest manner to solve the mesh size dependency problem; with such approach the damage localizes into a zone which is one element in height thus the best results are obtained when the damage location is known thus permitting to design mesh accordingly [10]. When a well-defined damage path is not defined a non-local approach is generally recommended [10], nevertheless the smeared crack formulation can still be used if a certain level of approximation is tolerated.

4 The interlaminar damage model

In order to simulate the onset and propagation of delaminations, three-dimensional special-purpose cohesive elements, formulated

according to the CZM, are used at interply interfaces. A *traction-separation* approach has been considered combining a strength-based criteria to predict the delamination onset with fracture mechanics to evaluate the subsequent propagation. The element is characterized by a very small thickness (h), representing the resin rich ply-to ply interfaces, and has four couples of matched nodes connected by nonlinear springs used to simulate the three different damage modes, that is (see Fig. 3) the opening mode (mode I), the sliding mode (mode II) and the tearing mode (mode III).

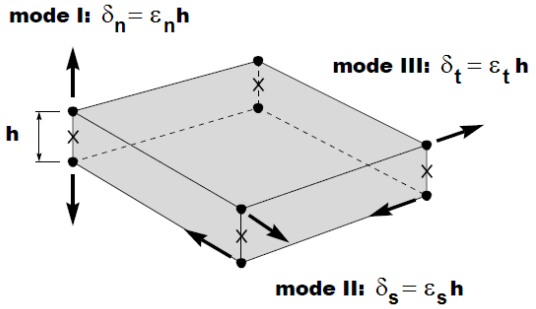


Fig. 3 A cohesive element with four couples of matched nodes. Image taken from [11].

The element constitutive equation relates the normal and shear stress components with the relevant nodal separations by means of a penalty stiffness matrix $[\mathrm{K}]$:

$$\{t_n, t_s, t_t\}^T = [\mathrm{K}]\{\delta_n, \delta_s, \delta_t\}^T. \qquad (18)$$

The nodal separations are obtained multiplying the strain components by the thickness of the cohesive element: $\{\delta_n, \delta_s, \delta_t\} = \{\varepsilon_n h, \varepsilon_s h, \varepsilon_t h\}$. The choice of the correct value of the penalty stiffness is an issue of great concern: a high initial penalty stiffness is necessary to hold the top and bottom faces together in the linear elastic range even though a very large penalty value may result in an ill-conditioned problem. Usually, a value around $K = 10^6$ N/mm^3 is used [12] in combination with thicknesses between 10^{-4} and 10^{-3} mm.

Delamination onset is linked to a quadratic interaction function involving the stress ratios:

$$\varphi = \left(\frac{\langle t_n \rangle}{T_n^0}\right)^2 + \left(\frac{t_s}{T_s^0}\right)^2 + \left(\frac{t_t}{T_t^0}\right)^2, \qquad (19)$$

where T_n^0, T_s^0 and T_t^0 represent material allowables, that is the peak value of the stress when the deformation is purely normal to the interface or purely in the first or the second shear direction, respectively. Delamination initiates when φ reaches a value of one; after delamination onset, the stiffness of the cohesive element is gradually reduced to zero depending on the value of a damage parameter D [13], which ranges from zero at onset, to one at the complete debonding between the two adjacent laminae.

When mixed-mode propagation is analysed, generally a Benzeggagh-Kenane (B-K) law or a Power law are the most used [13]. The B-K law is particularly useful when the critical fracture energies during deformation purely along the first and the second shear directions are the same, i.e. $G_{IIC} = G_{IIIC}$:

$$G_{IC} + (G_{IIC} - G_{IC})\left(\frac{G_{Sh}}{G_n + G_{Sh}}\right)^{\eta} = G_n + G_{Sh}, \qquad (20)$$

where G_n, G_s and G_t refer to the work done by the traction components and their conjugate relative displacements δ_i, $G_{Sh} = G_s + G_t$ and η is a parameter obtained experimentally.

5 Damage model implementation and verification

The three dimensional damage model described in sections 3 and 4 has been implemented into the Abaqus/Standard implicit code programming the intralaminar damage model in an User-defined MATerial (UMAT) subroutine with the aim to simulate damage propagation in layered composite material caused by low energy impacts.

Notwithstanding the fact explicit codes are the most widely used for low-velocity and ballistic impact simulations, the Hilber-Hughes-Taylor time integration scheme available within Abaqus/Standard has been preferred for its

unconditional stability and for its major control of convergence. The main disadvantage is the computational cost due to the fact that the stiffness matrix must be inverted and a set of simultaneous nonlinear dynamic equilibrium equations must be iteratively solved at each time increment.

Due to the aforementioned mesh sensitivity problem of the proposed intralaminar damage model, a sensitivity analysis has been preliminarily performed in order to verify the correct implementation of the smeared crack formulation within the UMAT subroutine.

A cube with a side of 1 mm has been modelled and loaded in tension in the longitudinal (fiber) direction. Three different mesh densities have been considered consisting of 1 element (l = 1 mm), 3^3 elements (l = 0.333 mm) and 5^3 elements (l = 0.2 mm). For the last two cases, the damage path has introduced only for the elements in a band in the middle of the cube perpendicular to the loading direction. Fig. 4 shows the load-displacement curve of the three simulation showing that the energy absorbed by the cube is correctly independent of the mesh refinement.

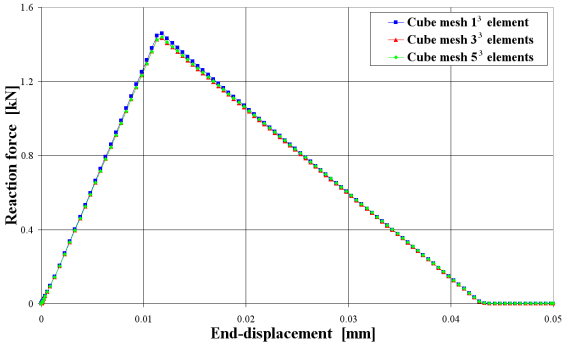


Fig. 4 Load-displacement curve for three different mesh densities.

6 Finite element simulation of a low-velocity impact

The proposed damage model has been subjected to a first verification and validation process by simulating low-velocity impact test case taken from the Literature [14]. The low-velocity impact test is realized following the specification described on the ASTM D5628 [15] on a composite specimen with a 1.8 mm thickness and a $[0/90/0/90/0]_S$ stacking sequence. The specimen is hold between two circular steel disks with a hole of 80 mm diameter under a drop tower machine with an aluminium impactor with a hemispherical cap of 8 mm radius. The impactor, its support and the load cell have a total mass of 1.205 kg. The impactor falls from a height of 0.5 m, obtaining an impact velocity of 3.13 m/s and an impact energy of 5.91 J.

6.1 Modelling strategy

In order to alleviate the computational costs, a circular portion of the specimen, with a diameter of 100 mm has been modelled in a clamped set-up; model dimensions have been further reduced by to one fourth by exploiting material and load symmetries of the problem.

The impactor has been modelled by means of a hemispherical analytical rigid surface and a node with a concentrated mass point has been associated to it. The boundary conditions are enforced in this point allowing only the motion of the impactor along the falling direction (z-axis). Full three-dimensional models have been developed employing eight nodes reduced integration solid elements (C3D8R) with interlaminar cohesive elements (COH3D8) placed between each couple of adjacent laminae. The quadratic nominal stress criterion has been used to determine the delamination onset, whereas the Benzeggagh-Kenane criterion (η=1.45) has been assumed in order to analyse the delamination propagation. Lamina material properties are reported in Table 1; intralaminar fracture toughness were not available and so values for similar materials have been chosen from the Literature. Interlaminar allowables and fracture toughness are reported in Table 2.

A surface-to-surface contact discretization with a small-sliding tracking approach has been used to simulate the contact between the impactor and the thin plate because it is more efficient and recommended for normal impactor direction [16]. The penalty method has been used as contact constraint enforcement method. The contact between adjacent layers has been

simulated in a similar manner thus accounting for contact nonlinearities [13].

Table 1 Lamina material properties.

Elastic Properties		Allowable strengths [MPa]		Fracture toughness [kJ/m^2]	
E_1	127.0GPa	X_T	1400	G_{1T}	80
$E_2 = E_3$	10.0GPa	X_C	930	G_{1C}	60
$G_{12} = G_{13}$	5.4GPa	Y_T	47	G_{2T}	0.2
G_{23}	3.0GPa	Y_C	130	G_{2C}	0.7
$\nu_{12} = \nu_{13}$	0.34	$S_{12} = S_{13}$	53		
ν_{23}	0.306	S_{23}	89		

Table 2 Interlaminar allowables and fracture toughness.

Allowable strengths [MPa]		Fracture toughness [kJ/m^2]	
T_n^0	62.3	G_{IC}	0.2
T_s^0	53	G_{IIC}	0.9
T_t^0	89	G_{IIIC}	0.9

The introduction of artificial damping has been necessary in order to control the problem convergence: a damping control parameter of $\alpha = -0.414$ has been used. This value guarantees an improvement of the convergence behaviour without significantly degrading the solution accuracy [13].

6.2 Numerical results

The mesh density is an issue of great concern so a preliminar mesh sensitivity analysis, considering an undamaged material, has been carried out. A temporal window of 1 ms and four different mesh densities in the contact zone have been considered: a coarse mesh (l = 2 mm), an intermediate mesh (l = 1 mm), a fine mesh (l = 0.5 mm) and a very fine mesh (l = 0.25 mm). Fig. 5 shows the numerical results in a force-time graph. The results shows that the average behaviour is the same for all the mesh densities. However, some differences can be noted. A finer mesh captures quite well the oscillations typical of the impact phenomenon while a coarser mesh, which suffers from high distortion of the elements, has a force-time curve characterized by small amplitude smoothed oscillations.

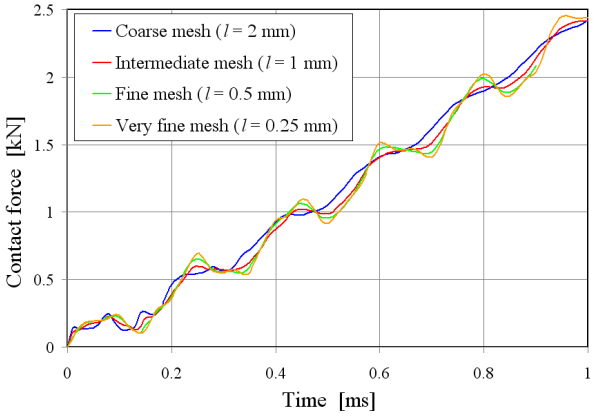


Fig. 5 Mesh sensitivity analysis, force-time curves.

Unsurprisingly, the better result is obtained with higher computational costs so a compromise is found by using a hybrid meshing technique: a fine density (l = 0.7 mm) is used in the contact zone and an intermediate density (l = 1.9 mm) is employed in the remaining part of the model. The final model consists of 10560 C3D8R solid elements and 3600 COH3D8 cohesive elements, with a total of 16775 nodes, Fig. 6.

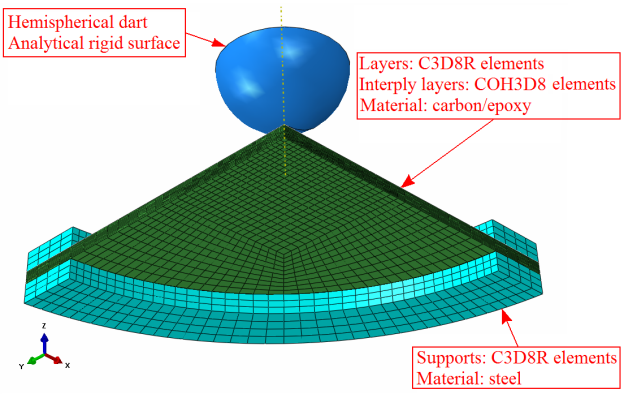


Fig. 6 Overview of the selected mesh.

The selected mesh has been analysed considering a quite tight convergence criterion with ratio R_n^α of the largest residual to the corresponding average flux norm equal to 0.1 [13]. Since this value did not allow to reach the convergence for the whole simulation a restart

was necessary employing a relaxed convergence criterion (a unit value has been used due to the fact that the unloading phase of the simulation is less affected by the damage phenomena).

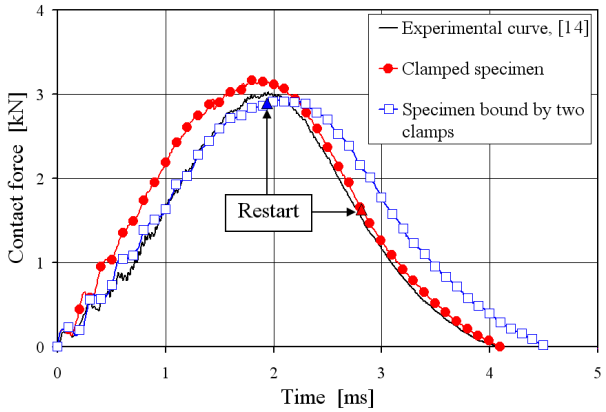


Fig. 7 Force-time curves.

Fig. 7 shows the numerical force-time curve for the case of the clamped specimen together with the experimental curve, whereas Fig. 8 shows the force-displacement curves. The overall numerical response of the clamped specimen is significantly stiffer than the experimental result, leading to a lower maximum displacement of the impactor.

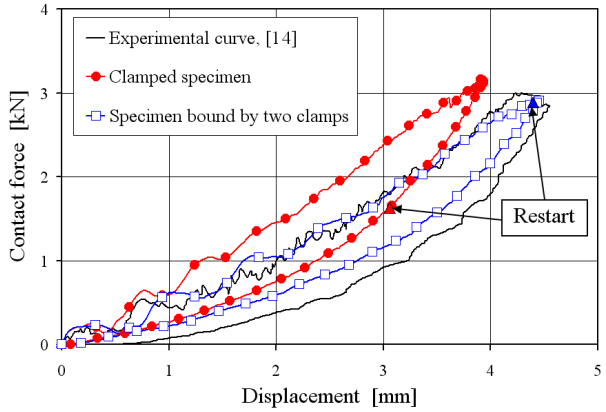


Fig. 8 Force-displacement curves.

Since ASTM D5628 also permits to test unclamped specimen, and as the real support condition for the test case under study were not precisely known, an additional numerical simulation has been carried out, considering the same specimen in a different support set-up. For this additional simulation the upper support disk has been removed and two clamps, modelled as analytically rigid surfaces with a diameter of 10 mm, have been introduced (Fig. 9).

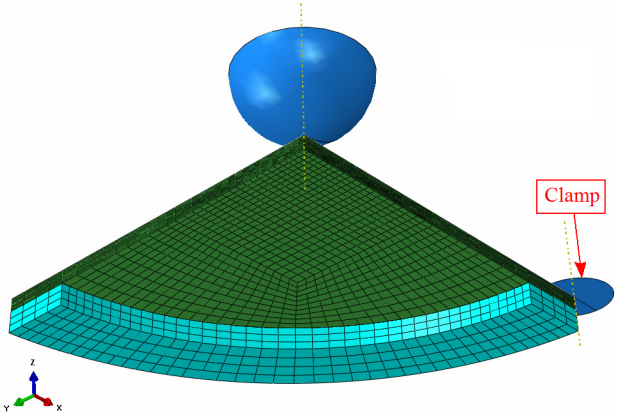


Fig. 9 Specimen bound by two clamps.

From Fig. 7 and Fig. 8, it can be noted that the response of the modified model is less rigid matches the experimental curve quite closely. Table 3 summarize the obtained numerical results, comparing them with the experimental data. Fig. 10 shows the comparison between the extension of delaminated zone obtained by the numerical simulation (clamped specimen) and that obtained by the experimental test. SDEG represents the damage parameter D as indicated in paragraph 4. A good agreement has been achieved.

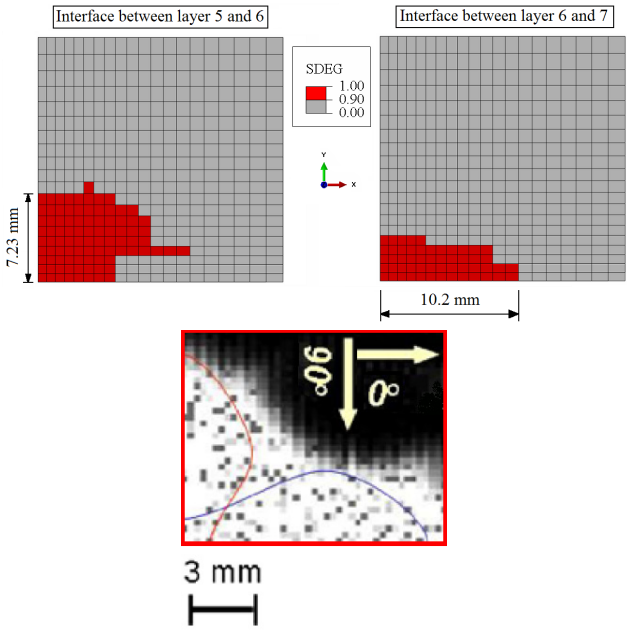


Fig. 10 Delaminations, comparison between experimental result [14] and numerical estimation.

Table 3 Summary of numerical prediction and experimental results.

	Max Displ. [mm]	Displ. error [%]	Max Force [N]	Force error [%]
Experimental	4.55	-	3015	-
Clamped specimen	3.94	-13.4	3170	5.1
Specimen bound by two clamps	4.46	-1.9	2930	-2.8

7 Conclusion

A three-dimensional intralaminar and interlaminar damage model has been proposed using the Continuum Damage Mechanics and the Cohesive Zone Model. A thin composite plate subjected to a low-velocity impact has been analysed. The clamped boundary conditions led to stiff numerical results, whereas the specimen bound by two clamps has shown a behaviour similar to the experimental result. The prediction of the extension of the delaminations is in good agreement with the experimental results.

References

[1] Hinton M.J. and Soden P.D., "Predicting failure in composite laminates: The background to the exercise", *Composite Science and Technology*, Vol. 58, 1998, pp. 1001-1010.

[2] Soden P.D., Hinton M.J. and Kaddour A.S., "A comparison of the predictive capabilities of current failure theories for composite laminates", *Composites Science and Technologies*, Vol. 58, 1998, pp. 1225-1254.

[3] Falzon B.G. and Apruzzese P., "Numerical analysis of intralaminar failure mechanisms in composite structures. Part I: FE implementation", *Composite Structures*, Vol. 93, 2011, pp. 1039-1046.

[4] Dávila C.G. and Camanho P.P., "Failure criteria for FRP laminates in plane stress", NASA Report, NASA/TM-2003-212663, pp. 1-23, 2003.

[5] Pinho S.T., Dávila C.G., Camanho P.P., Iannucci L. and Robinson P., "Failure models and criteria for FRP under in-plane or three-dimensional stress states including shear non-linearity", NASA Report, NASA/TM-2005-213530, pp. 1-64, 2005.

[6] Knight Jr. N.F., User-defined material model for progressive failure analysis, NASA Report, NASA/CR-2006-214526, pp. 1-84, 2006.

[7] Bažant Z.P. and Planas J., *Fracture and size effect in concrete and other quasibrittle materials*, CRC Press LLC, 1998, Chaps 8.

[8] Elder D.J., Thomson R.S., Nguyen M.Q. and Scott M.L., "Review of delamination predictive methods for low speed impact of composite laminates", *Composite Structures*, Vol. 66, 2004, pp. 677-683.

[9] Pinho S.T., Iannucci L. and Robinson P., "Physically-based failure models and criteria for laminated fibre-reinforced composites with emphasis on fibre kinking: Part II: FE implementation", *Composites Part A: applied science and manufacturing*, Vol. 37, 2006, pp. 766-777

[10] Zobeiry N., Forghani A., McGregor C., Vaziri R. and Poursartip A., "Progressive damage modeling of composite materials under both tensile and compressive loading regimes", *ECCOMAS Conference on Mechanical Response of Composites*, Porto (Portugal), 12-14 September 2007, P. P. Camanho et al. (eds.), pp. 179-195.

[11] Lopes C.S., Camanho P.P., Gürdal Z., Maimí P. and Gonzalez E.V., "Low-velocity impact damage on dispersed stacking sequence laminates. Part II: Numerical simulations", *Composite Science and Technology*, Vol. 69, 2009, pp. 937-947.

[12] Turon A., Dávila C.G., Camanho P.P. and Costa J., "An engineering solution for mesh size effects in the simulation of delamination using cohesive zone models", *Engineering Fracture Mechanics*, Vol. 74, 2007, pp. 1665-1682.

[13] Abaqus Analysis User's Manual Ver. 6.10-3, Dassault Systèmes, 2010.

[14] Tita V., de Carvalho J. and Vandepitte D., "Failure analysis of low velocity impact on thin composite laminates: Experimental and numerical approaches", *Composite Structures*, Vol. 83, 2008, pp. 413-428.

[15] ASTM D5628-10, "Standard test method for impact resistance of flat, rigid plastic specimens by means of a falling dart (tup or falling mass)", American Society for Testing and Materials, 2010.

[16] Kärger L., Baaran J., Gunnion A. and Thomson R., "Evaluation of impact assessment methodologies. Part I: Applied methods", *Composites: Part B*, Vol. 40, 2009, pp. 65-70.

CEAS 2011

Modelling and Simulation 2

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Numerical Design of Leading Edge Flow Control over Swept High-Lift Airfoil

S. Mahmood, P. Scholz and R. Radespiel
Technische Universität Braunschweig, Institut für Strömungsmechanik, Braunschweig, Germany

Keywords: *vortex generator jets, leading edge stall, swept high-lift airfoil*

Abstract

The present study aims at the numerical design of the leading edge flow control over a swept two-element high-lift airfoil using steady state blowing via vortex generator jets. The vortex generator jets are applied at the leading edge of the airfoil to prevent the leading edge type stall observed in the baseline cases without flow control. The numerical design involves the variation of several key flow control parameters such as the pitch angle, skew angle, blowing ratio, sense of rotation and vortex generator jets spacing. The results indicate that numerical design of leading edge flow control seems feasible for such wing arrangement. It shows an acceptable degree of stall suppression over the swept arrangement of high-lift airfoil.

1 Introduction

The use of pneumatic flow control devices like vortex generator jets (VGJs) have been studied quite often in the past to delay the stalling of an airfoil. This paper describes the numerical design of leading edge (L/E) flow control over a swept airfoil. The numerical work is performed within the project LEBox which is part of the European project Joint Technical Initiative for Aeronautics and Air Transport (CLEANSKY). The task is challenging due to difficulties to compute L/E type stall. The VGJ geometry for actuation at the L/E of the swept airfoil has been evaluated in terms of a variation of span-wise distance between the VGJ, favorable and unfavorable sense of rotation of the vortices with reference to the sweep-back angle φ and optimum size of the VGJ relative to the boundary layer. It is well known that boundary layers of swept airfoils feature a considerable twist which in turn strongly favors longitudinal vortices with a certain sense of rotation [1,2,3]. It is one goal of the studies made here to utilize this effect to increase the efficiency of L/E actuation with VGJs. It has been stated in [1,2] that the vortex should be oriented in a way such that the local vectors of the secondary flow point into the direction of the local 3D flow vector. Later on [3,4] confirmed, that the local vectors of the secondary flow should be aligned in the direction of the local 3D flow vector of the undisturbed boundary layer. This yields an increased effectivity of the so-called co-rotating configuration. The optimal spanwise spacing λ between the VGJs found by [4] was about $\lambda = 6 \cdot d$, where d being the diameter of the VGJ and [3] reported spanwise spacing

of $\lambda = 10 \cdot d$. Smaller spacings would require a larger mass flow of pressurized air. Therefore it seemed reasonable to rather head towards larger spacings and for the present study we choose $\lambda = 10 \cdot d$ as reported by [3].

2 Numerical Approach

The test case used is the DLR F15 high-lift airfoil [5] with a reference chord length of $c = 600mm$ as shown in Fig.1. It has been recognized that the flap setting has a large influence on the behavior of leading edge actuation [5,6]. The flap setting chosen for this airfoil is called fs#2 [6] and exhibits the L/E type stall with the flap flow largely separated.

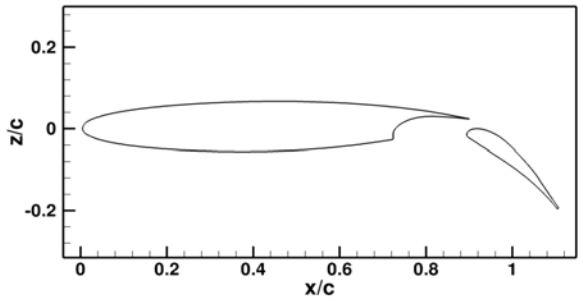


Fig.1 Slat-less DLR F15 high-lift airfoil

This fs#2 flap setting has a flap angle of δ_F=45°, gap g_F/c=2.7% and overlap ovl_F/c=0.5%.

The RANS computations for the present study have been performed with the *DLR TAU-code* [7] of the German Aerospace Centre. The *TAU-code* solves the three-dimensional compressible Reynolds-averaged Navier-Stokes equations. The *TAU-code* is flexible in solving different types of grids like hybrid meshes having prism layers and tetrahedral mixed with hexahedral cells. For the current numerical study mesh generation was carried out with the commercial package GRIDGEN V15.15 [12] by Pointwise Inc. The numerical setup used here involves the use of fully structured 2D grids to simulate the 2.5D infinite swept airfoil for the baseline (without flow control) case and by imposing spanwise periodic boundary conditions along with the sweep-back angle. For L/E flow control cases fully structured hexahedral 3D grids were composed of 13.0x10^6 million grid points. The estimation of 3D mesh was based on the linear scaling of grid points compared to the mesh used in [10,11] to simulate the effect of varying VGJ parameters. The flow control cases resolve a single VGJ with span-wise periodicity. The VGJs are located at $x/c = 1\%$ on the pressure side of the airfoil. The flow control computations have been performed with steady state blowing via a special *actuation* boundary condition available in the TAU-code. The numerical computations used 512 to 1024 CPUs to analyze the VGJs. Computational resources were provided by the Nordeutscher Verbund für Hoch- und Höchstleistungsrechnen (HLRN) [13]. Unswept baseline simulations were carried out with the Spalart Allmaras turbulence model (SA) [8] while the infinite swept baseline computations used both SA and Menter's Shear Stress turbulence model (MSST) [9]. For the 2.5D flow control computations, turbulence transport was simulated using the MSST turbulence model as it gives better prediction of maximum lift coefficient and proved to be superior in simulations of VGJs on a flat plate [10].

Fig.2 shows the fully structured grid in the periodic plane and three dimensional grid in the vicinity of the co-rotating VGJs. The 2D baseline computations were conducted at a free stream Mach number of Ma=0.15 with the Reynolds number of Re=2.0x10^6 based on the reference chord length. The sweep angle was set to $\varphi = 30°$. The 2.5D computations were performed on a slightly higher Mach number increased by a factor of $1/\cos\varphi$ to achieve Ma=0.15 normal to the leading edge of the airfoil. Note that all the 2.5D infinite swept baseline and flow control computations are presented in a coordinate system normal to the L/E of the airfoil. That is airfoil chord, angle of attack and velocity normal to the leading edge were used to compare aerodynamic coefficients.

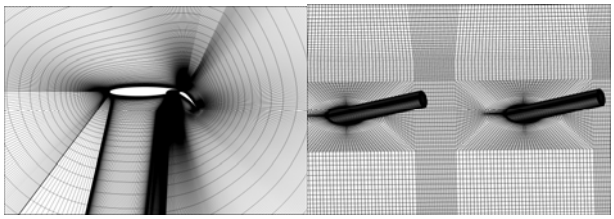


Fig. 2 Fully Structured Mesh in the periodic plane and mesh topology around VGJ's

Fig.3 shows the 2.5D arrangement for the high-lift airfoil with several arrays of co-rotating VGJs. The enlarged view in the vicinity of the L/E VGJs shows the geometric parameters considered in the present paper. The flow control parameters used are the pitch angle $(\alpha = 30°,15°)$ and the skew angle $(\beta = 90°,75°,105°)$. u_j is the jet velocity, λ defines the spacing between two VGJs, d is the diameter of the VGJs and U_∞ is the free-stream velocity with a sweep-back angle φ respectively. Note that only round jets were investigated.

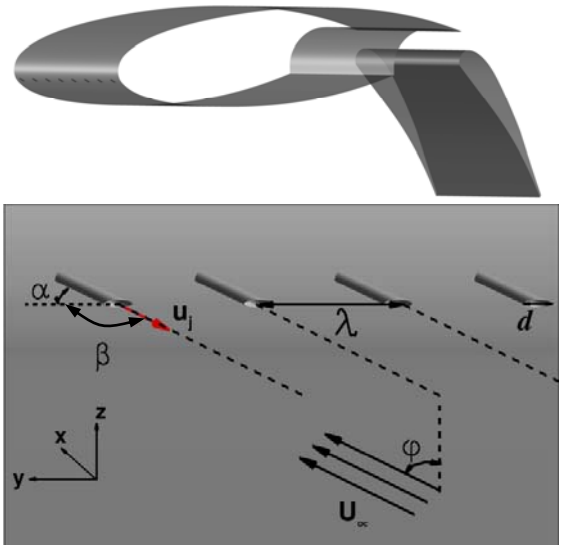


Fig. 3 L/E co-rotating VGJs flow configuration

3 Results

3.1 Grid convergence

For the present RANS computations, grid independence studies were carried out for the 2D unswept baseline case.

A set of four different grids were used to study the effect of grid-size and to obtain a grid-converged solution. The different grids consist of: Coarse grid [CG](5.0 x 10^4 grid points[gp]), Baseline grid [BG] (9.0 x 10^4 gp), Fine grid [FG](1.9 x 10^5 gp) and Very Fine grid [VFG](3.0 x 10^5 gp). Fig. 4 shows the influence of different grid sizes on the lift coefficient C_l and drag coefficient C_d of the airfoil. It can be seen that there is minor difference in the aerodynamic coefficients between the FG and VFG. This observation has been confirmed by the fig. 5 which shows the

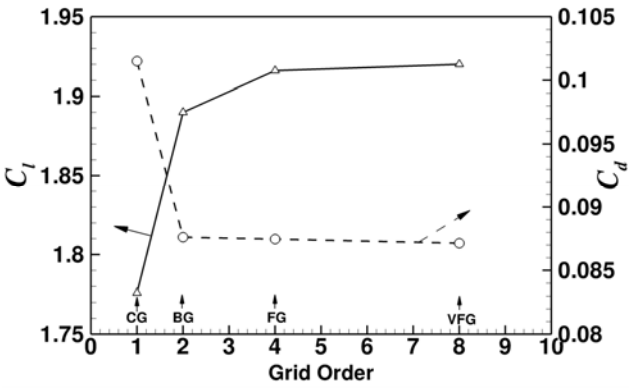


Fig. 4 Influence of grid converged result

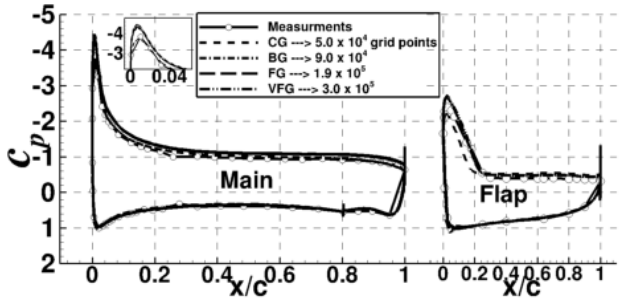


Fig. 5 Grid Converged pressure distribution

surface pressure distribution over the high-lift airfoil computed with different grid sizes. It can be observed that the difference in results from FG grid and VFG is small as seen in fig. 4 and a grid independent solution is obtained. So the FG grid has been chosen as the optimal grid to carry out both the 2.5D type infinite swept baseline and the L/E flow control computations, respectively.

3.2 Baseline flow

This section describes the 2D unswept and 2.5D infinite swept baseline computations over the high-lift airfoil. The 2D unswept baseline results were obtained using SA turbulence model. 2.5D swept baseline computations were carried out using both the SA and MSST turbulence models. The computations were performed in full-turbulent mode. Fig. 6 shows that both the 2D and 2.5D computations predict L/E type stall as observed in the experiments [6] for the 2D setup. SA model predicts much higher maximum C_l compared to the MSST

model. This over prediction is due to the inferior performance of SA model in regions of adverse pressure gradient and the separated regions. On the other hand MSST model performs generally better than the SA model. This argument is supported by the results shown in fig. 7 at constant C_l value of 2.2 as highlighted in fig. 6. The results show that the MSST yields a reasonable suction peak region of the main element and reproduces the flap separation very well.

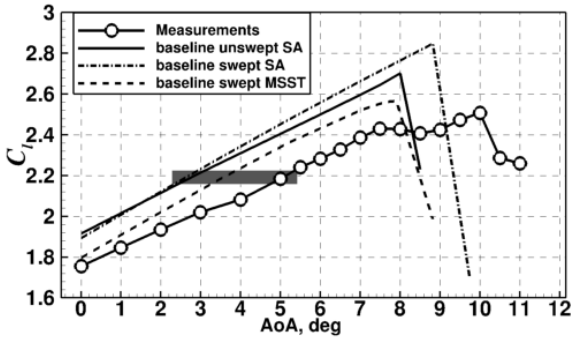


Fig. 6 Baseline unswept and swept lift coefficients

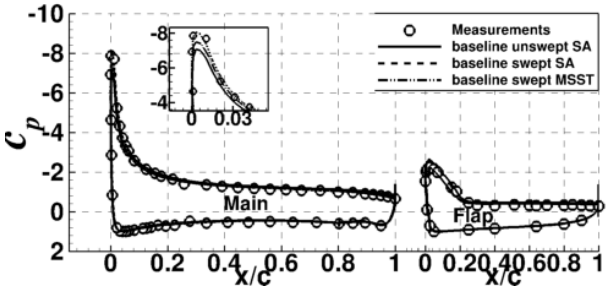


Fig. 7 Comparison of the pressure distribution at constant lift coefficient

3.3 L/E flow control

This section provides the results of the 2.5D flow control computations carried out for different VGJ configurations. As mentioned earlier the resolved VGJs were computed with MSST model and Unsteady RANS computations were performed to obtain the mean aerodynamic values. In order to optimize the VGJs for swept high-lift airfoil, one particular stalling angle of attack was studied with the various flow control parameters. As shown earlier in fig. 6 numerical computations predicted L/E type stall similar as observed in the experiments. So the stalling angle of attack (AoA = 8.5°) as shown in fig. 6, is selected for the flow control cases. Among the various numerically tested flow control settings, four optimized VGJ configurations are shown in fig. 8 and the results are compared with the baseline data of the MSST model. It is worth noticing that all the flow control computations suppressed the L/E type stall. The first flow control case is the one with favorable rotation relative to the sweep angle φ and it has a spanwise VGJ spacing of $\lambda = 10 \cdot d$. It preserves the increase in lift with increasing angle of attack and suppresses stall. The fourth flow control case shown in fig. 8 is the one with unfavorable orientation of the VGJ operating against the oncoming swept boundary layer. It proved to be less effective comparatively with the favorable one. This statement is further strengthened by the detailed inspection of the flow as shown below. Both the favorable and the unfavorable VGJ were computed with the pitch angle α=30° and skew angle β=90°.

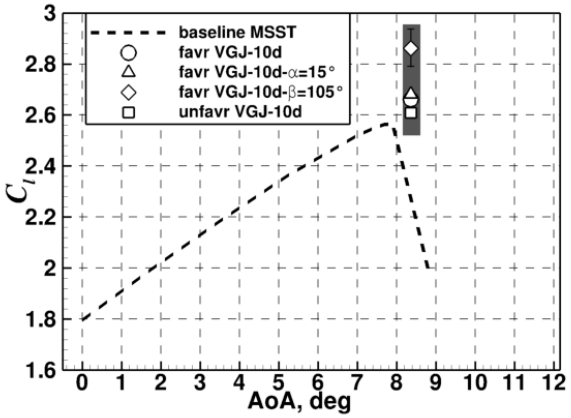


Fig. 8 Comparison of the lift coefficient between the baseline and various flow control parameter studies

The effect of varying the pitch angle was also computed. The results show that the pitch angle has an influence on the lift of the airfoil. The value of $\alpha = 30°$ proved to be the best configuration. The next flow control computation involved the varying skew angles. This parameter was tested because the baseline results around the maximum C_l region showed that the stream lines within the boundary layer are generally not aligned to the surface stream lines. Therefore another computation with skew angle $\beta = 105°$ was carried out by leaving the

pitch angle, jet diameter and spanwise spacing of VGJ unchanged. As can be seen from fig. 8 this arrangement proved to be the most effective one. The bars for the lift coefficient show the varying amplitude of the lift as computed by the URANS solver for this case. The maximum increase of $\Delta C_{l,Max} = 0.33$ was then obtained by choosing the parameters $\lambda = 10 \cdot d$, $\alpha = 30°$ and $\beta = 105°$.

Fig.9 shows the difference between the baseline and L/E flow control in terms of the normalized longitudinal velocity contours along with the stream lines. The result shows that the flow control prevents the separation of the flow observed in baseline case at AoA= 8.5°. Fig. 10 shows numerical representation of the L/E flow control over the upper surface of the swept airfoil. The plot shows the structured grid along with the numerical display of the longitudinal vortices colored by contours of constant vorticity. These contour surfaces are colored according to values of the pressure coefficient in order to display the boundary layer development.

Fig. 11 shows the spanwise averaged pressure and skin friction distribution in the vicinity of the L/E of the airfoil. The results obtained from the best flow control case with locally adapted skew angle produce a higher suction peak for the surface pressure and a larger positive peak for the skin friction coefficient.

3.3.1 Vortex strength

The strength of the vortices generated by the VGJs was judged by monitoring the peak vorticity and circulation in the cross-flow plane downstream of the VGJ. The circulation was calculated by integrating the vorticity over the cross-flow plane. Here we denote the peak vorticity by ω_{peak}, normalized with maximum peak vorticity ω_{max} in the vortex core of the best configuration with $\lambda = 10 \cdot d$, $\alpha = 30°$ and $\beta = 105°$. The circulation is denoted by the circulation symbol Γ. Fig. 12 shows the effectiveness of VGJs w.r.t to different sense of rotation and varying pitch and skew angles in terms of the vorticity and the circulation downstream of the VGJ. Both the normalized peak vorticity and circulation were plotted against the normalized surface arc length of the airfoil, where the valued of zero denotes the position of the VGJ. The result shows that the favorable VGJ yields higher vortex strength downstream of the actuation in comparison to the unfavorable VGJ. Also the effect of locally adapted skew angle provides higher peak vorticity and vortex strength. Another finding is the effect of using pitch angle $\alpha = 15°$ that gives much higher peak vorticity among the other flow control cases as shown in fig. 12 by an enlarged view in the vicinity of $\Delta\xi / c = 0.12 \sim 0.16$. This is supported by large increases of skin friction already shown in fig. 11.

3.3.2 Several Other VGJ configurations

Several other VGJ computations addressed the effect of different spanwise spacing λ, velocity ratios (VR) which is the ratio of jet velocity to the free-stream velocity and an influence of counter rotating VGJs (ctr-rot VGJ). All these computations were carried out with standard flow control parameters like pitch angle $\alpha = 30°$ and skew angle $\beta = 90°$ unless stated otherwise. The results are shown in fig. 13 where the comparison of maximum lift coefficient between the reference baseline and several other VGJ test cases is made. It can be concluded from the results that increase in spanwise spacing, $\lambda = 20 \cdot d$ decreases the lift considerably while decreasing the VGJ spacing $\lambda = 5 \cdot d$ does not increase the lift coefficient compared to $\lambda = 10 \cdot d$. So the initial spacing with $\lambda = 10 \cdot d$ proved to be an optimal one. Similarly, constant blowing at lower VRs=[5.04, 4.45] than the standard one VR=5.78 shows less lift.

4 Conclusion

This paper shows that the numerical design of L/E flow control over the swept high-lift airfoil can be done. The numerical baseline results predicted the L/E type stall with a given flap setting as observed in the experiment. VGJ computations with favorable orientation proved to give higher lift in comparison to the

unfavorable VGJ configuration. Spanwise spacing of VGJs was optimized with $\lambda = 10 \cdot d$, decrease or increase in VGJs spacing causes lift losses. Computations with varying pitch angle and adapted skew angle show potentials for increasing the lift coefficient. Future wind tunnel testing planned as part of the project will determine whether the computed gains can be realized in practice.

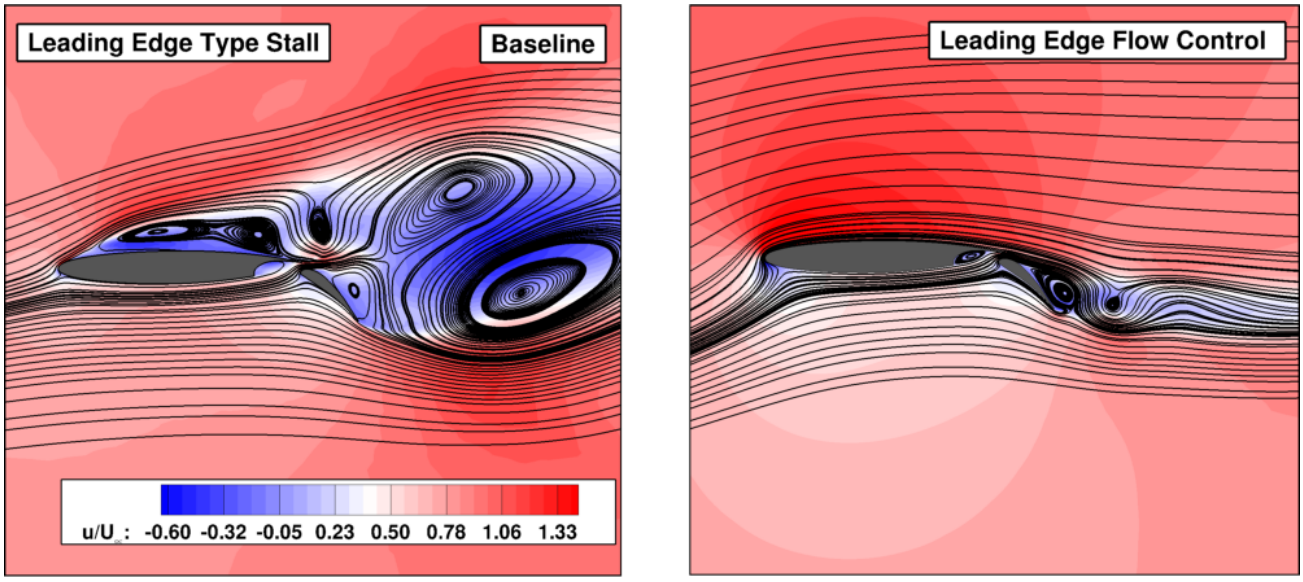


Fig. 9 Normalized velocity contours and streamlines over the airfoil exhibit the stall suppression due to L/E flow control

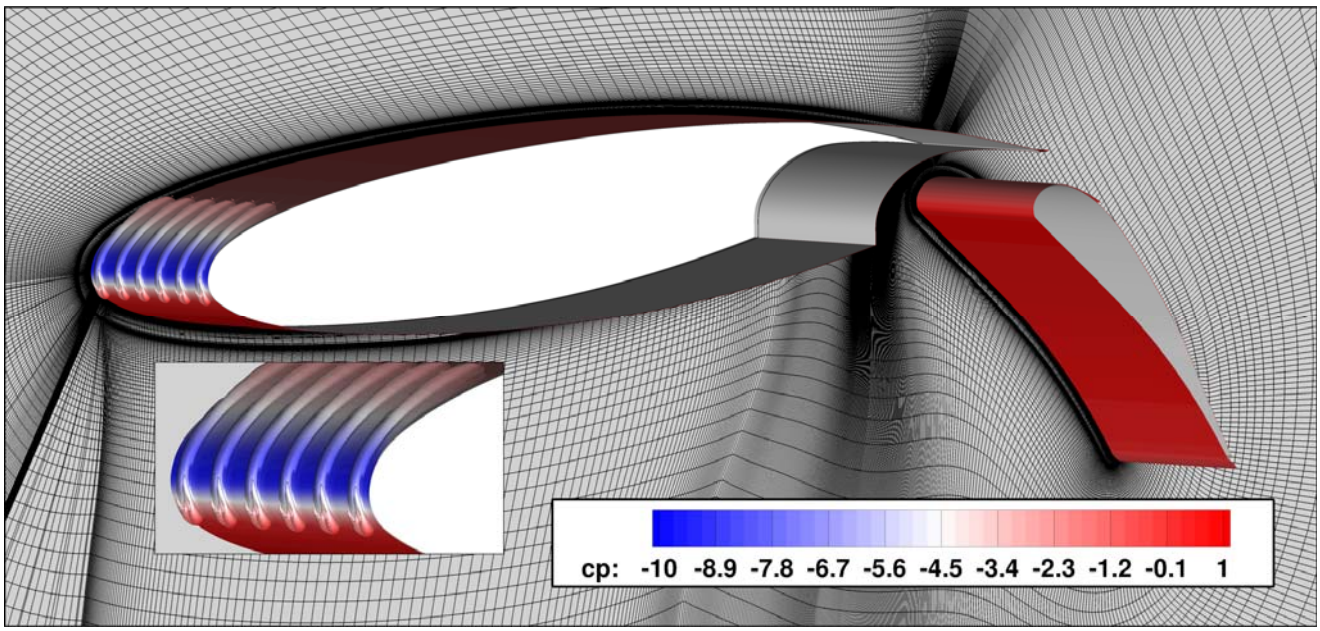


Fig. 10 Numerical representation of L/E flow control resulting in longitudinal vortices over the swept high-lift airfoil using vorticity-criterion

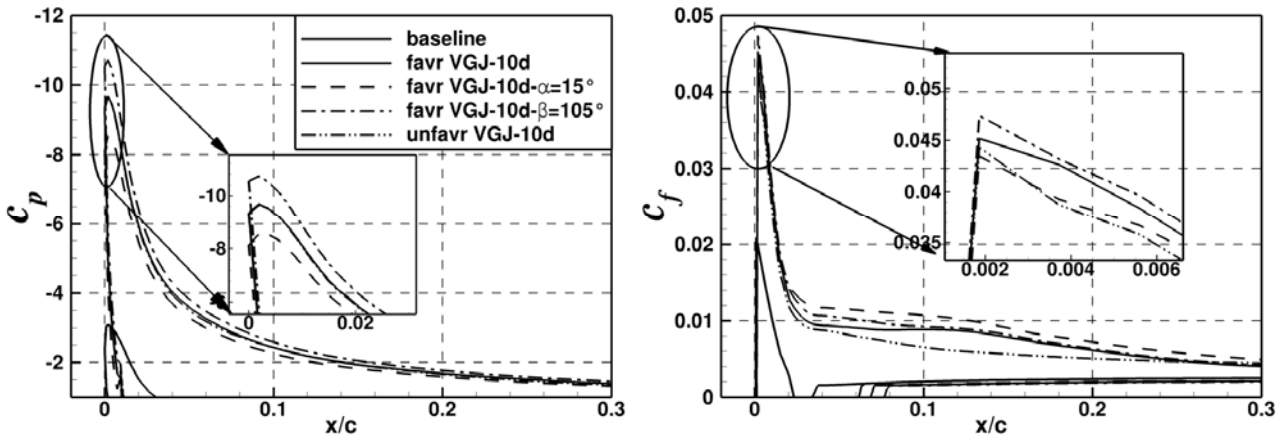


Fig. 11 Spanwise averaged pressure and skin-friction distribution in the vicinity of the L/E of the main element

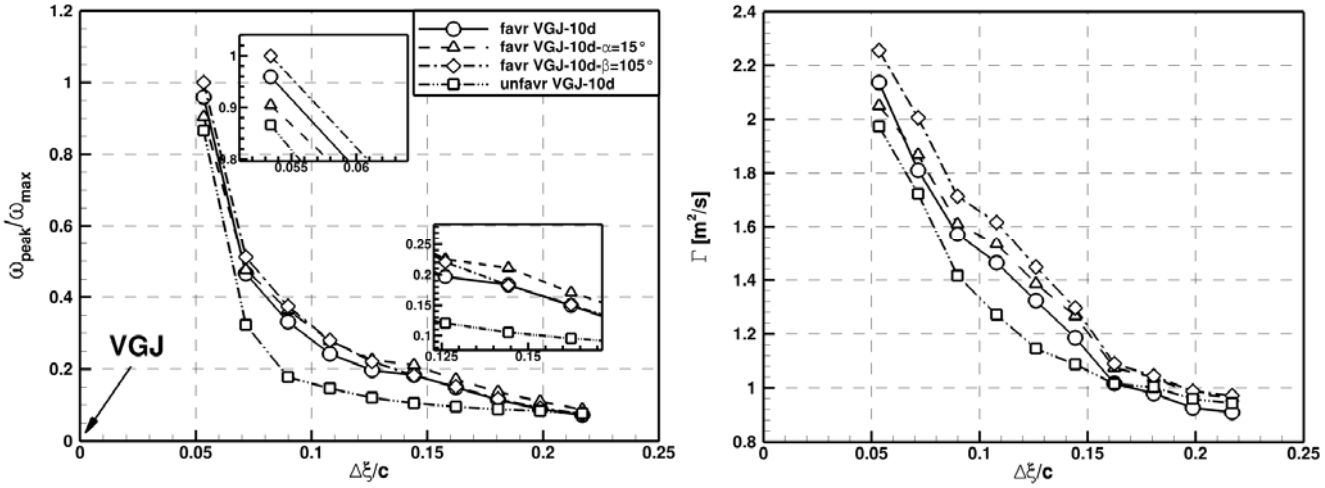


Fig. 12 Vorticity and circulation plot downstream of the VGJ

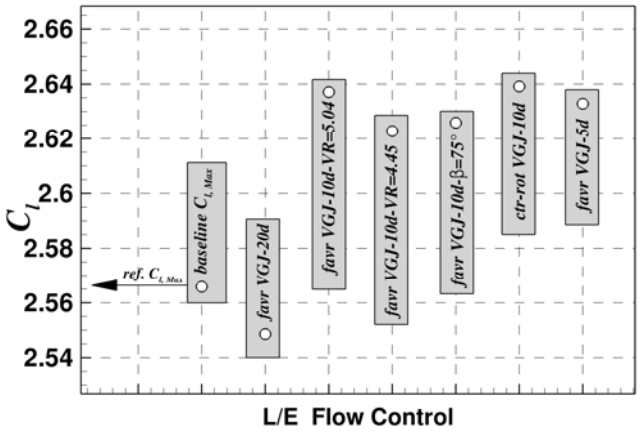


Fig. 13 Lift coefficient by various VGJ parameters

Acknowledgements

The project is approved and funded by European Union via Grant Agreement CS-GA-2009-255752. The support of the "Nordeutscher Verbund für Hoch- und Höchstleistungsrechnen (HLRN)" by providing the computational resources is gratefully acknowledged.

References

[1] Shizawa T. and Eaton J.K., "Interaction of a longitudinal vortex with a three-dimensional turbulent boundary layer", *AIAA Journal*, Vol. 30, No. 5, 1992a, pp. 1180-1181.

[2] Shizawa T. and Eaton J.K., "Turbulence

measurements for a longitudinal vortex interacting with a three-dimensional turbulent boundary layer", *AIAA Journal*, Vol. 30, No. 1, 1992b, pp. 49-55.

[3] Meunier M. and Brunet V., "High-Lift Devices Performance Enhancement Using Mechanical and Air-Jet Vortex Generators", *Journal of Aircraft*, Vol. 45, No. 6, 2008, pp. 2049-2061.

[4] Godard G. and Stanislas M., "Control of decelerating boundary layer. Part 3: Optimization of round jets vortex generators", *Aerospace Science and Technology*, Vol. 10, No. 6, 2006c, pp. 455-464.

[5] Wild J., Wichmann G., Haucke F., Peltzer I., and Scholz P., "Large Scale flow control experiments with the German Flow Control Network", *47th AIAA Aerospace Sciences Meeting*, Orlando, USA, AIAA 2009-0530, 5-8 Jan, 2009.

[6] Scholz P., Kähler C.J., Radespiel R., Wild J. and Wichmann G., "Active control of Leading-Edge Separation within the German Flow Control Network", *47th AIAA Aerospace Sciences Meeting*, Orlando, USA, AIAA 2009-0529, 5-8 Jan, 2009.

[7] Schwamborn D., Gerhold T. and Heinrich R., "The DLR TAU-Code: Recent Applications in Research and Industry", *In: P. Wesseling, E. Oate, J. Priaux (Eds), ECCOMAS CFD, Egmond aan Zee*, The Netherlands, 2006.

[8] Spalart P.R. and Allmaras S.R., "A one-equation turbulence model for aerodynamic flows", *La Rech.*, Aerospatiale No.1, 1994, pp. 5-21.

[9] Menter F.R., "Zonal two-equation k-w turbulence model for aerodynamic flows", *AIAA paper*, AIAA 93-2906, 1993.

[10] Mahmood S. and Radespiel R., "RANS simulation of jet actuation in a boundary layer flow using Chimera grids," *Deutscher Luft- und Raumfahrtkongress*, Aachen, Germany, 8-10 Sep, 2009.

[11] Scholz P., Mahmood S.S., Casper M., Radespiel R., Sitzmann M., and Ciobaca V., "Experimental and Numerical Investigations on the Control of Airfoil Stall using Vortex Generator Jets", *5th Flow Control Conference*, Chicago, Illinois, AIAA 2010-4250, 28 June-01 July, 2010.

[12] Commercial CFD software by POINTWISE, Inc. http://www.pointwise.com/gridgen

[13] Norddeutscher Verbund für Hoch und Höchsleistungsrechnen, http://www.hlrn.de, 2011.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Straight Wing and a Forward Swept Wing Compared with a Curved Planform Wing in the Transonic Regime

M. R. Chiarelli, G. Lombardi and A. Nibio
Department of Aerospace Engineering – University of Pisa, Italy

Keywords: *drag reduction, pressure drag, shock wave drag, curved planform, CFD analyses*

Abstract

The effects of the planform on the aerodynamic drag have been numerically investigated and compared, in the high subsonic regime and low angles of attack, for three wing models: a straight wing, a forward swept wing and a forward curved planform wing. The semispan (0.7 m), the aspect ratio (5.7), the taper ratio (0.4) and wing sections (NACA 0012) are the same.

For the straight and swept wing models some experimental results are available in the literature. CFD analyses have been executed using three structured grids: the code STAR–CCM+® has been used in the research.

Drag polar curves, at Mach number 0.7, confirm that the planform of a wing strongly influences the pressure distribution and the drag coefficient: with a lift coefficient ranging through –0.25 and 0.25 the curved planform wing develops the minimum of the drag coefficient thanks to reduction of both pressure drag and shock wave effects.

1 Introduction

The physical effects associated with sweep of wings, valid also for higher flight Mach numbers, are well explained in [1]: the theoretical arguments are based on the concept of a swept wing of infinite aspect ratio, the so called *infinite sheared wing*.

In [1] three distinct physical effects of sweep are highlighted: (a) the shape of the streamlines over the sheared wing are curved in planview (this fact partially occurs also over a swept wing of finite span), (b) the perturbation velocity varies more slowly with Mach number as the angle of sweep is increased and (c) the critical conditions are reached only when the velocity component normal to the direction of sweep is sonic. As described in [1] (pag. 114), following a conjecture proposed by W. G. Bickley [2] the "criterion" (c) can be generalized by introducing the concepts of isobars: on the surface of a swept wing of arbitrary shape the isobars are, in general, curved lines and can be said that the critical conditions are reached when the local velocity component normal to the isobars, i.e. in the direction of the local pressure gradient, reaches the local velocity of sound. Further considerations arise about the isobar patterns on the surface of a swept wing: it is beneficial to design swept wings to have quite regular isobars pattern and *fully-swept isobars* along the span ([1], pag. 223).

Following the over mentioned criteria (b) and (c) based on physical observations and on realistic assumptions, from engineering point of view, a study has been executed on wing with a curved planform, that is a wing with a curved leading edge, a curved trailing edge and with a

tip of finite dimensions (i.e. without a vertex at the wing tip). In previous works [3] and [4], by comparing two wings, a backward swept wing and a backward curved wing with the same aspect ratio and the same wing sections, some preliminary numerical results show that, in the transonic regime, the drag coefficient is strongly influenced by the planform; in particular the curved wing shows the lowest drag coefficient value between the two because the shock wave effects are reduced. In [3] for a wing-body model of a modern transport aircraft a CD (the drag coefficient) reduction of about 4% has been estimated at cruise flight conditions at Mach number of 0.85.

Few examples exist of lifting systems or wing model configurations with unconventional planform, whose boundaries, in a planview, are curved lines: first of all the delta wings of the famous Aerospatial/BAC Concorde [5], with a low values of the aspect ratio; delta wings with curved leading edges in supersonic flow are also treated in [6]; a fundamental study about the plan forms of swept wings in transonic flow has been executed in [7] (also cited in [1]); a study of a blended wing-body with curved plan forms is shown in [8]; the effects of curved planform on the induced drag are discussed in [9] and more recently some publications as [10] (an experimental study on wing models with curved plan forms under subsonic condition) or [11] are available. In the basic work of Lock R.C. [7], adopting a linearised theory for lifting surfaces at Mach=1, the results of a study on the loading distribution (chord-wise and span-wise) on swept wings with curved leading edges are shown. In this case the trailing edge of the wings is assumed to be a straight line and for all the geometries examined the wing tip is always a vertex. For this reason the isobars do not follow in a uniform way the curved shape of the leading edge and along the wing span they are not arranged as the maximum possible value of sweep corresponding to the leading edge itself. On the other hand, as said before, this is a requirement to better reproduce the two-dimensional behavior of the infinite sheared wing with the aim to retard the critical conditions establishment on the surface of a swept wing as stated in [1]. Further examples of wings with unconventional boundaries can be found in [1] (see Figures 5.19 and 5.20 of Chapter 5).

In all the cited works a lack of emphasis persists about the effects that a curved planform or also a curved leading edge produces on the form drag component and/or on the shock wave drag component of a swept wing. These effects are, on the other hand, the basic objectives of the present study and in the next paragraphs some numerical results will be shown which define, in a preliminary way, some suitable technical improvements in the aerodynamic design of swept wings oriented to the drag reduction in the high subsonic flight regime of modern aircrafts. In the present research a comparison of three wing models has been carried out at Mach 0.7: (1) a straight wing, (2) a forward swept wing and (3) a forward curved wing. To set up and to confirm the reliability of numerical models (the Star-CCM+® code has been used in the research) the experimental results (e.g. pressure measurements), available for the straight wing and for the forward swept wing ([12], [13] and [14]), have been used as reference data.

2 The CFD Models of the Wings

The numerical models of three half-wings have been constructed starting from CAD archives and transferring the geometrical information to the pre-processor GAMBIT®. The study of the wing geometries, the straight ([12], [13]), the forward swept ([12], [13], [14]) and the forward curved one, has been carried out according to the geometries shown in Fig. 1. At the root section the leading edge of the curved wing has the same slope of the leading edge of the forward swept wing. The semi-span (0.7 m), the aspect ratio (5.7), the taper ratio (0.4) and wing sections (NACA 0012) are the same for the three wings. All the wing models are untwisted. In the span direction the curved wing has the same chord distribution of the other two wings. In a CAD code the parameters used to define the shape of the leading edge of the curved wing have been fixed. A conic arc of

parabolic type has been adopted. The sweep angle of the leading edge at the root section is equal to -21,37 degrees (as said this angle is the same both for the forward swept and the curved wings). At the tip of the curved wing the sweep angle of the leading edge is fixed to -45 degrees and also the in plane position of the point on the nose of the wing tip has been assigned (Fig. 1).

To execute the fluid dynamic analyses three structured grids composed of about 1,500,000 hexa cells have been built. In order to take into account the viscosity effects, a turbulent and viscous flow conditions have been assumed during CFD analyses. A standard K-ε model has been used to describe turbulence.

The surface grids of the forward swept and curved wings are shown in Fig. 2 and Fig. 3a. On the surfaces of both the upper and lower skins of wings a number of 100 and 30 subdivisions have been assigned in the chord direction and in the span direction respectively.

Figure 3b shows a detail of the surface grid at the tip of curved wing model. The trailing edges of the three wings have been modeled by means of a thin surface: along the height of this surface a number of 8 subdivisions have been assigned.

In Fig. 4a and Fig. 4b are represented the boundaries of the sub-volumes and the grid volume of the curved wing model respectively.

The free stream conditions have been fixed for all the free faces of the grid volume, and a wall boundary condition has been assigned to the plane of symmetry to simulate the effects of wall interference of the transonic wind tunnel.

In the Table 1 the geometry data of wings and the flow reference conditions adopted in the numerical analyses are summarized.

3 Validation of the CFD Analyses

According to the available experimental tests results ([12], [13]) a set of analyses at Mach 0.7 for different values of the angle of attach α have been executed on the three wing models. The analyses have been carried out on a dual core PC adopting the parallel computing capabilities of the fluid dynamic code. For each run has not been fixed an a priori convergence criteria, the analysis has been stopped by executing a control on the stability of the drag and lift coefficients asymptotic values. In general this condition has been reached on for a cumulative iterations number of 2500 or 3000 units.

Due to the symmetry of the wing sections only the positive values of the angle of attack have been investigated in the range of $\alpha = 0°$ to $\alpha = 4°$.

In order to validate grids topology and grids refinement a comparison between numerical and experimental data, relevant to the pressure coefficient distributions, has been executed both for the straight wing and swept forward wing.

The location of the pressure orifices on the test coupons of wings were positioned according to the description available in [12]. In the same sections position along the semi-span of wings (that is at 4%, 18%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 97% of the semi-span) a series of graphs has been drawn to carried out the comparison of local values of pressure coefficients.

Also a comparison of the sectional lift coefficient along the semi-span has been executed as made in the previous work [14].

A comparison of the drag characteristics of the wings has not been executed because of the lack of experimental measurements in the tests carried out in the transonic wind tunnel.

Drag polar curves of the three wings have been numerically evaluated and compared in the present research and a detailed analysis of the results (pressure drag contributes, shock wave effects, skin friction effects) will be shown and discussed in the next paragraphs.

3.1 Straight Wing

The comparison between numerical and experimental pressure coefficients on the upper surface of the straight wing model is represented in the graphs of Fig. 5, Fig. 6, Fig. 7 and Fig. 8. The graphs refer to the data obtained for Mach=0.7 and $\alpha = 4°$.

Along the semi-span of the straight wing, the numerical pressure coefficients data agree very well with the experimental ones. The rise of pressure across the shock wave is also evident in the figures. In this case the shock wave is not

present in the root zone as can be seen from Fig. 5 and as will be shown later on looking at the supersonic zone computed around the upper surface of the wing. From the present results it can be concluded that the assumption of wall boundary condition at the root describes in a good fashion the experimental test conditions which, locally, depended on the boundary layer effects developing at the support plate of the wing models.

Figure 9 shows the span distribution of the numerical and the experimental sectional lift coefficients of the straight wing: the two distributions are very similar and this fact confirms the reliability of present CDF model.

3.2 Forward Swept Wing

The comparison between numerical and experimental pressure coefficients on the upper surface of the forward swept wing model is represented in the graphs of Fig. 10, Fig. 11, Fig. 12 and Fig. 13. The graphs refer to the data obtained for Mach =0.7 and $\alpha = 4°$.

In the present case a good approximation of experimental data has been obtained in the zone near the root of the wing, while, along the semi-span, a sensible difference in the suction peak at the leading edge of the swept wing persists.

In the Fig. 14 the span distribution of the numerical and the experimental sectional lift coefficients of the forward swept wing is shown. The numerical data underestimate the experimental data especially in the central part of the wing span.

In the Fig. 15 the numerical and the experimental lift curves of the swept wing are compared. A good approximation has been obtained for the global values of the lift coefficient CL (an error of 4% persists for the angle of attack $\alpha = 4°$). The numerical values of CL underestimate the experimental ones as the graphs of Cp and the graph of sectional lift coefficient also demonstrate. From a practical point of view, that is the objective of the present research, this fact is favorable: by means of the present structured grids, adopting the viscous turbulent flow assumptions, the numerical efficiency of the forward swept wing will underestimate the experimental efficiency of the tested wing model. On the contrary, in the cited paper [14] using an inviscid Eulerian flow solver and unstructured grids the numerical lift curve overestimates the experimental one (see Fig. 6 in [14]). These observations allow us to conclude that the CFD models of the wings used in the present research will give realistic and good results from an engineering point of view: with a good amount of reasonableness the last assertion can be considered valid also for CFD model of the forward curved wing.

4 Comparison of the Wings Performaces

The validation carried out for the straight and forward swept CFD models has encouraged the construction of similar grid model for the curved wing (see Fig. 3a, Fig. 4a and Fig. 4b).

In the present research the analysis of the curved wing has been executed to demonstrate the effects of the planform shape on the flow characteristics around the wing and to compare the aerodynamic drag, numerically evaluated, of the three wing configurations.

First of all, with reference to the technical criteria treated in Section 1, qualitatively analyses of the shape of streamlines and of the isobars pattern on the three wings will be shown later on. Moreover, a quantitative analysis of the results, looking at the drag components, will allow us to carry out a detailed comparison about the behaviors, at fixed value of the lift coefficient CL, of a curved planform, a swept and a straight wing having the same aspect ratio and the same section profiles.

4.1 Streamlines and Isobars Patterns

The streamlines patterns on the upper surfaces of the wings are represented in Fig. 16, Fig. 17 and Fig. 18. These graphs refer to the same value of CL=0.2 for the three wings. Swept and curved wings develop a streamlines pattern, mainly in the center of wing span, in accordance with the *sheared wing* concept. In the case of straight wing the streamlines are quite aligned with the asymptotic flow direction. Moreover by observing with some shrewdness

the graphs of swept and curved wings it can be seen that in the case of the curved wing the streamlines are more *sheared* with respect to the swept wing case, especially from the middle of the wing span to the tip region of the wing (in this zone in fact the sweep angle of the wing is increasing).

A similar situation occurs if the isobars patterns are compared (Fig. 19, Fig. 20, Fig. 21). In this case the sweep angle distributions of the isobars are much more intelligible and in the case of the curved wing not only persist a *fully-swept isobars* pattern but, along the span direction, the sweep angles of isobars are greater than the swept wing case. This means that for a curved wing the direction of pressure gradient varies along the span and the component of the velocity in this direction (which is normal to the isobar layout) reduces its value with respect a swept wing: the critical condition is then retarded. In fact Fig. 22, Fig. 23 and Fig. 24 demonstrate that the dimensions of the sonic zone around the curved wing are smaller then other two wings.

4.2 Drag Polar Curves of Wings

Figure 25 shows the drag polar curves of the three wings in the positive range of CL (the curves are symmetric) at Mach=0.7. The curved wing develops lower values of CD (higher values of efficiency) at low values of the angles of attack ($-3° < \alpha < 3°$; $-0.25 < CL < 0.25$). In the figure fourth order polynomials, which interpolate the numerical data, are shown.

The estimation of the drag coefficient reduction of the curved wing case *vs* the swept wing case can be carry out for each value of CL. In two particular situations the following results have been computed:

(1) CL=0 CD reduction = 3.78%;

(2) CL=0.2 CD reduction = 1.81%.

To better investigate the effect of planform shape of wings, the components of the drag coefficient have been evaluated by means the CFD code for the previous conditions (1) and (2). The data have been summarized in a table format (from Table 2 to Table 7).

Both pressure and shear components of drag reduce their total values in the sequence: (i) straight wing, (ii) swept wing, (iii) curved wing. The pressure drag and the shear components of the curved wing are smaller than the swept wing case also at CL=0 (in this case no shock waves persist on the wings): it can be said that in a subsonic flow the curved planform shape strongly modifies the pressure distribution on the wing surfaces and also the boundary layer effects are partially reduced. Moreover, at CL=0.2, when a transonic flow persists on the three wings (see Fig. 22, Fig. 23, Fig. 24), a similar behavior can be observed by comparing the Table 7 with the Table 5 and the Table 6.

References

[1] Küchemann D., *The aerodynamic design of aircraft*, Pergamon, London, 1978, ISBN 0-08-020514-3, pp. 103-220.

[2] Bickley W. G., "Critical conditions for compressible flow", ARC R&M 2330, 1946.

[3] Chiarelli M. R., Cagnoni M., Ciabattari M., De Biasio M., Massai A., "High aspect ratio wing with curved planform: CFD and FE analyses", *27th International Congress of the Aeronautical Sciences*, Paper ID 452, CD-Rom Proceedings ISBN 978-0-9565333-0-2, Optimage Ltd, Edinburgh, UK, 2010.

[4] Chiarelli M. R., Cagnoni M., Ciabattari M., De Biasio M., Massai A., "Curved Planform Wings with a High Aspect Ratio for Aircrafts Operating in the Transonic Regime", Pub. No. WO/2010/150212, World Intellectual Property Organization, December 2010.

[5] Taylor J. W. R., *Jane's All the World's Aircraft*, Pilot Press Ltd, 1977.

[6] Vaughn M. E., Burkhalter J. E., "Pressure Loading on Curved Leading Edge Wings in Supersonic Flow", *Journal of Aircraft*, Vol. 23, No. 7, 1986, pp. 574-581.

[7] Lock R. C., "Design of wing plan forms for transonic speeds", *The Aeronautical Quarterly*, Royal Aeronautical Society, London, GB, Vol. 12, No. Part 1, 1961, pp. 65-93.

[8] Holdaway G. H., Mellenthin J. A., "Evaluation of Blended Wing-Body Combinations with Curved Plan Forms at Mach Numbers Up to 3.50", NASA no. 19980227180, 1960.

[9] Ashenberg J., Weihs D., "Minimum Induced Drag of Wings with Curved Planform", *Journal of Aircraft*, Vol. 21, No. 1, 1984, pp. 89-91.

[10] Van Dam C.P., Vijgen P.M.H.W., Holmest B.J., "Experimental Investigation on the Effect of

Crescent Planform on Lift and Drag", *Journal of Aircraft*, Vol. 28, No. 11, 1991, pp. 713-720.

[11] Gratzer L. B., "Curved Wing Tip", Pub. No. US 2010/0181432 A1, Patent Application Publication, US, July 2010.

[12] Lombardi G., "Experimental Study on the Aerodynamic Effects of a Forward-Sweep Angle", *Journal of Aircraft*, Vol. 30, No. 5, 1993, pp. 629-635.

[13] Buresti G., Lombardi G., Morelli M., "Pressure measurements on different canard-wing configurations in subsonic compressible flow", Department of Aerospace Engineering, Pisa, ADIA 91-4, 1991.

[14] Lombardi G., Salvetti M. V., Morelli M., "Appraisal of Numerical Methods in Predicting the Aerodynamics of Forward-Swept Wings", *Journal of Aircraft*, Vol. 35, No. 4, 1998, pp. 561-568.

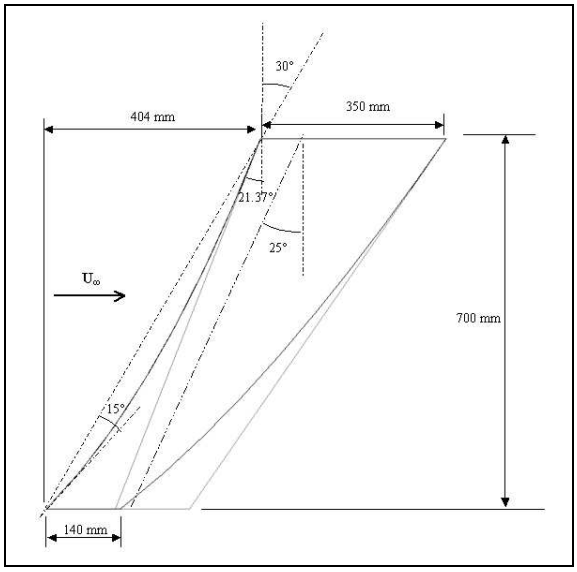


Figure 1 Sketch of the wings geometries (planview)

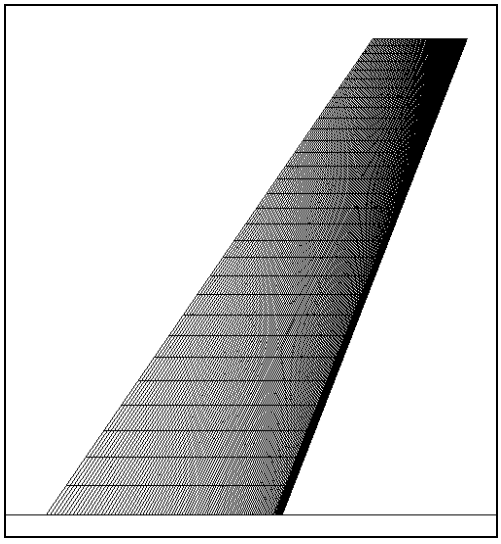


Figure 2 Surface grid of the forward swept wing

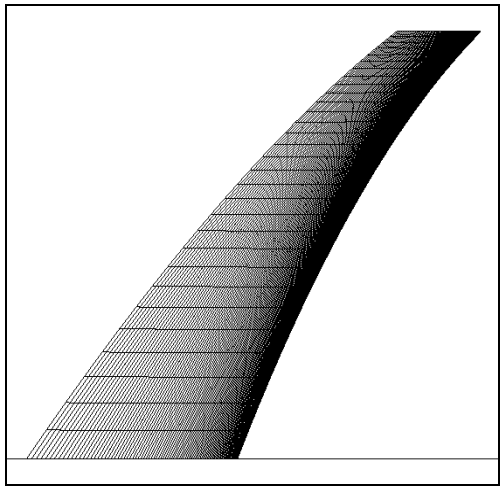


Figure 3a Surface grid of the forward curved wing

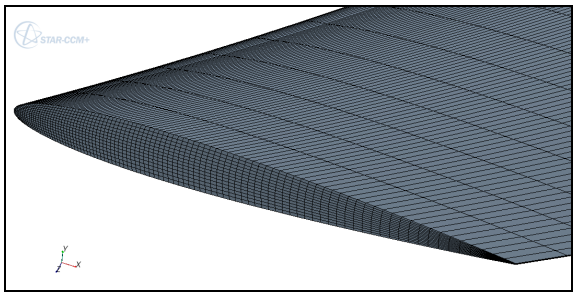


Figure 3b Detail of the surface grid at the tip of the forward curved wing

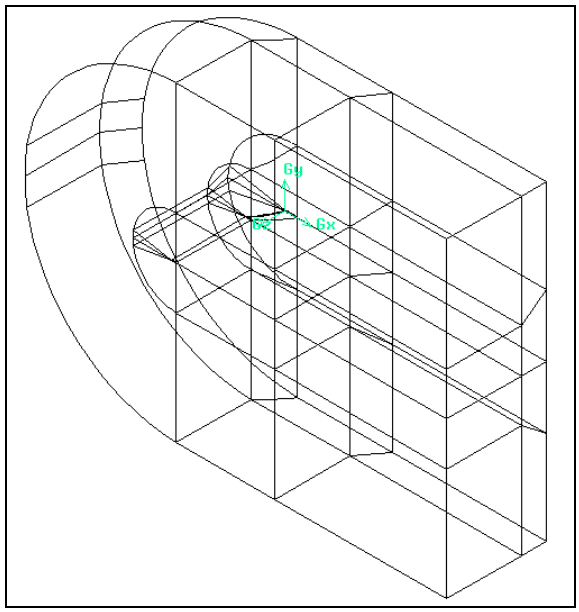


Figure 4a Sketch of the sub-volumes boundaries used in the construction of the grids (curved wing model)

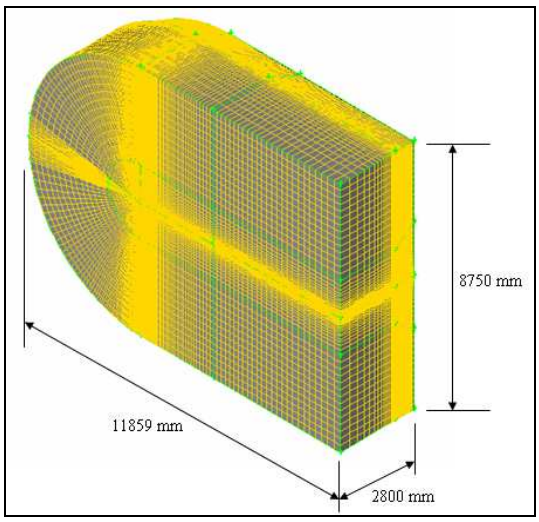


Figure 4b Sketch of the complete grid volume used in the CFD analyses (curved wing model)

Table 1 Geometry parameters of wings and flow reference conditions

Wing reference surface	0.171 m^2
Half Wing Span (b/2)	0.7 m
Sweep angle at the leading edge: Straight wing	4.29°
Sweep angle at the center line (25% c): Forward Swept wing	-25°
Sweep angle at the leading edge: Forward Swept wing	-21.37°
Sweep angle at the leading edge: Root section of the Curved wing	-21.37°
Sweep angle at the leading edge: Tip section of the Curved wing	-45°
Root chord of wings	0.350 m
Tip chord of wings	0.140 m
Mean Aerodynamic Chord	0.260 m
Aspect Ratio of wings (AR)	5.71
Mach Number	0.7
Reynolds Number	~2.8e+06
Reference Density	1.225 kg/m^3
Reference Velocity	238.2 m/s

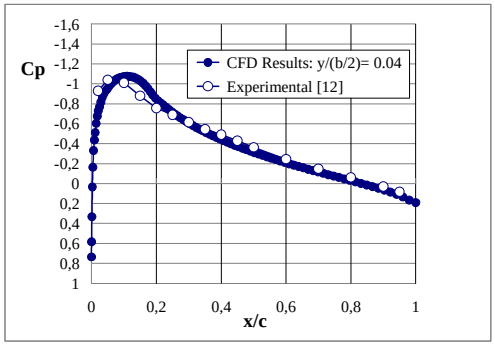


Figure 5 Straight wing model: Cp distribution on the upper surface of the section at 4% of the wing semi-span (α = 4°, Mach =0.7).

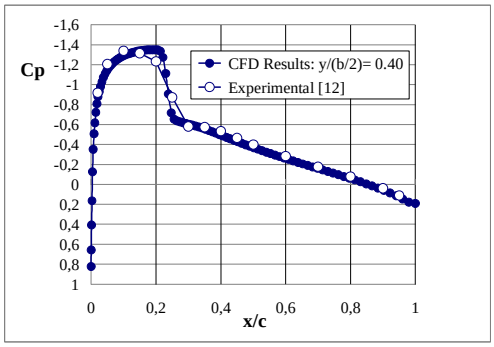


Figure 6 Straight wing model: Cp distribution on the upper surface of the section at 40% of the wing semi-span (α = 4°, Mach =0.7).

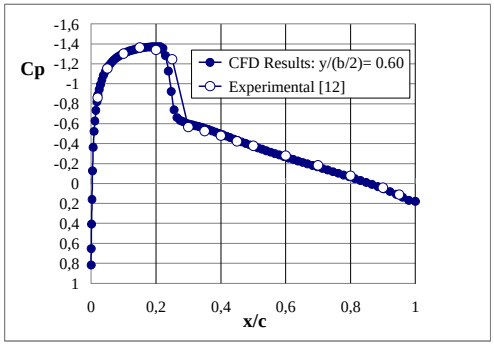


Figure 7 Straight wing model: Cp distribution on the upper surface of the section at 60% of the wing semi-span (α = 4°, Mach =0.7).

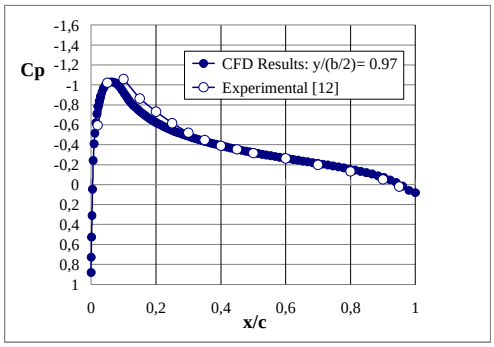


Figure 8 Straight wing model: Cp distribution on the upper surface of the section at 97% of the wing semi-span (α = 4°, Mach =0.7).

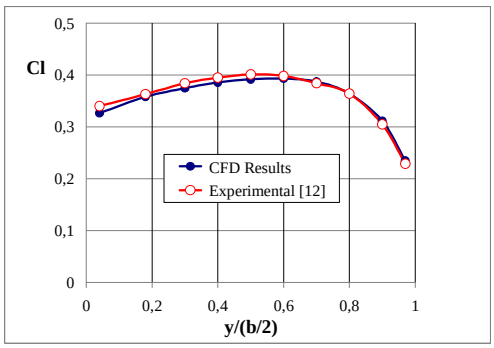


Figure 9 Comparison of the sectional lift coefficient for the straight wing model (α = 4°, Mach =0.7).

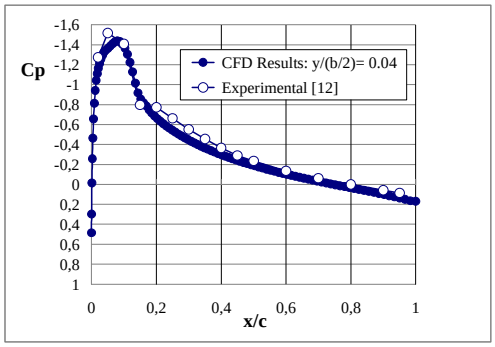


Figure 10 Forward swept wing model: Cp distribution on the upper surface of the section at 4% of the wing semi-span (α = 4°, Mach =0.7).

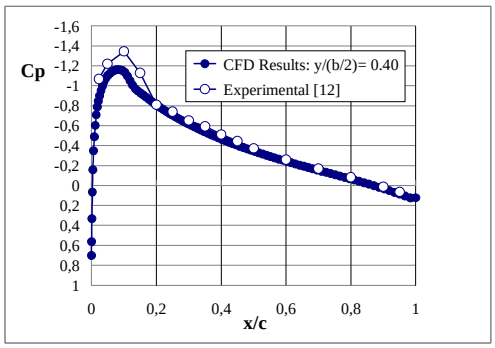


Figure 11 Forward swept wing model: Cp distribution on the upper surface of the section at 40% of the wing semi-span (α = 4°, Mach =0.7).

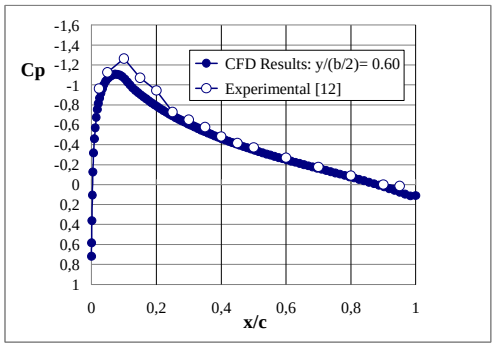


Figure 12 Forward swept wing model: Cp distribution on the upper surface of the section at 60% of the wing semi-span (α = 4°, Mach =0.7).

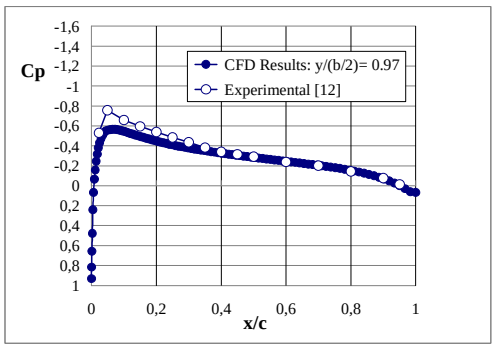


Figure 13 Forward swept wing model: Cp distribution on the upper surface of the section at 97% of the wing semi-span (α = 4°, Mach =0.7).

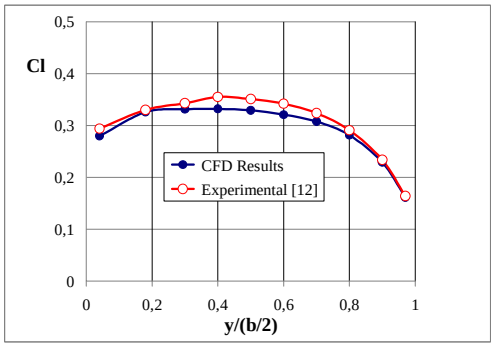


Figure 14 Comparison of the sectional lift coefficient for the swept wing model (α = 4°, Mach =0.7).

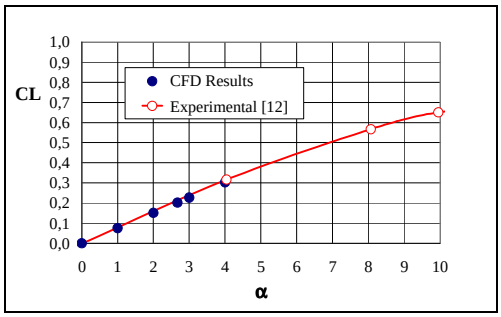


Figure 15 Comparison of the lift curves for the forward swept wing model (Mach = 0.7).

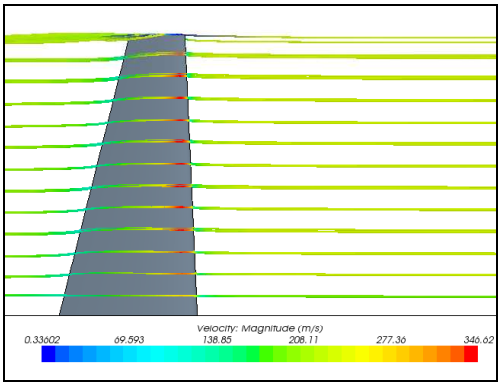


Figure 16 Streamlines on the upper surface of the straight wing model (Mach = 0.7, CL = 0.2, α = 2.23°)

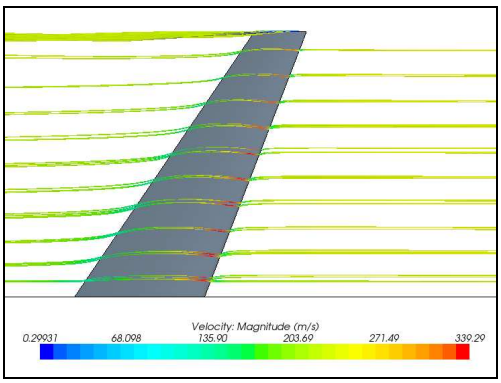


Figure 17 Streamlines on the upper surface of the swept wing model (Mach = 0.7, CL = 0.2, α = 2.67°)

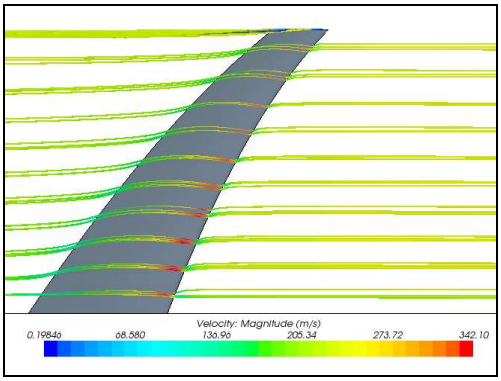


Figure 18 Streamlines on the upper surface of the curved wing model (Mach = 0.7, CL = 0.2, α = 2.94°)

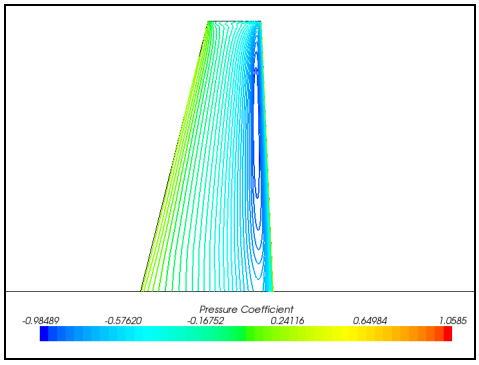


Figure 19 Isobars on the upper surface of the straight wing model (Mach = 0.7, CL = 0.2, α = 2.23°)

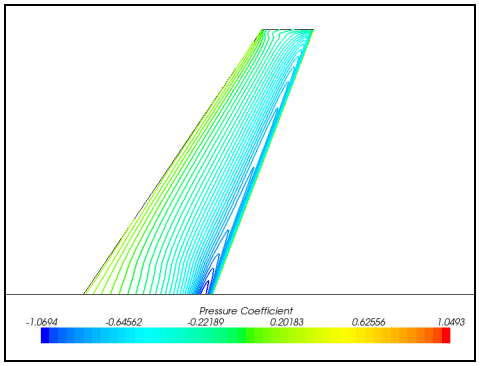


Figure 20 Isobars on the upper surface of the swept wing model (Mach = 0.7, CL = 0.2, α = 2.67°)

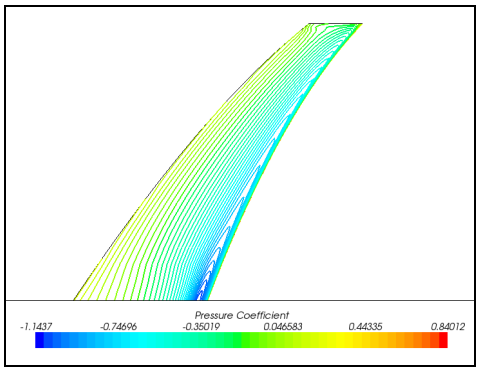


Figure 21 Isobars on the upper surface of the curved wing model (Mach = 0.7, CL = 0.2, α = 2.94°)

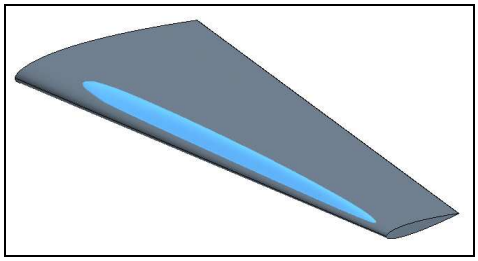


Figure 22 Extension of the Supersonic Region in the fluid flowing on the upper surface of the straight wing (Mach = 0.7, CL = 0.2, α = 2.23°)

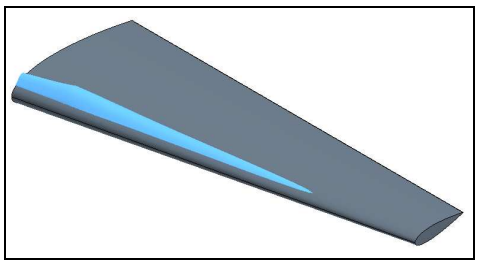


Figure 23 Extension of the Supersonic Region in the fluid flowing on the upper surface of the swept wing (Mach = 0.7, CL = 0.2, α = 2.67°)

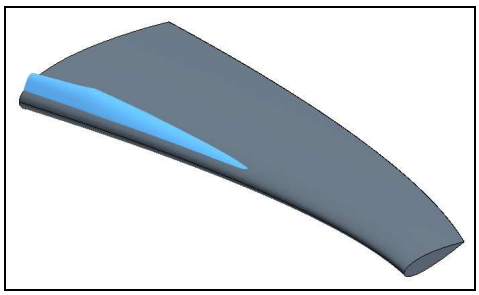


Figure 24 Extension of the Supersonic Region in the fluid flowing on the upper surface of the curved wing (Mach = 0.7, CL = 0.2, α = 2.94°)

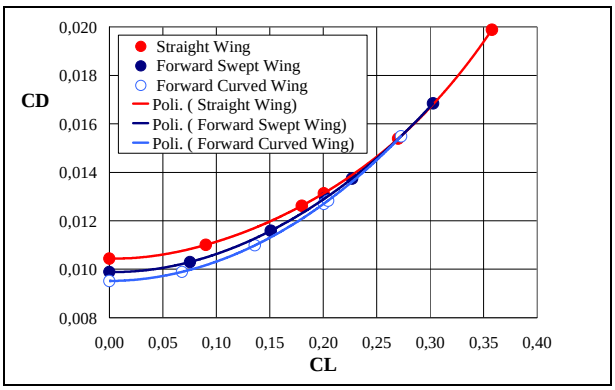


Figure 25 Drag Polar Curves of Wings (Mach = 0.7)

Table 2 Components of the Drag Coefficient – *Straight Wing Model* (Mach = 0.7, CL = 0)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	1.0	0.0	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-6.727e-04	2.991e-05	-6.428e-04
Upper Surface	2.405e-03	3.132e-03	5.536e-03
Tip Surface	0.0	1.136e-05	1.136e-05
Lower Surface	2.405e-03	3.132e-03	5.537e-03
Total	**4.137e-03**	**6.305e-03**	**1.044e-02**

Table 3 Components of the Drag Coefficient – *Swept Wing Model* (Mach = 0.7, CL = 0)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	1.0	0.0	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-4.943e-04	2.673e-05	-4.676e-04
Upper Surface	2.064e-03	3.110e-03	5.174e-03
Tip Surface	0.0	8.157e-06	8.157e-06
Lower Surface	2.064e-03	3.110e-03	5.174e-03
Total	**3.634e-03**	**6.255e-03**	**9.889e-03**

Table 4 Components of the Drag Coefficient – *Curved Wing Model* (Mach = 0.7, CL = 0)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	1.0	0.0	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-4.386e-04	2.290e-05	-4.157e-04
Upper Surface	1.870e-03	3.091e-03	4.961e-03
Tip Surface	0.0	8.299e-06	8.299e-06
Lower Surface	1.870e-03	3.091e-03	4.961e-03
Total	**3.302e-03**	**6.213e-03**	**9.515e-03**

Table 5 Components of the Drag Coefficient – *Straight Wing Model* (Mach = 0.7, CL = 0.2, $\alpha = 2.23°$)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	9.992e-01	3.886e-02	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-6.678e-04	2.861e-05	-6.392e-04
Upper Surface	-3.556e-03	3.150e-03	-4.060e-04
Tip Surface	0.0	1.128e-05	1.128e-05
Lower Surface	1.110e-02	3.072e-03	1.417e-02
Total	**6.875e-03**	**6.262e-03**	**1.314e-02**

Table 6 Components of the Drag Coefficient – *Swept Wing Model* (Mach = 0.7, CL = 0.2, $\alpha = 2.67°$)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	9.989e-01	4.662e-02	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-4.869e-04	2.401e-05	-4.629e-04
Upper Surface	-4.728e-03	3.105e-03	-1.623e-03
Tip Surface	0.0	7.958e-06	7.957e-06
Lower Surface	1.195e-02	3.060e-03	1.501e-02
Total	**6.738e-03**	**6.197e-03**	**1.293e-02**

Table 7 Components of the Drag Coefficient – *Curved Wing Model* (Mach = 0.7, CL = 0.2, $\alpha = 2.94°$)

	Chord Direction	Thickness Direction	Span Direction
Direction Vector	9.987e-01	5.130e-02	0.0
Components of the Drag Coefficient			
	Pressure	Shear	Net
Trailing Edge	-4.286e-04	1.850e-05	-4.100e-04
Upper Surface	-5.107e-03	3.067e-03	-2.040e-03
Tip Surface	0.0	8.061e-06	8.061e-06
Lower Surface	1.210e-02	3.050e-03	1.514e-02
Total	**6.557e-03**	**6.144e-03**	**1.270e-02**

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Transient Spoiler Aerodynamics using a conservative adaptive-grid ALE scheme

Alberto Guardone, Dario Isola and Giuseppe Quaranta
Dipartimento di Ingegneria Aerospaziale. Politecnico di Milano, Italy.

Keywords: ALE Formulation, Unsteady Aerodynamics, Spoiler Opening

Abstract

Two dimensional computations on moving control surfaces are presented. A novel numerical scheme for solving the compressible Euler equations on dynamic meshes with adaptation is used to simulate the opening of an airfoil spoiler. At each time level the boundary geometries are updated and a new mesh is obtained by a suitable combination of "elastic" deformation, edge-swapping, grid refinement and coarsening allowing to preserve the initial grid spacing. The Arbitrary Lagrangian-Eulerian formulation of the Euler equations is used to perform computations on dynamic grids without resorting to any explict interpolation phase. The grid modifications caused by the changes in topology due to edge-swapping or the insertion/removal are interpreted as continuous (in time) deformations of suitably defined node-centred finite volumes. The solution on the new grid is therefore computed simply by integrating of the governing equations. Preliminary computations of the dynamic loads on the airfoil due to the opening of the spoiler, in which the flow-field is characterized by unsteady flow separation at the movable surface trailing edge, are presented and compare fairly well with the reference experimental and numerical results.

1 Introduction

The correct investigation of hinge moments during the opening of control surfaces such as aileron, high-lift devices and spoilers is of the utmost importance in the preliminary design of aircraft, not only for the overall aerodynamic behaviour, but also for the general aircraft sizing. The adoption of high (Computational Fluid Dynamics) CFD models can be beneficial to the design of such aerodynamic appendicies and may be required for the correct assessment of the unsteady aerodynamic loads when movable surfaces are used as means for aeroelastic control systems. [1]

In the present paper the problem of the deployment of a spoiler from an airfoil is studied. Due to the large displacements of the boundaries, that characterize these problems, suitable techniques must be adopted to preserve the grid quality during the computaitons. Indeed the computational domain is continuously changing its shape and a new grid complying with the new geometry must be constructed at each time level. One popular approach to solve such problem are the so called over-set grids [3, 9] where the adoption of multiple overlapping meshes allows to overcome the difficulties arising from the complexity of the geometry.

Nodes displacement and local topological modifications (edge-swapping, node insertion and removal) are adopted to achieve large displacement of the boundaries. The overall grid

quality is fairly well maintained and the grid spacing is controlled such that the size of the elements is proportional to the distance from the solid bodies. As a result the computational mesh is moving during the simulation, so the standard finite volume Eulerian formulation of the flow equations must be dropped in favor of the Arbitrary LagrangianEulerian (ALE) approach, in which the control volumes are allowed to change in shape and position as time evolves.

In the present work the Euler equation in the ALE framework are discretized with a node centered finite volumes scheme. [4] The interpretation of grid modifications as continuous deformation, first introduced by Muffo et al. [7], allows to compute the solution at the current time level by employing standard ALE techniques. The same approach is here extended to the case of node insertion and deletion. Moreover, high order time integration schemes, like standard BDF techniques and others, can be implemented very easily. This results in a novel strategy to the simulation of unsteady aerodynamics with moving domain with variable connectivity and number of nodes, without explicit interpolation of the solution.

2 SPOILER DEPLOYMENT

The opening phase of a spoiler located on the on the upper surface of an airfoil is considered. Such control surfaces are used to decrease lift and increase drag in order to control the aircraft roll or as air-brakes. Complex chimera techniques have been adopted to compute the steady flow past three dimensional wings with open spoilers [6, 2]. However, the time accurate dynamic simulation of the spoiler opening, even in two spatial dimensions it is a very challenging problem from the numerical standpoint. [9]. Such kind of simulation can be of interest, since spoilers are considered as possible device for vibrations and flutter control [5] or separation control. [8]

Preliminary computations of the fast deployment of a spoiler from a NACA0012 are presented in order to assest the validity of the proposed approach. For this test case the airfoil is modified to insert a spoiler on the upper side. The location of the leading edge of the spoiler is $0.7c$, where c is the airfoil chord, and the length is $0.1c$. The upper surface of the spoiler is simply defined as the upper part of the NACA0012 at the same chordwise location. A $0.0015c$ gap between the spoiler and the airfoil main body is present.

The steady state, corresponding to an asymptotic Mach number of 0.25, is first computed. As shown in Figure 3(a), the flowfield is reasonably symmetric despite the presence of the small gap. The control surface is then set in motion applying an upward rotation $\Delta\theta = 90$ deg around the hinge, located at $0.7c$, with the following smooth time law

$$\theta = \frac{\Delta\theta}{2}\left[1 - \cos\left(\frac{\pi}{T_{deploy}}\min(t, T_{deploy})\right)\right]$$

where $T_{deploy} = 4.4$. The evolution of the computational grid is shown in Figure 3(left) and the grid at each time steps is updated using the adaption scheme already presented. It must be noted that the opening of the spoiler causes a significant changes in the topology of the mesh. The results in terms of streamlines and pressure coefficient values in the flowfield are shown in Figure 3(right). The development of a main vortex structure plus two secondary vortices behind the spoiler can be observed. It must be stressed that in the present inviscid simulations the separation point is fixed by the sharp end of the spoiler.

In figure 1 the lift coefficient C_L versus the non dimensional time t/T_{deploy} si shown. The maximum value of lift coefficient is obtained for $t/T_{deploy} \sim 0.5$ which agrees fairly well with the reference Navier-Stokes computations [9] and the experimental data [10]. The value of the maximum lift coefficient however differs from both the numerical and the experimental references. The C_L curve shows a rapid decay for $t/T_{deploy} > 0.5$ without the oscillatory behaviour that can be observed in the references.

As shown by this preliminary simulation the adopted strategy seems promising and it is capable to capture the most important flow features. More accurate results, with solution driven adaptation, are currently under study

and a better agreement with the reference is expected.

3 FINAL REMARKS

A novel flow solver for unsteady compressible flows with dynamic adaptive meshes was presented. The proposed approach is based on the interpretation of simple mesh adaptation techniques, namely, edge swapping, node insertion and node deletion, in terms of the continuous modification (in time) of suitably defined finite volumes surrounding the nodes of the triangulation. Therefore, the updated value of the unknowns at the grid nodes is automatically recovered by using an Arbitrary Lagrangian Eulerian (ALE) description of the flow equations, without resorting to an explicit interpolation step.

As a preliminary result, the numerical simulation of the opening process of a spoiler was carried out, to test the capabilities of both the grid alteration algorithm and the flow solver. These preliminary results increase the confidence of the authors on the possibility of increasing the agreement with the references adopting suitable solution driven adaptation techniques that are currently under study.

References

[1] L. Cavagna, S. Ricci, and A. Scotti. Active aeroelastic control over a four control surface wind tunnel wing model. In *International Forum on Aeroelasticity and Structural Dynamics IFASD-2007*, 2007.

[2] G. Fillola, M. Le Pape, and M. Montagnac. Numerical simulations around wing control surfaces. In *24th ICAS International Congress of the aeronautical sciences*, 2004.

[3] O. Hassan, K. A. Sørensen, K. Morgan, and N. P. Weatherill. A method for time accurate turbulent compressible fluid flow simulation with moving boundary components employing local remeshing. *International Journal for Numerical Methods in Fluids*, 53(8):1243–1266, 2007.

[4] D. Isola, A. Guardone, and G. Quaranta. An ale scheme without interpolation for moving domain with adaptive grids. In *40th Fluid Dynamics Conference and Exhibit, AIAA*, 2010.

[5] J. Rivera Jr, B. Dansberry, R. Bennett, M. Durham, and W. Silva. Naca 0012 benchmark model experimental flutter results with unsteady pressure distributions. In *33rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, volume No. AIAA Paper 92-2396, 1992.

[6] R. Mertins, E. Elsholz, S. Barakat, and B. Colak. 3d viscous flow analysis on wing-body-aileron-spoiler configurations. *Aerospace Science and Technology*, 9(6):476–484, 2005.

[7] D. Muffo, G. Quaranta, and A. Guardone. Compressilbe fluid-flow ale formulation on changing topology meshes for aeroelastic simulations. In *Proceedings of the* 26th *ICAS Congress*, Anchorage, Alaska, September 14–19 2008. Sumbitted.

[8] M. Schatz, T. Knacke, F. Thiele, R. Meyer, W. Hage, and D. Bechert. Separation control by self-activated movable flaps. In *42nd AIAA Aerospace Sciences Meeting and Exhibit*, pages 5–8, 2004.

[9] C. P. van Dam, R. Chow, J. R. Zayas, and D. E. Berg. Computational investigations of small deploying tabs and flaps for aerodynamic load control. *Journal of Physics: Conference Series*, 75(1):012027, 2007.

[10] W. W. H. Yeung, C. Xu, and W. Gu. Reduction of transient adverse effects of spoilers. *Journal of Aircraft*, 34(4):479–484, July-August 1997.

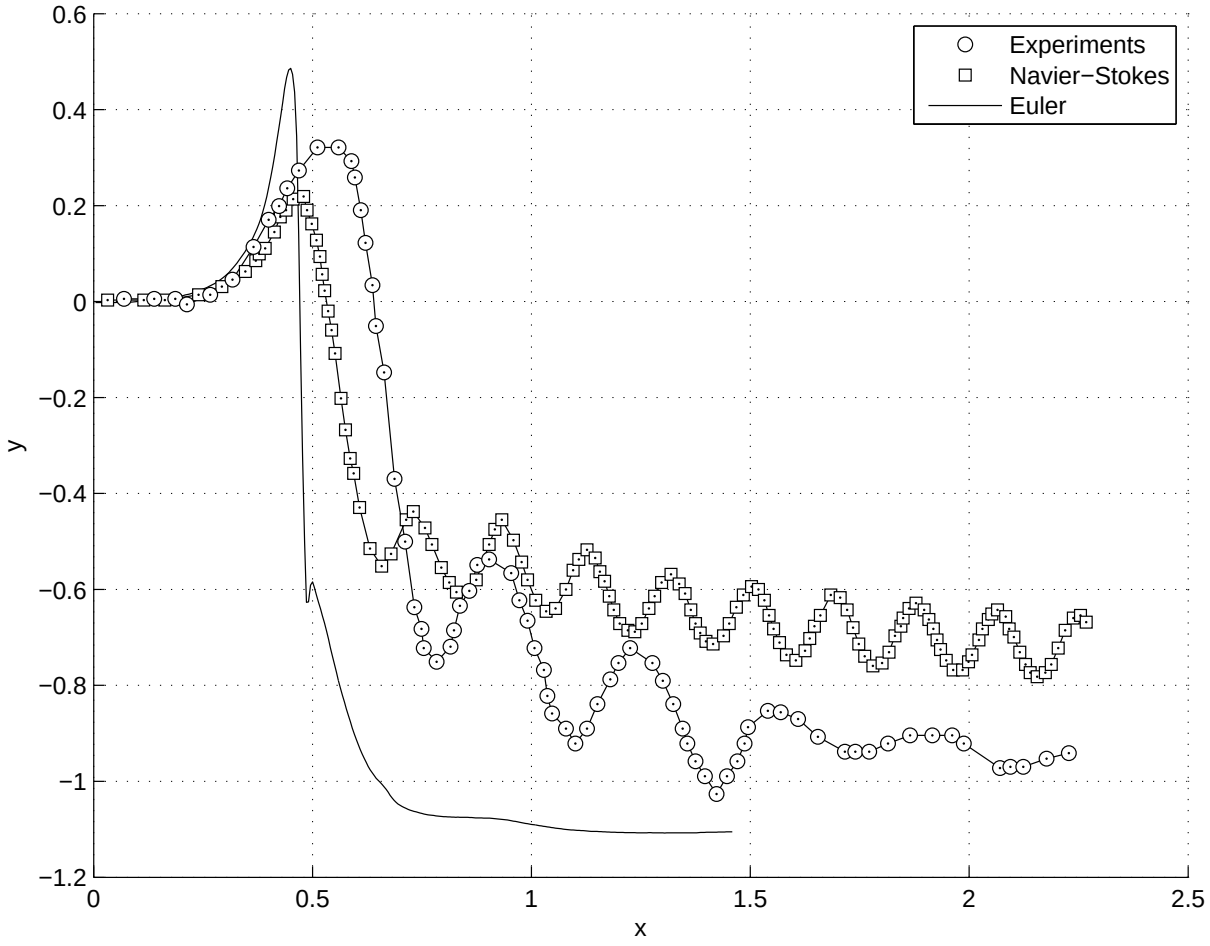
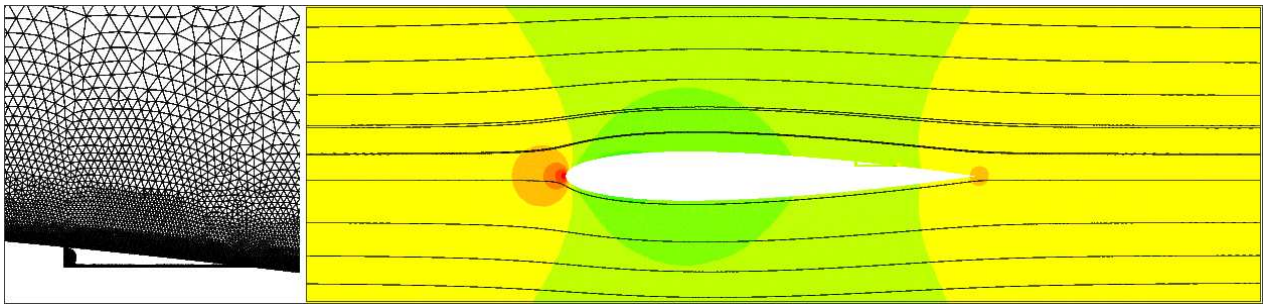
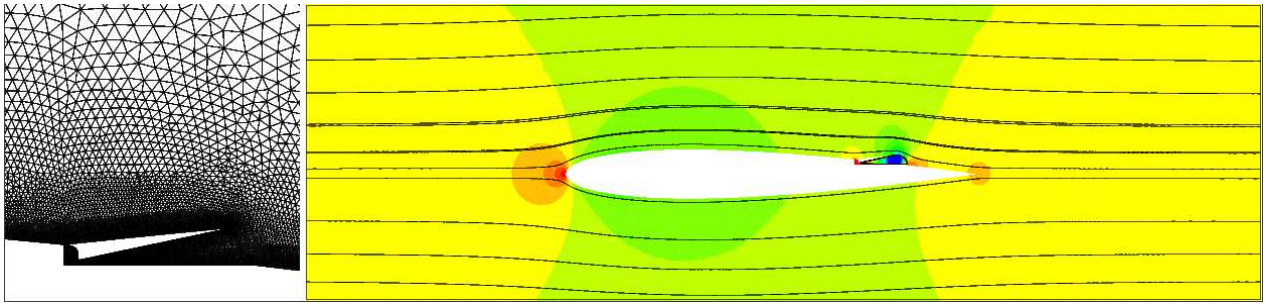


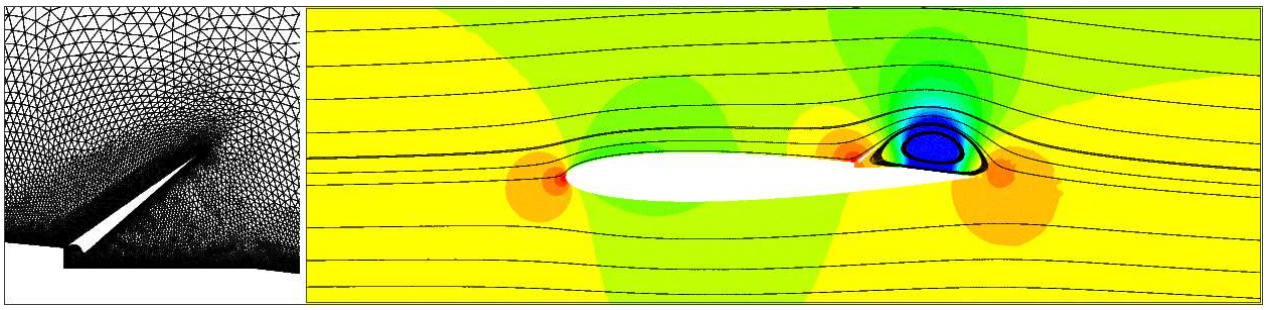
Figure 1: Lift coefficient versus the non dimensional time. References by Chow et al. [9] and Yeung et al.[10]



(a) $i = 0$, $t/t_{ref} = 0$

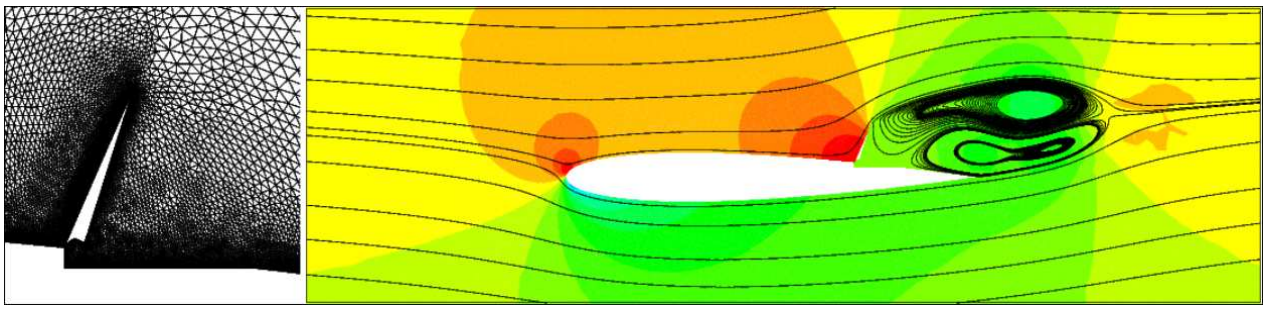


(b) $i = 50$, $t/t_{ref} = 0.25$

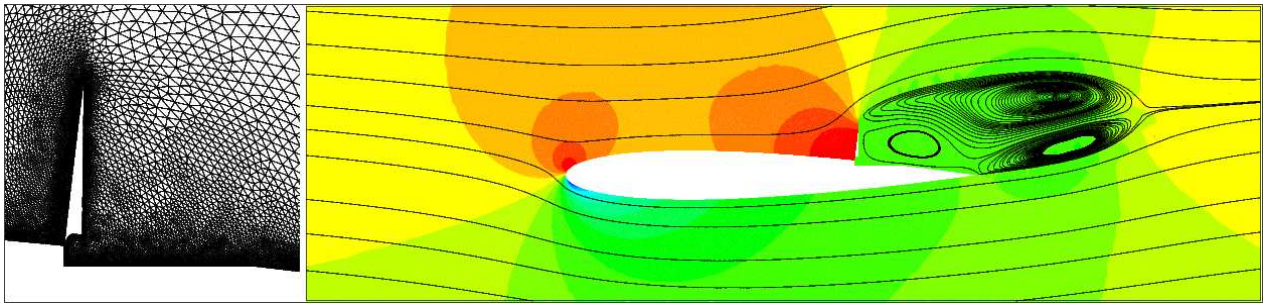


(c) $i = 100$, $t/t_{ref} = 0.5$

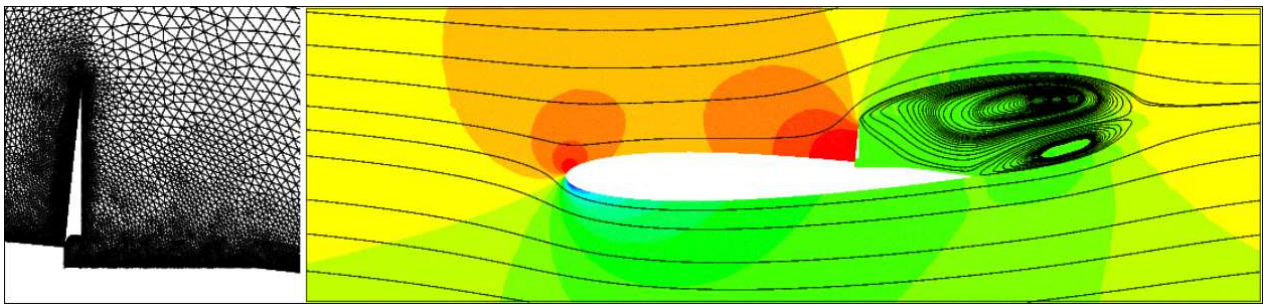
Figure 2: Pressure coefficient around the modified NACA0012 airfoil and close up of the mesh. (a) The initial C_p distribution on the the lower and upper side of the airfoil are very similar despite the small gap. The pressure inside the narrow slit is constant, and the air is essentially still. (b) When the spoiler is deployed the variations in the pressure field are mostly located in the proximites of the spiler. (c) The C_p increases slightly on the bottom side of the airfoil main body while decreases on the upper side, therefore the stagnation point is shifted towards the lower side. A strong vortex is trapped between the spoiler and the trailing edge. The resulting lift coefficient is positive.



(a) $i = 150$, $t/t_{ref} = 0.75$



(b) $i = 200$, $t/t_{ref} = 1$



(c) $i = 300$, $t/t_{ref} = 1.5$

Figure 3: Pressure coefficient around the modified NACA0012 airfoil and close up of the mesh. Cont. (a) The pressure coefficient on the upper side of the main body of the airfoil is strongly increased while it drops on the lower side. A constant C_p region is located past the spoiler that is almost completely deployed. (b-c) The movement is completed and the overall forces on the body are positive downward. A large separation region is located past the spoiler and the airfoil.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Prediction of Aerodynamic Propeller/Wing Interaction with a Panel Code Using Embedded Actuator Disk Model

G. Droandi and G. Gibertini
Dipartimento di Ingegneria Aerospaziale – Politecnico di Milano
Via La Masa 34, 20156, Milano, Italy

e–mail: droandi@aero.polimi.it and giuseppe.gibertini@polimi.it

Keywords: Actuator disk, Propeller–Wing interaction, Panel code, Velocity potential

Abstract

An inviscid actuator disk model is embedded in a three-dimensional low-order panel method for inviscid incompressible flow in order to study the aerodynamic interaction between propeller and wing. The actuator disk model predicts the time–averaged induced velocities in the slipstream of a propeller with an arbitrary radial distribution of load. The model is constructed by superposition of four vorticity distributions, by neglecting the radial contraction of the vortex tube and by supposing a fixed wake for the propeller. Two different experiments have been used to validate the actuator disk model embedded in the panel code.

1 Introduction

The problem of calculating aerodynamic loads developing on arbitrary three dimensional bodies is of high interest in aeronautical industries. In this context, the calculation of aerodynamic loads involves the knowledge of velocity fields about the aircraft, and a wide range of numerical methods are nowadays available to solve the flow field around it. The problem of calculating the velocity fields becomes more complex when a propeller is taken into account, because it is necessary to compute the propeller's influence on the aircraft. The aerodynamic interaction between propeller and wing can be calculated using a boundary integral panel method incorporating some kind of propeller model. Since boundary integral methods give good numerical results in short (computational) time, they are often used in the initial design of aircrafts. In these cases, only some properties of the propeller are known, therefore it is necessary to combine a panel method with a propeller actuator disk that does not require a detailed panel model of the propeller. In the present paper, an analytical solution for the actuator disk proposed by Conway in [1] is embedded in COMPA (COMPlet Aircraft aerodynamic computation), that is a three-dimensional low-order panel code written and developed at Politecnico di Milano. The method described by Conway is based on the construction of the velocity fields induced by a ring vortex as integrals over the allowed values of the separation constant of the eigensolutions of the Laplace's equation in cylindrical coordinates. The ring vortex solutions are combined to give the solutions for general actuator disks. Integral representation for velocities can be obtained in terms of elliptic integrals. The integral representations of induced velocity components in cylindrical coordinates can always be evaluated thanks to a polynomial expansion approx-

imating the radial distribution of the axial velocity on the actuator disk plane. The actuator disk model is coupled to COMPA, hence when a propeller is taken into account, the velocity field induced by it is introduced in the solving equation as a modification of the free stream (as shown by Conway and Su in [2]). In order to validate the actuator disk model that has been implemented, two different experimental studies on interaction effects between propeller and wing are taken into account: the first one by Stuper [3], and the second one by Chiaramonte, Favier, Maresca and Agnes [4] and [5].

2 Actuator Disk Theory

It is well known that for an arbitrary unsteady incompressible flow the istantaneous velocity fields are linearly related to the istantaneous vorticity distribution thanks to the Biot–Savart law. Therefore the steady flow induced by the time–average of the vorticity distribution gives the time–average of the velocity fields. From a physical point of view, a rotating propeller develops a distribution of forces along each blade that depends on blade geometry, angular velocity and flow field conditions. These forces distribution can be related to the circulation by the Kutta-Joukowski theorem. The circulation jump due to the rotating propeller may be expressed by straight lifting lines which lie in the disk plane. Hence, the time–averaged velocity field induced by a rotating propeller can be correctly represented by an actuator disk model (Conway [1]). Therefore, in order to obtain a general three–dimensional solution for rotating propeller, an actuator disk model can be constructed by superposition of some appropriate vorticity distributions. As shown by Conway in [1], four different vortex distributions can be used:

1 a ring vortex system (from the disk to downstream infinity);

2 a hub vortex of constant strength (from the disk centre to downstream infinity);

3 a radial vortex system on the disk plane;

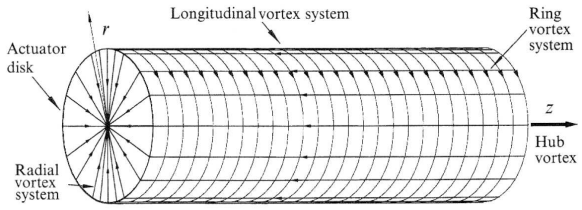


Figure 1: Vortex systems for the actuator disk model.

4 a longitudinal vortex system normal to (1) and with the same strength of (2).

The four vorticity distributions involved in the actuator disk model are shown in Fig. 1. In the linear limit, by neglecting the radial contraction of the vortex tube and by taking into account an arbitrary radial loading, a general solution may be written in terms of velocity components (or, optionally, velocity potential). In cylindrical polar coordinates (r, ϕ, z), the azimuthal velocity component $V_\phi(r, z)$ is found to be:

$$\begin{cases} V_\phi(r,z) = \dfrac{\Gamma_s(r)}{2\pi r}, & r < R, \\ V_\phi(r,z) = 0, & r > R, \end{cases} \tag{1}$$

where R is the external propeller radius. The axial and radial velocity components ($V_r(r, z)$ and $V_z(r, z)$) can be obtained by the superposition of a radial distribution of vortex tubes of constant radius and surface ring vortex density $\gamma(r)$. Starting from the Biot–Savart law, the integral forms of axial and radial velocity components, written in terms of Bessel functions, are:

$$\begin{aligned} V_z(r,z) = & \frac{1}{2}\int_0^R \int_0^\infty \gamma(\hat{r}) \left(2 - e^{-s|z|}\right) \\ & \times \hat{r} J_1(s\hat{r}) J_0(sr)\, ds d\hat{r}, \\ & z \geq 0, \end{aligned} \tag{2}$$

$$\begin{aligned} V_z(r,z) = & \frac{1}{2}\int_0^R \int_0^\infty \gamma(\hat{r})\, e^{-s|z|} \\ & \times \hat{r} J_1(s\hat{r}) J_0(sr)\, ds d\hat{r}, \\ & z < 0, \end{aligned} \tag{3}$$

$$\begin{aligned} V_r(r,z) = & -\frac{1}{2}\int_0^R \int_0^\infty \gamma(\hat{r})\, e^{-s|z|} \\ & \times \hat{r} J_1(s\hat{r}) J_1(sr)\, ds d\hat{r}. \end{aligned} \tag{4}$$

It can be observed that in the actuator disk plane, for $z = 0$, Eq. 2 and Eq. 3 give the same expression for the axial velocity component. Since the strength of vortex tubes $\gamma(r)$ is zero for $r > R$, the limit of radial integration can be replaced by ∞. By taking the first–order Hankel transform of $\gamma(\hat{r})$, the strength of vortex tubes $\gamma(r)$ can be related to the axial velocity distribution at the actuator disk by the following relation:

$$\frac{dV_z\left(r,0\right)}{dr} = -\frac{\gamma\left(r\right)}{2}. \tag{5}$$

Thanks to Eq. 5, Eq. from 2 to 4 become:

$$\begin{aligned} V_z\left(r,z\right) = &-\int_0^\infty \int_0^\infty \frac{dV_z\left(\hat{r},0\right)}{d\hat{r}} \left(2 - e^{-s|z|}\right) \\ &\times \hat{r} J_1\left(s\hat{r}\right) J_0\left(sr\right) ds d\hat{r}, \\ &z \geq 0, \end{aligned} \tag{6}$$

$$\begin{aligned} V_z\left(r,z\right) = &-\int_0^\infty \int_0^\infty \frac{dV_z\left(\hat{r},0\right)}{d\hat{r}} e^{-s|z|} \\ &\times \hat{r} J_1\left(s\hat{r}\right) J_0\left(sr\right) ds d\hat{r}, \\ &z < 0, \end{aligned} \tag{7}$$

$$\begin{aligned} V_r\left(r,z\right) = &\int_0^\infty \int_0^\infty \frac{dV_z\left(\hat{r},0\right)}{d\hat{r}} e^{-s|z|} \\ &\times \hat{r} J_1\left(s\hat{r}\right) J_1\left(sr\right) ds d\hat{r}. \end{aligned} \tag{8}$$

Therefore, Eq. from 6 to 8 can be integrated by parts and Bessel identities and some properties of Henkel transforms may be applied to obtain useful expressions for axial and radial velocity components (for an exhaustive explanation see Conway [1] and [6]):

$$\begin{aligned} V_z\left(r,z\right) = &2V_z\left(r,0\right) - \int_0^{R_a} \int_0^\infty V_z\left(\hat{r},0\right) \\ &\times e^{-s|z|} s\hat{r} J_0\left(s\hat{r}\right) J_0\left(sr\right) ds d\hat{r}, \\ &z \geq 0, \end{aligned} \tag{9}$$

$$\begin{aligned} V_z\left(r,z\right) = &\int_0^{R_a} \int_0^\infty V_z\left(\hat{r},0\right) \\ &\times e^{-s|z|} s\hat{r} J_0\left(s\hat{r}\right) J_0\left(sr\right) ds d\hat{r}, \\ &z < 0, \end{aligned} \tag{10}$$

$$\begin{aligned} V_r\left(r,z\right) = &-\int_0^{R_a} \int_0^\infty V_z\left(\hat{r},0\right) \\ &\times e^{-s|z|} s\hat{r} J_0\left(s\hat{r}\right) J_1\left(sr\right) ds d\hat{r}. \end{aligned} \tag{11}$$

Eq. from 9 to 11 can be easily solved if the radial distribution of axial velocity at the actuator disk can be approximated with an even polynomial distribution, and closed–form solution can be obtained, such as in the case of an elliptic distribution. When the axial velocity distribution at the disk can be expressed by a general polynomial distribution, such as:

$$V_z\left(r,0\right) = \sum_{n=0}^{N} A_n r^n, \tag{12}$$

the solution for axial and radial velocity components can be given in terms of elementary functions that can be evaluated numerically. In the present work, the general case is analyzed and a modification of the method given by Conway in [1] to solve the problem numerically is presented.

We first consider the axial component of the velocity $V_z(r,z)$. The integral expression that appears in Eq. 9 and 10 can be written in the following form:

$$\begin{aligned} I^z\left(r,z\right) = \mp \frac{\partial}{\partial z} \int_0^{R_a} \int_0^\infty V_z\left(\hat{r},0\right) e^{-s|z|} \\ \times \hat{r} J_0\left(s\hat{r}\right) J_0\left(sr\right) ds d\hat{r}. \end{aligned} \tag{13}$$

By performing the s integration, substituting Eq. 12 in Eq. 13 and writing in explicit form the Legendre function that results from the integration (see Gradshteyn and Ryzhik [7] and Abramowitz and Stegun [8]), we obtain:

$$\begin{aligned} &I^z\left(r,z\right) = \\ &\frac{|\,z\,|}{\pi} \int_0^\pi \sum_{n=0}^{N} A_n \int_0^{R_a} \frac{\hat{r}^{n+1}}{\left(X\left(\hat{r},t,r,z\right)\right)^{3/2}} d\hat{r} dt, \end{aligned} \tag{14}$$

where, for simplicity, we have defined the quantity $X(\hat{r},t,r,z) = \hat{r}^2 + r^2 + z^2 - 2r\hat{r}\cos t)$. Conway [1] suggests to calculate the indefinite radial integral $I_n^z(\hat{r},t,r,z)$ that appears in Eq. 14 exploiting the recursive relation shown by Gradshteyn and Ryzhik [7] and Morita [9]. The correct expression of the recursive relation is the following:

$$
\begin{aligned}
&I_n^z\left(\hat{r},t,r,z\right)= \\
&+\frac{\hat{r}^n}{\left(n-1\right)\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}} \\
&+\frac{\left(2n-1\right)r\cos t}{n-1}I_{n-1}^z\left(\hat{r},t,r,z\right) \\
&-\frac{n\left(r^2+z^2\right)}{n-1}I_{n-2}^z\left(\hat{r},t,r,z\right).
\end{aligned}
\tag{15}
$$

To start the recursive process, it is necessary to give the correct initial values to the starting integrals (for $n=0$ and $n=1$). Following Gradshteyn and Ryzhik [7], we have:

$$
\begin{aligned}
&I_0^z\left(\hat{r},t,r,z\right)= \\
&-\frac{r^2+z^2-r\hat{r}\cos t}{\left(r^2\sin^2 t+z^2\right)\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}},
\end{aligned}
\tag{16}
$$

$$
\begin{aligned}
&I_1^z\left(\hat{r},t,r,z\right)= \\
&-\frac{\left(z^2-r^2\cos 2t\right)\hat{r}+\left(r^2+z^2\right)r\cos t}{\left(r^2\sin^2 t+z^2\right)\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}} \\
&+\ln\left[\left(2\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}+\hat{r}-r\cos t\right)\right].
\end{aligned}
\tag{17}
$$

The integral $I^z(r,z)$ of Eq. 14 can then be written as follow:

$$
I^z\left(r,z\right)=\frac{\mid z\mid}{\pi}\int_0^{\pi}\left[\sum_{n=0}^{N}A_n\left(\Delta_n^z\left(t,r,z\right)\right)\right]dt,
\tag{18}
$$

where:

$$
\Delta_n^z\left(t,r,z\right)=I_n^z\left(R_a,t,r,z\right)-I_n^z\left(0,t,r,z\right).
\tag{19}
$$

From Eq. 18 and Eq. 19 it is clear that the degree of the polynomial approximating the radial distribution of axial velocity at the disk may affect the final value of the integral $I^z(r,z)$. In order to decrease this effect, we use the additivity of integration on the interval $[0,R_a]$ when we substitute the recursive formula of Eq. 15 in Eq. 14. By proceeding this way, the interval of integration $[0,R_a]$ can be divided into an arbitrary number of subintervals $[\hat{r}^{In},\hat{r}^{Out}]$ when the integration over $\hat{r}$ is performed. Moreover, for each subinterval, a polynomial approximating the corresponding piece of the function $V_z(r,0)$ has to be computed (note that the degrees of a polinomial computed for a given subinterval may be different from each other). When the procedure described above is applied to Eq. 14, a new summation over the index k appears inside the integral expression. This summation allows to evaluate the influence on the considered point (that is strictly related to the parameter r) of all the pieces in which the $V_z(r,0)$ is divided. Hence, the final formula for the integral $I^z(r,z)$ of Eq. 14 is:

$$
\begin{aligned}
&I^z\left(r,z\right)= \\
&\frac{\mid z\mid}{\pi}\int_0^{\pi}\left[\sum_{k=1}^{K}\sum_{n_k=0}^{N_k}A_{n_k,k}\Delta_{n_k,k}^z\left(t,r,z\right)\right]dt,
\end{aligned}
\tag{20}
$$

where:

$$
\begin{aligned}
&\Delta_{n_k,k}^z\left(t,r,z\right)= \\
&I_{n_k,k}^z\left(\hat{r}_k^{In},t,r,z\right)-I_{n_k,k}^z\left(\hat{r}_k^{Out},t,r,z\right).
\end{aligned}
\tag{21}
$$

The axial velocity component $V_z(r,z)$ is:

$$
V_z\left(r,z\right)=2V_z\left(r,0\right)-I^z\left(r,z\right),\quad z\geq 0,
\tag{22}
$$

$$
V_z\left(r,z\right)=I^z\left(r,z\right),\quad z<0.
\tag{23}
$$

The solution for the radial velocity component $V_r(r,z)$ can be deduced in a very similar way as the axial velocity. If we consider Eq. 8 and we calculate the s integral, we obtain a new integral relation where a second order Legendre function appears. Exploiting the Legendre function (see Gradshteyn and Ryzhik [7] and Abramowitz and Stegun [8]) Eq. 8 becomes:

$$
\begin{aligned}
&V_r\left(r,z\right)= \\
&\frac{1}{\pi}\sum_{n=1}^{N}\int_0^{R_a}\int_0^{\pi}\frac{nA_n\hat{r}^n\cos t}{\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}}dtd\hat{r}.
\end{aligned}
\tag{24}
$$

As before, the radial integral $I_n^r(\hat{r},t,r,z)$ that appears in Eq. 24 can be calculated with a recursive formula by Gradshteyn and Ryzhik [7] and Morita [9], that is:

$$
\begin{aligned}
&I_n^r\left(\hat{r},t,r,z\right)= \\
&+\frac{\hat{r}^{n-1}}{n}\left(X\left(\hat{r},t,r,z\right)\right)^{1/2} \\
&+\frac{\left(2n-1\right)r\cos t}{n}I_{n-1}^r\left(\hat{r},t,r,z\right) \\
&-\frac{\left(n-1\right)\left(r^2+z^2\right)}{n}I_{n-2}^r\left(\hat{r},t,r,z\right).
\end{aligned}
\tag{25}
$$

The starting values for the recursion are:

$$
\begin{aligned}
&I_0^r\left(\hat{r},t,r,z\right)= \\
&\ln\left[2\left(\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}+\hat{r}-r\cos t\right)\right],
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
&I_1^r\left(\hat{r},t,r,z\right)= \\
&\left(X\left(\hat{r},t,r,z\right)\right)^{1/2}+r\cos t I_0^r\left(\hat{r},t,r,z\right).
\end{aligned}
\tag{27}
$$

If we divide the interval of integration $[0,R_a]$ into an arbitrary number of subintervals, the final expression for the radial velocity component of Eq. 24 is:

$$
\begin{aligned}
&V_r\left(r,z\right)= \\
&\frac{1}{\pi}\int_0^{\pi}\cos t\left[\sum_{k=1}^{K}\sum_{n_k=1}^{N_k}n_k A_{n_k,k}\Delta_{n_k,k}^r\left(t,r,z\right)\right]dt,
\end{aligned}
\tag{28}
$$

where:

$$
\begin{aligned}
&\Delta_{n_k,k}^r\left(t,r,z\right)= \\
&I_{n_k,k}^r\left(\hat{r}_k^{In},t,r,z\right)-I_{n_k,k}^r\left(\hat{r}_k^{Out},t,r,z\right).
\end{aligned}
\tag{29}
$$

3 Actuator Disk And Panel Code

COMPA is a three-dimensional low-order velocity-based panel method for inviscid incompressible flow. The code is based on the classical Hess and Smith method [10] and [11]. This method has been modified following the idea proposed by Maskew and Woodward [12] to have a right distribution of singularities (sources and vortex rings) on lifting bodies. The code solves the Dirichlet problem for the velocity, requiring the equivalence between the velocity vector ($\mathbf{V}=\nabla\varphi$) and the free stream velocity vector ($\mathbf{V}_\infty$) in the normal direction ($\mathbf{n}$) on the body surfaces; the wake is considered rigid and aligned to the free stream velocity vector. Hence the boundary condition that must be satisfied is:

$$
\nabla\varphi\cdot\mathbf{n}=-\mathbf{V}_\infty\cdot\mathbf{n}. \tag{30}
$$

The perturbation potential φ can be written in the integral form, as shown by Hess and Smith [10], and bodies can be represented by distributions of sources and vortex rings. To solve the integral equation, the surfaces of the bodies are discretized into quadrilateral panels on which constant surfaces source (of strength σ) are specified. To represent a lifting body however, a correct number of vortex rings (of strength Γ) have to be placed on the mean line identified by its surface panels (this approach is described by Maskew and Woodward in [12] and is a little modification of the original Hess method). The discretized linear equation is:

$$
\begin{aligned}
&\sum_{k=1}^{N_s}B_{kj}\sigma_k+\sum_{i=1}^{N_v}A_{ij}\Gamma_i=-\mathbf{V}_\infty\cdot\mathbf{n}, \\
&j=1,N_s+N_v.
\end{aligned}
\tag{31}
$$

In Eq. 31, N_s is the number of surface panels on which sources are distributed, N_v is the number of panels of wake and bodies on which vortex rings are located (Kutta condition is imposed when a trailing panel is computed), B_{kj} and A_{kj} are the influence coefficient matrices representing the influence of source and vortex ring distribution on panel k at the collocation point j. Equation 31 represents a linear system of equations that can be easily solved with an iterative method; COMPA use alternatively a PCG (preconditioned conjugate gradient method) or a LU decomposition with partial pivoting and row interchanges.

In order to account for the velocity fields induced by the propeller and its associated wake, the inviscid flow induced by an exact actuator disk model for a non–uniform distribution of load along the blades is calculated with the method proposed in the previous section (Sec. 2). The actuator disk model has been obtained by the superposition of some appropriate vorticity distributions under the hypothesis that the

wake produced by the propeller is fixed (see Conway and Su [13]). For this reason, the velocity fields induced by the actuator disk on each control point of the surface panels of the body can be computed neglecting its presence. The coupling between COMPA and the actuator disk solution can then be done by introducing the velocity fields induced by the actuator disk as a perturbation of the free stream velocity vector (Conway and Su [2]). Hence Eq. 30 becomes:

$$\nabla\varphi \cdot \mathbf{n} = -\left(\mathbf{V}_\infty + \mathbf{V}_{AD}\right) \cdot \mathbf{n}, \qquad (32)$$

where $\mathbf{V}_{AD}$ is the velocity induced by the propeller in cartesian coordinates. It follows that the influence coefficient matrices B_{kj} and A_{kj} of Eq. 31 are not modified by introducing the actuator disk solution in Eq. 30.

4 Numerical Results

In order to validate the actuator disk model that has been implemented in COMPA, two different experimental studies on interaction effects between propeller and wing are taken into account: the first one by Stuper [3], and the second one by Chiaramonte, Favier, Maresca and Agnes [4] (and also [5] and [14]).

In the first case, the wind tunnel measurements were made on a wing of symmetrical section (Gottingen 409 airfoil) having a span of $0.8m$, a chord of $0.2m$ between two circular end–plate of radius $0.16m$. The free stream velocity for the measurements is about $30m/s$. A small high–speed electric motor placed in front of the wing drives a two blade propeller; the distance between the propeller and the leading edge of the wing is $0.075m$. The blade airfoil and the thrust coefficient ($C_t = T/(\rho_\infty A(\Omega R)^2)$) of the propeller are not given by Stuper. Since the propeller has a radius of $0.075m$ and its advance ratio is $J = V/(nD) = 0.47$, assuming that blades have a Clark-Y airfoil (for blade geometry, see Fig. 23 in [3]), the blade element momentum theory gives a computed C_t of 0.0038. Experimental results in terms of global wing lift coefficient and load distribution in the spanwise direction are given by Stuper without any corrections. However, end–plates and wind tunnel walls seem to have non–neglegible effects on the results. Therefore, to compare experimental data with numerical results, we decided to build the computational mesh only for motor and wing and to analyze a wing with an aspect ratio ($\lambda = 4.5$) greater than the one used in the experiments ($\lambda = 4$). Anyway only the portion of the wing that corresponds to the experimental model has been weighed. Moreover, it has been noted from experimental results that for higher angles of attack an abrupt change in the slope of $C_L - \alpha$ curve of the wing occurs. This phenomenon is related to viscous effects that become non–neglegible for $\alpha \geq 6°$. Since COMPA is a panel method for inviscid incompressible flow and no boundary layer calculation is performed to take into account viscous effects, the results presented here are related only to low angles of attack. Results in terms of the x component of the velocity in the weighed portion of the wing, for $\alpha = 4°$, are shown in Fig. 2 and Fig. 3, respectively without and with propeller (the motor is not shown). For this configuration the distributions of load on the wing in spanwise direction with and without propeller are reported in Fig. 4, in which numerical results are compared with experimental data (Stuper [3]). In terms of numerical results, the effect of the propeller on the wing is clear when comparing Fig. 2 and Fig. 3. Anyway, Fig. 4 shows that the code gives results that are slightly overrated than the experimental data. Comparison between experimental and numerical results shows that the actuator disk model embedded in COMPA gives good results in terms of load distribution in spanwise direction of the wing for low angles of attack. However, when the propeller is present, in some sections of the wing there are differences between experimental data and numerical results. These differences seem to be related to the actuator disk model embedded in COMPA and its hypotheses because the wake produced by the propeller is supposed to be fixed and the radial contraction of the vortex tube is neglected.

In the second case, experiments were conducted on a model scale of propeller/nacelle/wing ensemble mounted in the open chamber of a subsonic wind tunnel.

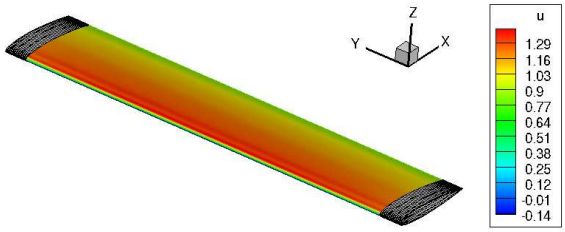


Figure 2: Distribution of V_x over the wing for $\alpha = 4°$ without propeller.

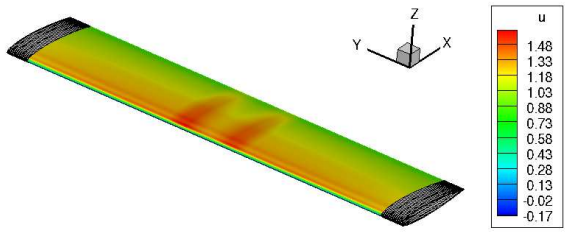


Figure 3: Distribution of V_x over the wing for $\alpha = 4°$ with propeller.

The wing has a span of $1.7m$ and a chord of $1.02m$; the nacelle is located in the middle of the wing. The propeller is located in front of the wing and the distance between the hub and its leading edge is $0.255m$ in the direction of the wing chord. The wing has a RA18–43N1L1 airfoil section of 17% thickness (see Fratello, Favier, Maresca and Barbi [5] for geometry details). The four blade propeller is powered by an electrical motor embedded in the nacelle. The four blades are cambered NACA64A408 airfoil sections of variable thickness and chord (the blade geometry is reported by Favier, Ettaouil and Maresca [14]). The radius of the propeller is $0.425m$. The operating conditions that we considered for the calculation are the following: wing angle of attack $\alpha = 0°$, free stream velocity $V_\infty = 17.2m/s$, propeller advance ratio $J = 0.89$, angular velocity $\Omega = 142.63Hz$ and pitch angle of the blade $\alpha_0 = 32.5°$ (referred to the section located at 75% of the blade). Since the nacelle shape is not well described neither in [4] nor in [5] by the authors (only approximate layouts are given), we decided to consider only the wing in our numerical simulations. In Fig. 5 the distribution of the x component of the velocity on the wing is reported, while in Fig. 6 the same quantity is shown when the propeller is considered.

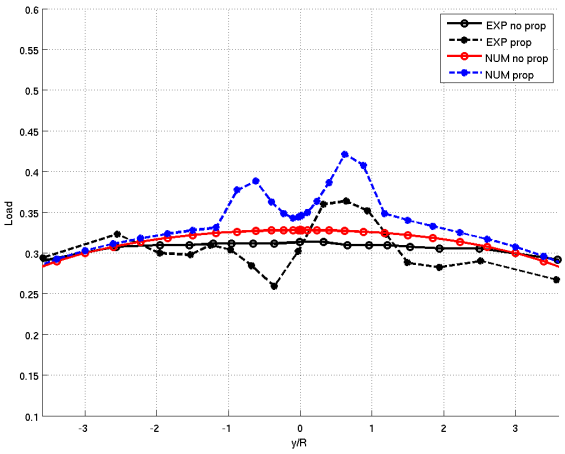


Figure 4: Span distribution of load over the wing for $\alpha = 4°$ with and without propeller. Comparison between experimental and numerical results.

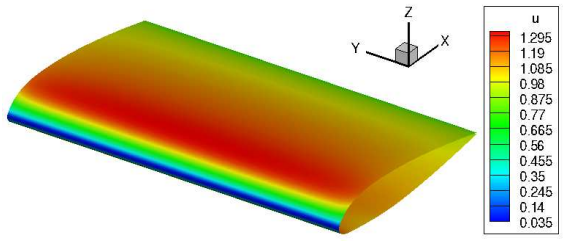


Figure 5: Distribution of V_x over the wing for $\alpha = 0°$ without propeller.

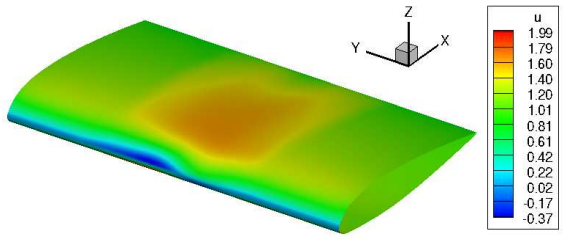


Figure 6: Distribution of V_x over the wing for $\alpha = 0°$ with propeller.

Figures from 7 to 14 give the comparison between the mean value of the measured pressure coefficient (averaged over the rotational period of the blades, Chiaramonte, Favier, Maresca and

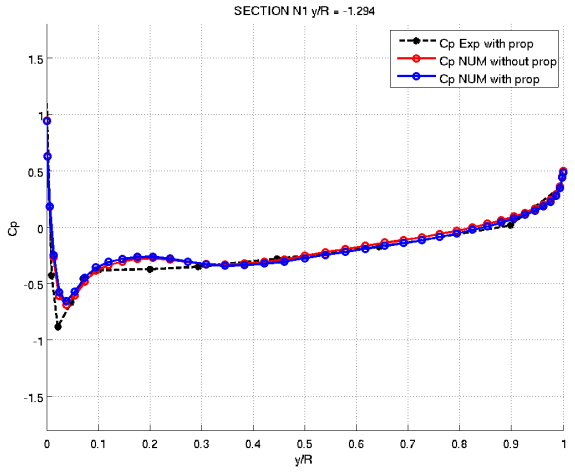


Figure 7: Mean pressure distribution on the lower side of the wing.

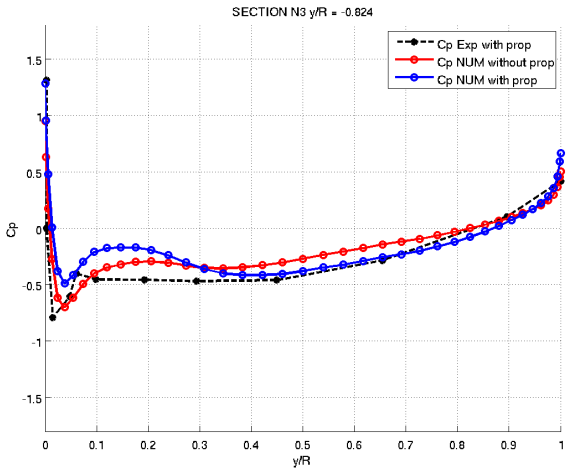


Figure 9: Mean pressure distribution on the lower side of the wing.

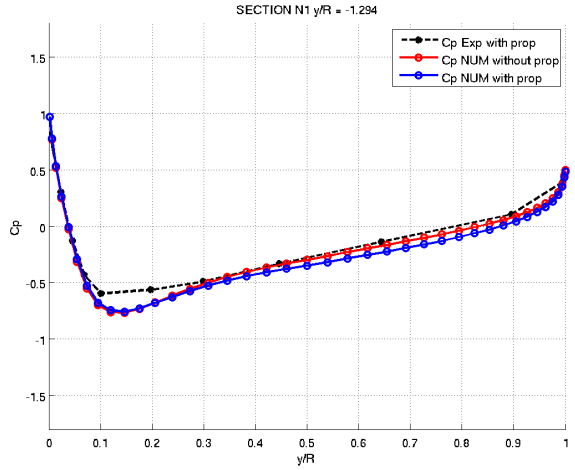


Figure 8: Mean pressure distribution on the upper side of the wing.

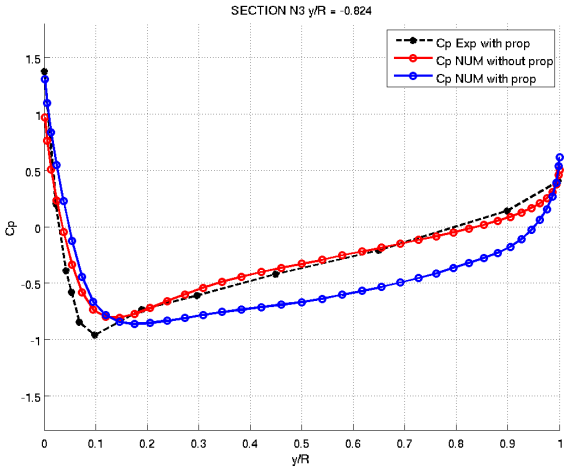


Figure 10: Mean pressure distribution on the upper side of the wing.

Agnes [4]) and the pressure coefficient evaluated numerically with COMPA (the section numbers are the same of [4]) in different sections of the span of the wing. Numerical results are given for each section for both cases with (red line) and without (blue line) propeller. Sections number 1 and 10 are outside the wake of the propeller while sections number 3 and 8 are inside the slipstream of the actuator disk and near the rim of the propeller's blade. The little differences between experimental and numerical mean pressure coefficient in all sections can likely be related to the actuator disk model's hypotheses, as in the previous case. While in sections number 1 and 10 we have good agreement between experimental and numerical data, in sections number 3 and 8 the code seems to overrate the pressure coefficient in lower and upper side respectively.

5 Conclusions

A modification of the analytical solution for the inviscid actuator disk proposed by Conway in [1] is embedded in the low–order velocity–based panel code COMPA for inviscid incompressible flow. The actuator disk model predicts

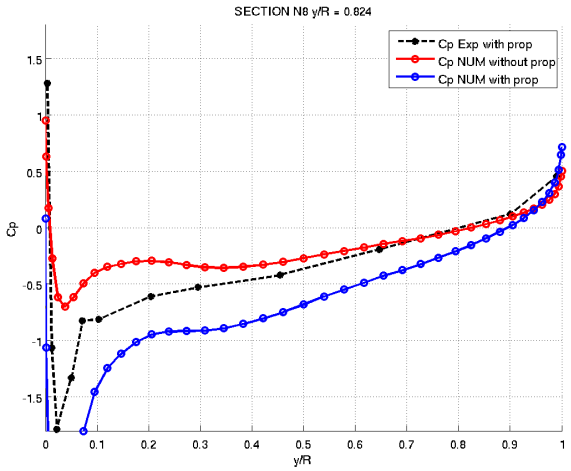


Figure 11: Mean pressure distribution on the lower side of the wing.

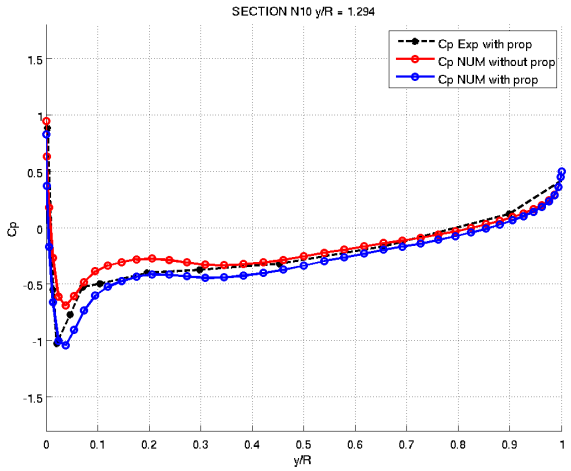


Figure 13: Mean pressure distribution on the lower side of the wing.

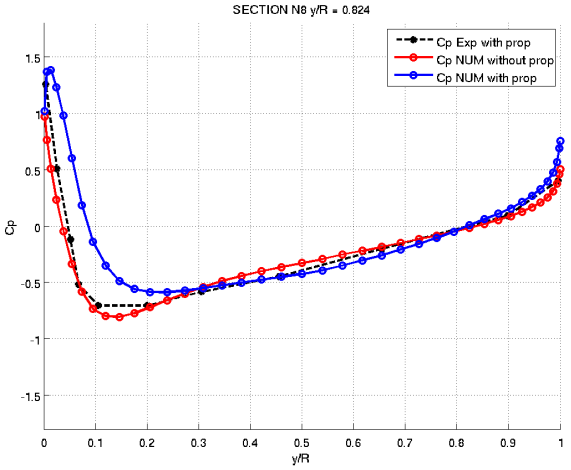


Figure 12: Mean pressure distribution on the upper side of the wing.

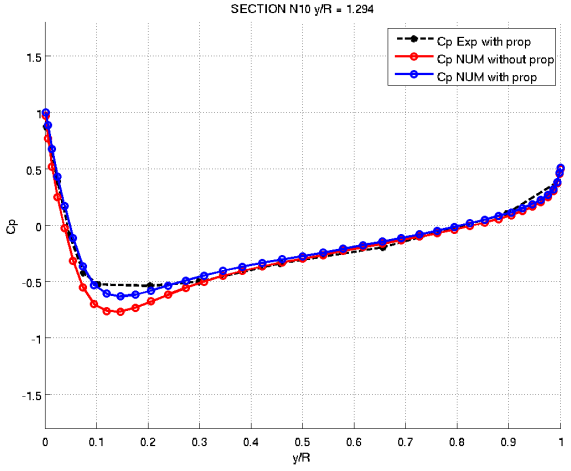


Figure 14: Mean pressure distribution on the upper side of the wing.

the time–averaged induced velocities in the slipstream of a propeller and can be used to study the aerodynamic interaction between propeller and wing, for example. The combined method has been checked by comparison with two different experiments (the first one by Stuper [3], and the second one by Chiaramonte, Favier, Maresca and Agnes [4]). Good results have been found in terms of the spanwise distribution of load on the wing of the first experiment (Stuper [3]) and in terms of the mean pressure distributions at different spanwise sections of the wing of the second case (Chiaramonte, Favier, Maresca and Agnes [4]). Nevertheless, some differences have been found between numerical results and experimental data in both cases and they seem to be related to the actuator disk model's hypotheses. The radial contraction of the vortex tube can be taken into account by modifing the actuator disk model, as shown by Conway in [15]. A more sophisticated modification is required to include the velocity fields induced by the propeller and its associated wake (supposed rigid in this model) in the influence coefficient matrices B_{kj} and A_{kj} of Eq. 31.

In any case, the actuator disk model embed-

ded in COMPA can be useful to predict aerodynamic interaction between propeller and wing with good results in terms of global effects in short computational time. On the other hand, istantaneus phenomena due to the istantaneous position of the wake cannot be studied.

References

[1] Conway, J. T., "Analytical Solution For The Actuator Disk With Variable Radial Distribution Of Load," *Journal of Fluid Mechanics*, Vol. 297, 1995, pp. 327–355.

[2] Conway, J. T. and Su, J., "PMAL Propeller-Induced Asymmetric Flow Calculations For Aurora Aircraft Using Embedded Non–Linear Actuator Disks," *Canadian Aeronautics and Space Journal*, Vol. 46, No. 1, March 2000, pp. 20–27.

[3] Stuper, J., "Effect of the Propeller Slipstream on Wing and Tail," Technical Memorandums TM–874, NACA, Washington D. C., August 1938.

[4] Chiaramonte, J.-Y., Favier, D., Maresca, C., and Agnes, A., "Unsteady Interactional Effects Between A Propeller And A Fixed Wing," *AIAA 1991–3231*, 9th Applied Aerodynamics Conference, Baltimore, Maryland, USA, 23–25 September 1991, pp. 231–245.

[5] Fratello, G., Favier, D., Maresca, C., and Barbi, C., "Experimental And Numerical Study Of The Propeller/Fixed Wing Interaction," *Journal of Aircraft*, Vol. 28, No. 6, 1991, pp. 365–373.

[6] Conway, J. T., "Analytical Solution For The General Non-Axisymmetric Linearized Actuator Disk," *AIAA 20003–3521*, 21th Applied Aerodynamics Conference, Orlando, Florida, USA, 23–26 June 2003, pp. 1–14.

[7] Gradshteyn, I. S. and Ryzhik, I. M., *Tables Of Integrals, Series and Products*, Academic Press, INC., 24/48 Oval Road, London NW1, 1980.

[8] Abramowitz, M. and Stegun, I. S., *Handbook Of Mathematical Functions, With Formulas, Graphs And Mathematical Tables*, Dover, 1972.

[9] Morita, T., "Numerical Calculation of the Elliptic Integrals of the First and Second Kinds with Complex Modulus," *Interdisciplinary Information Sciences*, Vol. 6, No. 1, 2000, pp. 67–74.

[10] Hess, J. L., "Calculation Of Potential Flow About Arbitrary Three Dimensional Lifting Bodies," Report MDC J5679–01, Douglass-McDonnell, Long Beach, California, October 1972.

[11] Hess, J. and Smith, A., "Calculation Of Potential Flow About Arbitrary Bodies," *Progress in Aeronautical Sciences*, Vol. 8, 1967, pp. 1–138.

[12] Maskew, B. and Woodward, F., "Symmetrical Singularity Model For Lifting Potential Flow Analysis," *Journal of Aircraft*, Vol. 13, No. 9, September 1976, pp. 733–734.

[13] Conway, J. T. and Su, J., "PMAL Flow Calculations For Aurora Aircraft Using Non–Axisymmetric Propeller Actuator Disks," *Canadian Aeronautics and Space Journal*, Vol. 49, No. 1, March 2003, pp. 1–9.

[14] Favier, D., Ettaouil, A., and Maresca, C., "Numerical And Experimental Investigation Of Isolated Propeller Wakes In Axial Flight," *Journal of Aircraft*, Vol. 26, No. 9, 1989, pp. 837–846.

[15] Conway, J. T., "Exact Actuator Disk Solution For Non-Uniform Heavy Loading And Slipstream Contraction," *Journal of Fluid Mechanics*, Vol. 365, 1998, pp. 235–267.

CEAS 2011

Moon Exploration

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Numerical Investigation on Low-Energy Transfers to the Moon through Isomorphic Mapping

Mauro Pontani, Marco Giancotti, and Paolo Teofilatto
Scuola di Ingegneria Aerospaziale - University of Rome "La Sapienza"

Keywords: Earth-Moon transfers, Restricted three-body problem, Capture orbits, Lagrange points, Low-energy transfers

Abstract

This paper is concerned with the topological study of low-energy trajectories to the Moon, with emphasis to long-term lunar capture orbits, in the context of the classical circular restricted three-body problem. An original cylindrical representation of the phase space, associated with the third body dynamics, is proposed. Periodic orbits, transit orbits, and trajectories asymptotic to the periodic orbit around the Lagrange point L_1 are mapped through this cylindrical representation. Conley's theorem on the topology of lunar capture trajectories is employed to locate the initial conditions of potential interest, i.e. those generating long-term capture lunar orbits. This approach allows finding a number of long-term capture trajectories. One of these trajectories exhibits special features, in terms both of capture duration and of geometric distribution. A simple strategy – that nevertheless minimizes propellant consumption – to transform low-energy trajectories into the final (supposedly desired) circular orbit around the Moon is then proposed.

1 Introduction

The determination of low-energy trajectories to the Moon, based on the transit in proximity of the collinear Lagrange points L_1 and L_2, have been studied for a long time [1, 2] and some space missions have already taken advantage of the results of these studies. Two examples of this kind of missions are represented by the European Smart-1 mission and by the Japanese Hiten mission.

This paper is concerned with a topological study of the transfer trajectories to the Moon, with particular interest to the low energy trajectories that allow performing long-term lunar captures. In general the spacecraft trajectory oscillates, in the sense that the spacecraft passes continuously from the neighborhood of the Earth to that of the Moon (and vice versa). A fundamental topological result stated by Conley [3] is focused on the location of capture trajectories in the phase space. This theorem is recalled and applied through an original, convenient three-dimensional cylindrical representation of trajectories in the phase space, based on the special form of the equations of motion developed by Birkhoff [4]. Feasibility of trajectories for a real mission is investigated,

and a simple strategy to perform the insertion of a spacecraft into a final, circular, operational orbit around the Moon is then proposed.

2 Circular restricted three-body problem

The circular restricted three-body problem (CR3BP) models the dynamics of three bodies with masses m_1, m_2, and m_3, under the assumption that $m_1 > m_2 \gg m_3 \approx 0$. This means that the mass of the third body (i.e. the spacecraft), m_3, is considered negligible, whereas the remaining massive bodies, termed the primaries, describe counterclockwise circular orbits around their center of mass. This model approximates the dynamics of a spacecraft subject to the gravitational attraction of the Earth and of the Moon (that represent the primaries), in the context of a possible lunar mission.

The problem is conveniently described in synodic coordinates, which represent a coordinate system rotating with the two primaries. The *x*-axis connects the two primaries and is directed from the Earth to the Moon, whereas the *y*-axis lies in the plane of their motion. This problem is analyzed by employing canonical units, which represent a set of convenient normalized units. The distance unit (DU) is the (constant) distance between the two primaries $(1\ \text{DU} = 384400\ \text{km})$, whereas the time unit (TU) is such that the two primaries complete a single orbit in a period equal to 2π TU $(1\ \text{TU} = 375190\ \text{sec})$. If μ_E and μ_M denote respectively the gravitational parameter of the Earth and of the Moon, these definitions of TU and DU imply that $\mu_E + \mu_M = 1\ \text{DU}^3/\text{TU}^2$. After introducing the parameter $\mu \triangleq \mu_M/(\mu_E + \mu_M)$ $(= 0.01215510)$, it is straightforward to rewrite the gravitational parameters of the two primaries as:

$$\mu_E = 1-\mu \quad , \quad \mu_M = \mu \qquad \left(\text{in DU}^3/\text{TU}^2\right) \tag{1}$$

Their respective positions along the *x*-axis are $x_E = -\mu$ and $x_M = 1-\mu$ (in DU).

2.1 Nonlinear dynamics

In the synodic reference frame, which is illustrated in Fig. 1, let (x, y) and (v_x, v_y) denote the coordinates of position and velocity. The equations of motion of the planar CR3BP are the following [5]:

$$\dot{x} = v_x \quad , \quad \dot{y} = v_y \tag{2}$$

$$\dot{v}_x = \frac{\partial \Omega}{\partial x} + 2v_y \quad , \quad \dot{v}_y = \frac{\partial \Omega}{\partial y} - 2v_x \tag{3}$$

where

$$\Omega = \frac{x^2+y^2}{2} + \frac{1-\mu}{\sqrt{(x+\mu)^2+y^2}} + \frac{\mu}{\sqrt{(x+\mu-1)^2+y^2}} \tag{4}$$

The *Jacobi integral*, whose value is referred to as the *Jacobi constant* and denoted with *C*, is an integral of the system [5]:

$$C = 2\Omega - \left(v_x^2 + v_y^2\right) \tag{5}$$

C is associated with the energy of the dynamical system. Zero velocity curves are the geometrical loci where $v_x = v_y = 0$, and constitute the boundary of the region where the spacecraft motion is allowed. Along the *x*-axis, three equilibrium points (the collinear Lagrange points) exist for the problem at hand [5].
The existence of the Jacobi integral allows replacing Eqs. (3) with a single equation for γ, which represents the angle that the velocity vector forms with the *x*-axis. In fact, due to Eq. (5), the two components v_x and v_y can be rewritten as

$$v_x = \sqrt{2\Omega - C}\cos\gamma \quad , \quad v_y = \sqrt{2\Omega - C}\sin\gamma \tag{6}$$

As a result:

$$\tan\gamma = \frac{v_y}{v_x} \tag{7}$$

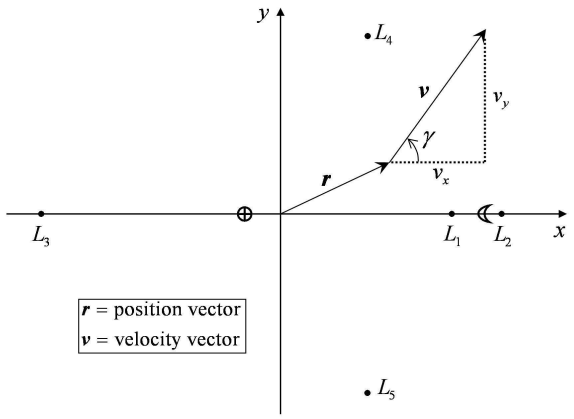


Fig. 1. Synodic reference frame

The time derivative of γ can be obtained:

$$\dot{\gamma} = \frac{\Omega_y \cos\gamma - \Omega_x \sin\gamma}{\sqrt{2\Omega - C}} - 2 \qquad (8)$$

Hence, Eqs. (6) and (8), also referred to as the *Birkhoff's equations*, represent an alternative set of equations of motion for the CR3BP.

2.2 Linear dynamics around L_1

The third body dynamics can be investigated in the neighborhood of the Lagrange point L_1 by linearizing the equations of motion (2)-(3). This section is concerned with Conley's description of the linear dynamics around the collinear Lagrange points [3]. The dynamical state of the third body when it is at rest in L_1 is assumed as the reference state and denoted with $\boldsymbol{X}_{L1}$. The coordinates of position and velocity with respect to $\boldsymbol{X}_{L1}$ are denoted with (ξ,η,u,w) and are given by:

$$\begin{aligned} \xi &= x - x_{L1} & \eta &= y \\ u &= v_x & w &= v_y \end{aligned} \qquad (9)$$

If $\boldsymbol{\chi} \triangleq [\xi \quad \eta \quad u \quad w]^T$ the linear equations of motion (around L_1) are (formally):

$$\dot{\boldsymbol{\chi}} = \mathbf{A}\boldsymbol{\chi} \qquad \text{where} \qquad \mathbf{A} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right|_{\boldsymbol{X}_{L1}} \qquad (10)$$

The general solution of the linear system (10) can be written in terms of the eigenvectors $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4\}$ of $\mathbf{A}$ and of the respective eigenvalues, $\{\alpha,-\alpha,i\beta,-i\beta\}$, to yield:

$$\xi = c_1 e^{\alpha t} + c_2 e^{-\alpha t} + 2|c_3|\cos(\beta t + \phi) \qquad (11)$$

$$\eta = \sigma c_1 e^{\alpha t} - \sigma c_2 e^{-\alpha t} - 2\tau|c_3|\sin(\beta t + \phi) \qquad (12)$$

$$u = \alpha c_1 e^{\alpha t} - \alpha c_2 e^{-\alpha t} - 2\beta|c_3|\sin(\beta t + \phi) \qquad (13)$$

$$w = \alpha\sigma c_1 e^{\alpha t} + \alpha\sigma c_2 e^{-\alpha t} - 2\beta\tau|c_3|\cos(\beta t + \phi) \qquad (14)$$

where

$$\sigma \triangleq \frac{2\alpha}{1-\alpha^2-\lambda^2} \qquad \tau \triangleq \frac{2\beta}{1+\beta^2-\lambda^2}$$
$$\lambda^2 \triangleq \frac{1-\mu}{|x_{L1}+\mu|^3} + \frac{\mu}{|x_{L1}+\mu-1|^3} \qquad (15)$$

The constants c_1, c_2, and c_3 $\left(=|c_3|e^{i\phi}\right)$ depend on the dynamical state at the initial time t_0 (which is set to 0) and on the spacecraft energy (directly related to *C*). This issue is addressed at the end of this subsection.

In the context of the linear equations of motion of the CR3BP, due to the form of Eqs. (11)-(14), orbits can be classified as:

(i) *periodic orbits*, if $c_1 = c_2 = 0$: these orbits are clockwise elliptic trajectories around L_1

(ii) *asymptotic orbits*, if $c_1 = 0$ and $c_2 \neq 0$ (asymptotic orbits converging into the periodic orbit) or if $c_1 \neq 0$ and $c_2 = 0$ (asymptotic orbits departing from the periodic orbit)

(iii) *bouncing orbits*, if $c_1 c_2 > 0$: these orbits do not cross the region where L_1 is located

(iv) *transit orbits*, if $c_1 c_2 < 0$: these orbits cross the region where L_1 is located

The orbits of interest for the present study are the asymptotic orbits.

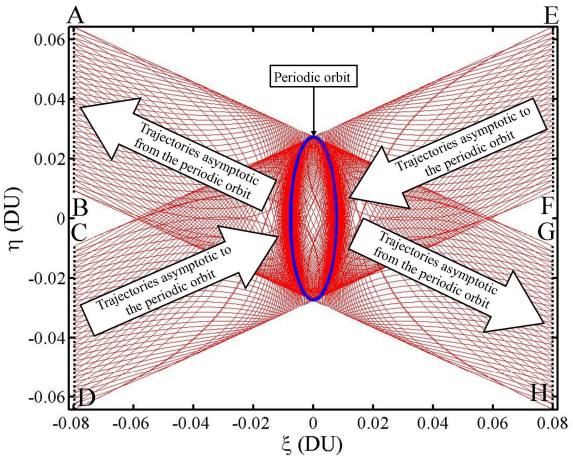


Fig. 2. Periodic orbit and asymptotic trajectories (linear dynamics)

2.3 Cylindrical representation of trajectories

For a given value of the Jacobi constant *C* (corresponding to a specified energy level) through Birkhoff's equations the trajectory of the third body can be conveniently calculated by employing the coordinates (x, y) and the angle γ in place of (x, y, v_x, v_y). This allows for a geometrical, visual representation of two-dimensional trajectories in the phase space. In this work a new, convenient representation is produced by defining an isomorphism into new coordinates (X, Y, Z) as follows:

$$X = x \qquad Y = \rho\cos\gamma \qquad Z = \rho\sin\gamma \qquad \left(\rho \triangleq y - y_{min}\right) \tag{16}$$

where y_{min} is set to a convenient value (-0.104236 DU in this study, with $C = 3.185$ DU^2/TU^2) such that $\rho \geq 0$ for the entire trajectory. Through Eq. (16) the instantaneous position and velocity of the third body correspond to a single point in the transformed (phase) space (X, Y, Z). The mapping (16) can allow detecting homoclinic and heteroclinic connections among distinct trajectory arcs. The existence of such connections is a crucial aspect in the restricted three-body problem, with possible profitable implications for spaceflight [6].

The same transformation can also be used to represent the periodic orbit and the associated asymptotic trajectories found with the linearized equations of motion (10), as shown in Fig. 3, which portrays the cylindrical mapping of the orbits illustrated in Fig. 2. Asymptotic trajectories are represented as a four-fold surface. The cross sections of this surface at $\xi_0 = -0.08$ DU and $\xi_0 = 0.08$ DU define four disks, conventionally named D_{1E}, D_{2E}, D_{1M}, and D_{2M}. These disks identify (in the phase space) the asymptotic orbits that depart from the segments CD and EF and those converging into the segments AB and GH of Fig. 2. In particular, Fig. 4 illustrates the disk D_{2E} (at $\xi_0 = -0.08$ DU). If η_0 is specified, then the two values $\gamma_0^{(min)}$ and $\gamma_0^{(max)}$ shown in Fig. 4 correspond to the pair of asymptotic orbits that depart from $\left(\xi_0, \eta_0, \gamma_0^{(min)}\right)$ and $\left(\xi_0, \eta_0, \gamma_0^{(max)}\right)$ and converge into the periodic orbit. It can be demonstrated that transit orbits departing from (ξ_0, η_0) fulfill the inequality $\gamma_0^{(min)} < \gamma_0 < \gamma_0^{(max)}$ and correspond to the points that form the arc AB. It is straightforward to deduce that all the points contained in the disk D_{2E} of Fig. 4 are associated with transit trajectories, whereas the external points correspond to bouncing orbits. Therefore, through the cylindrical mapping asymptotic trajectories are represented as a four-fold, "pair-of-pants" surface that separates transit orbits from bouncing trajectories. Transit trajectories are generated from points contained in the "pair-of-pants" surface and flow through it; bouncing orbits are external to this surface and do not cross the region where the Lagrange point L_1 is located.

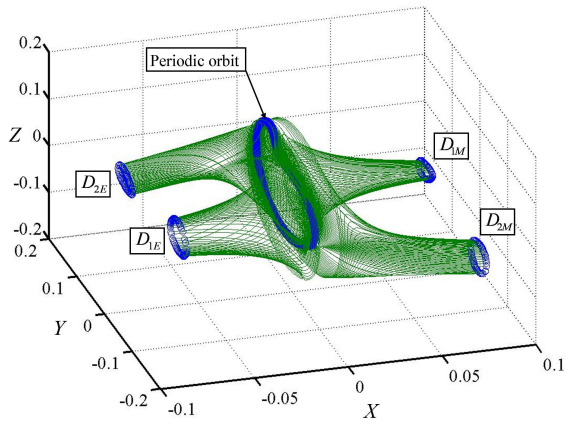


Fig. 3. Cylindrical representation of the periodic orbit and of the asymptotic trajectories (linear dynamics)

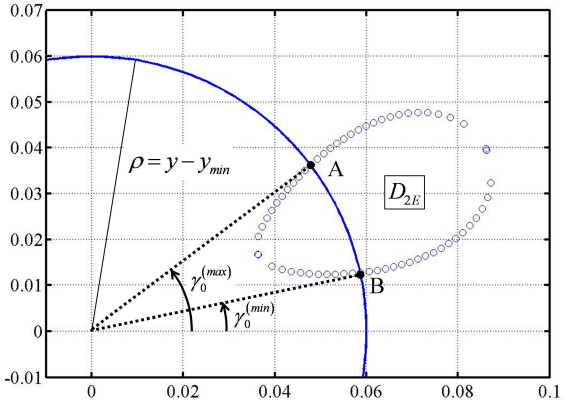


Fig. 4. Disk D_{2E} of the cylindrical representation ((*Y*,*Z*)-plane, cross section at $\xi_0 = -0.08$)

3 Lunar capture orbits

In the context of the CR3BP, lunar capture orbits are orbits that remain indefinitely in the neighborhood of the Moon, without crossing the region where L_1 is located. The determination of (non-periodic) lunar capture orbits is extremely challenging, as emphasized in the scientific literature [7, 8, 9]. Periodic, asymptotic and capture orbits represent three special classes of trajectories that exhibit an interesting, non-oscillating behavior. They can be characterized under the topological point of view through the use of Conley's theorem.

Conley's theorem states that [3]: "if a crossing asymptotic orbit exists then near any such there is a capture orbit". In the remainder of this paper this assertion will be used together with the cylindrical representation of trajectories to locate long-term lunar capture orbits.

3.1 Long-term lunar capture orbits

As a theoretical premise for the task of determining lunar capture orbits, the following property of the solutions of the CR3BP is recalled: the trajectory obtained by forward integration of the equations of motion (2)-(3) with (generic) initial conditions $(x_0, y_0, v_{x0}, v_{y0})$ is symmetrical – with respect to the *x*-axis – to that attained by backward integration of the same equations with initial conditions $(x_0, -y_0, -v_{x0}, v_{y0})$. Formally, let $(x_1(t_0+\Delta t), y_1(t_0+\Delta t), v_{x1}(t_0+\Delta t), v_{y1}(t_0+\Delta t))$ be the trajectory generated from $(x_0, y_0, v_{x0}, v_{y0})$ at the time $(t_0+\Delta t)$ and $(x_2(t_0-\Delta t), y_2(t_0-\Delta t), v_{x2}(t_0-\Delta t), v_{y2}(t_0-\Delta t))$ the trajectory generated from $(x_0, -y_0, -v_{x0}, v_{y0})$ at the time $(t_0-\Delta t)$. This means that

$$(x_1(t_0), y_1(t_0), v_{x1}(t_0), v_{y1}(t_0)) \equiv (x_0, y_0, v_{x0}, v_{y0})$$

$$(x_2(t_0), y_2(t_0), v_{x2}(t_0), v_{y2}(t_0)) \equiv (x_0, -y_0, -v_{x0}, v_{y0})$$

Hence, for any $\Delta t > 0$, the following symmetry relations hold:

$$\begin{aligned} x_1(t_0+\Delta t) &= x_2(t_0-\Delta t) \\ y_1(t_0+\Delta t) &= -y_2(t_0-\Delta t) \\ v_{x1}(t_0+\Delta t) &= -v_{x2}(t_0-\Delta t) \\ v_{y1}(t_0+\Delta t) &= v_{y2}(t_0-\Delta t) \end{aligned} \qquad (17)$$

This property can be profitably used to locate suitable initial conditions that yield lunar capture orbits. In fact, due to Conley's theorem,

in the phase space the boundary conditions for lunar capture orbits are located in the neighborhood of the states associated with trajectories asymptotic to the periodic orbit around L_1. This means that these boundary conditions can be sought on the boundary of the disk D_{2M}, which in fact corresponds to states (approximated to first order) converging asymptotically into the periodic orbit. Therefore long-term capture orbits (in principle, even stable lunar orbits) could be found by backward integration of the dynamic equations with initial conditions located on the boundary of the disk D_{2M}. However, because of the symmetry described above, this backward integration generates trajectories symmetrical to those generated by forward integrating the equations of motion with initial conditions located on the boundary of the disk D_{1M}. This is due to the fact that each point P_{2M} on the boundary of D_{2M} corresponds to a symmetric point P_{1M} belonging to the boundary of D_{1M}. Hence, the existence of long-term capture orbits can be investigated making reference to the disk D_{1M} as the geometrical locus (in the phase space) where suitable initial conditions are located.

Figures 5(a) and 5(b) illustrate the disks D_{1E} and D_{2E} (cross sections of the surface portrayed in Fig. 3 at $\xi_0 = -0.02$ DU) and the disks D_{1M} and D_{2M} (cross sections of the same surface at $\xi_0 = 0.02$ DU), respectively. The initial conditions for the numerical integration of the Birkhoff's equations correspond to 500 points along the boundary of D_{1M}. By definition, evasion from lunar capture occurs at the time t_f, when the spacecraft crosses (directed toward the Earth) the plane $x = x_{L1}$. The number of loops around the Moon that the spacecraft completes before evading from capture can be computed.

Figure 6 reports the number of loops around the Moon before evasion for these 500 trajectories, whereas Fig. 7 shows the corresponding evasion times. With this method

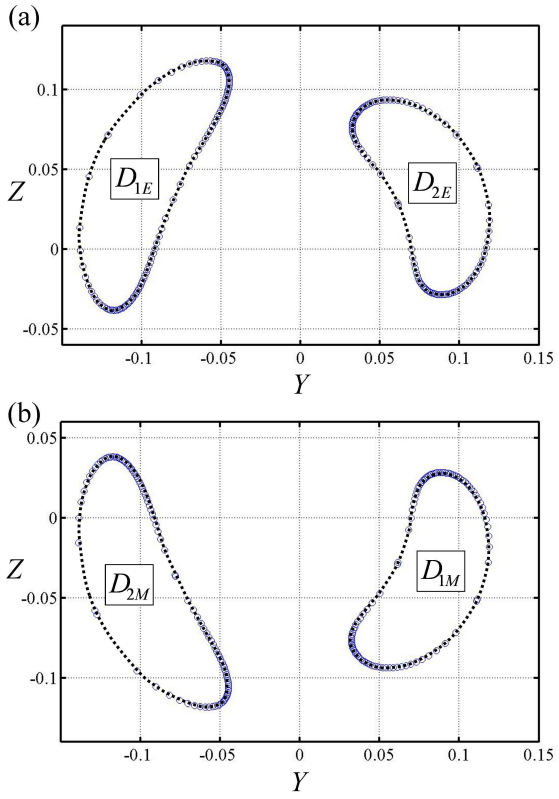


Fig. 5. Disks of the cylindrical representation; cross sections at $\xi_0 = -0.02$ (a) and $\xi_0 = 0.02$ (b)

a single trajectory that orbits the Moon for about 8.5 sidereal years has been determined. This trajectory is portrayed in Fig. 8 and exhibits very peculiar geometric features. First, it is symmetrical with respect to the *x*-axis (except in the portion where the third body is close to L_1, just before evading from capture). Secondly, this trajectory is extremely "dense" in special regions of the (x, y) plane, in the sense that many trajectory arcs are grouped, to form macroscopically wide stripes separated by regions not crossed by any trajectory arc.

This example orbit is not a capture orbit, because it leaves the Moon after 8.5 years. However it is longer than most of the other orbits that have been simulated, and therefore it may be conjectured that it is located close to a capture orbit. This conjecture can be substantiated by employing a local optimization algorithm, with the intent of determining orbits

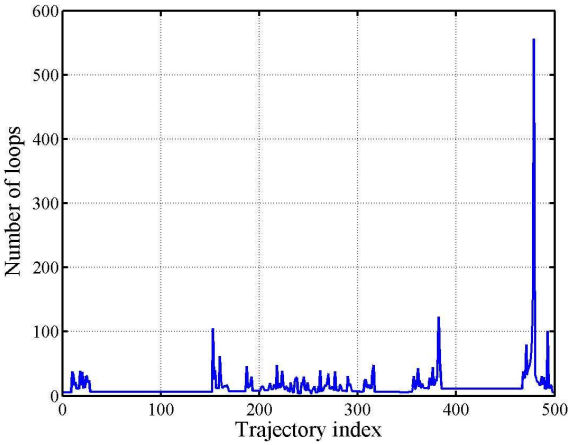


Fig. 6. Number of loops around the Moon for the trajectories departing from the boundary of D_{1M}

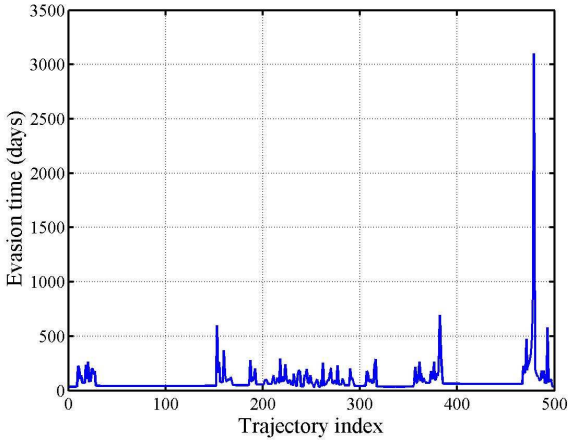


Fig. 7. Evasion time for the trajectories departing from the boundary of D_{1M}

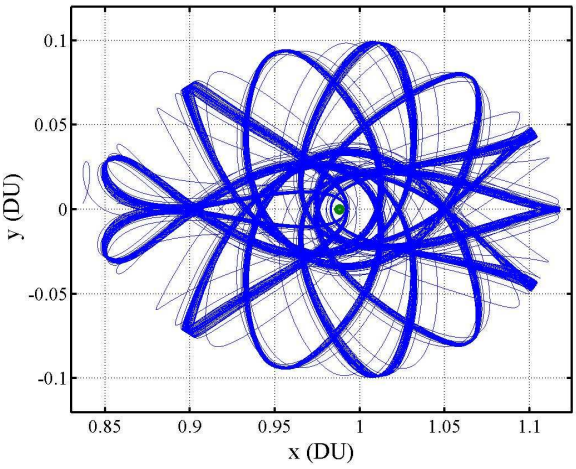


Fig. 8. Long-term capture trajectory

with increasing capture durations. This means that the objective function to be minimized for a similar algorithm is $J = -t_C$ (where t_C denotes the capture duration). Starting from the state corresponding to orbit A, the numerical optimizer (i.e. the MATLAB routine *fmincon*) finds a succession of orbits characterized by increasing capture durations (cf. Table 1). Apparently, the algorithm does not exhibit a convergent behavior (in reasonable computational times) to any locally minimizing result (i.e. to any maximum-duration capture orbit) and this circumstance can be assumed as indicative of the existence of an actual capture orbit, which can be numerically approached with an arbitrary accuracy.

An analysis of the periselenium altitudes shows that actually several trajectories are infeasible because they would impact the surface of the Moon. Nevertheless, from the theoretical point of view the results achieved in this section demonstrate the validity of the method for determining long-term capture orbits, on the basis of topological arguments.

3.2 Operational lunar orbits

Transfer trajectories to the Moon are usually aimed at inserting a spacecraft into a terminal specified lunar orbit, in general a circular orbit. None of the trajectories calculated in the preceding section terminates into the desired circular orbit. Orbital maneuvers are then needed to perform a transfer to the desired final (circular) orbit, whose specified altitude is set to 100 km in this case. In this context, the Moon radius is taken into account, in the sense that if a portion of trajectory follows an impact then it is neglected. A two-impulse Hohmann-like transfer is selected as the transfer strategy that allows transforming the trajectory into the final circular orbit. For each trajectory computed in Section 3.2 all the aposelenia and periselenia are determined and the corresponding Hohmann-like transfer is evaluated. Then, for each trajectory the orbital maneuver (i.e. the aposelenium) that minimizes the characteristic

Table 1. Iterations of the optimization algorithm

Iter.	x (DU)	y (DU)	γ (rad)	t (days)
1	8.568929195145340e-1	3.383246244880390e-3	1.913217098842190e-1	3102
2	8.568929095145339e-1	3.383246244880390e-3	1.913217098842190e-1	3134
3	8.568929195145340e-1	3.383236244880390e-3	1.913217098842190e-1	3159
4	8.568929195145340e-1	3.383246244880390e-3	1.913216998842190e-1	3285
5	8.568928951004714e-1	3.383221830817890e-3	1.913216854701565e-1	3509
6	8.568928951004714e-1	3.383211830817890e-3	1.913216854701565e-1	3581
7	8.568929195204944e-1	3.383197422715855e-3	1.913216687607437e-1	3589
8	8.568929195204944e-1	3.383197422715855e-3	1.913216787607437e-1	3589
9	8.568929242287879e-1	3.383185221644342e-3	1.913216809727949e-1	4701
10	8.568929243264907e-1	3.383185031025871e-3	1.913216810446838e-1	5286
11	8.568929243333040e-1	3.383185019112240e-3	1.913216810461537e-1	5949
12	8.568929243333040e-1	3.383185019112316e-3	1.913216810461536e-1	6444

velocity of the transfer is selected. Figure 9 reports the minimum value of the characteristic velocity for each trajectory. The most economical trajectory is associated with a total velocity change of 0.671 km/sec. Of course captures of longer duration imply a greater number of aposelenia where the first impulse can be applied and this circumstance leads to potential savings in terms of characteristic velocity of the maneuver. However, long-term capture orbits are often impact trajectories. Only the portion of the entire trajectory preceding the impact is to be taken into account and this actually can limit the benefit of using long-term capture orbits to perform a real mission.

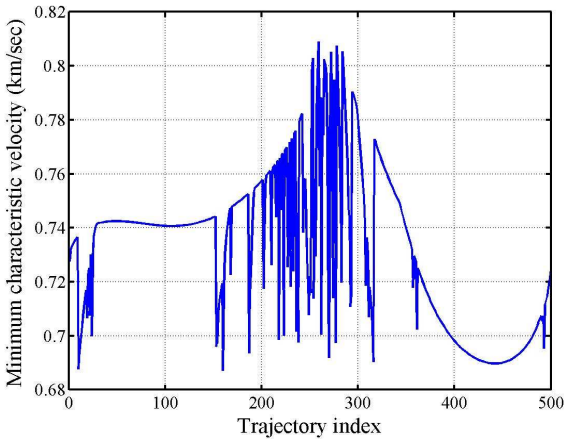


Fig. 9. Minimum characteristic velocity for injection into the terminal orbit

4 Conclusions

This paper investigates the topology of the low-energy trajectories to the Moon, with particular interest to long-term capture lunar trajectories. This goal is achieved by introducing an original cylindrical representation of the phase space that describes the third body dynamics.

The linear dynamics around the Lagrange point L_1 is used to study the mapping among four regions of the phase space. These regions and the orbits therein contained have a straightforward visual representation if the cylindrical mapping is employed. With this method other relevant aspects of the flow can be observed directly, such as the existence of homoclinic and heteroclinic orbits.

Using Conley's theorem, it is found that long-term capture orbits are to be sought along the boundary of one of the disks of the cylindrical representation.

The numerical results attained with the method described in the paper demonstrate that relatively long-term capture orbits exist. Moreover, the numerical investigation proves that periselenia of decreasing altitude are obtained as the number of loops increases. An orbit with special features, in terms both of evasion time (8.5 sidereal years) and of spatial distribution, has been found.

As transfer trajectories to the Moon are usually aimed at inserting a spacecraft into a terminal circular lunar orbit, Section 3.3 is devoted to determining a simple strategy to

achieve this result. A two-impulse Hohmann-like transfer is selected as the transfer maneuver that allows transforming the low-energy trajectory into the final circular orbit.

References

[1] Belbruno E., *Capture Dynamics and Chaotic Motions*, Princeton University Press, Princeton, 2004

[2] Belbruno E., Miller J., "Sun Perturbed Earth to Moon Transfers with Ballistic Capture", *Journal of Guidance, Control, and Dynamics*, Vol. 16, pp. 770-775, 1993

[3] Conley C., "Low Energy Transit Orbits in the Restricted Three-Body Problem", *SIAM Journal of Applied Mathematics*, Vol. 16, No. 4, pp. 732-746, 1968

[4] Birkhoff G. D., "The Restricted Problem of Three Bodies", *Rend. Circ. Matem. Palermo*, t. XXXIX (1° sem. 1915), pp. 265-334, 1915

[5] Szebeheli V., *Theory of Orbits*, Academic Press, New York, NY, 1967

[6] Koon W. S., Lo M. W., Marsden J. E., Ross S. D., "Heteroclinic Connections between Periodic Orbits and Resonance Transitions in Celestial Mechanics", *Chaos*, Vol. 10, pp. 427-469, 2000

[7] Egorov V., "The Capture Problem in the Three Body Restricted Orbital Problem", *Artificial Satellites Earth*, Vol. 3, pp. 3-17, 1961

[8] Heppenheimer T. A., Porco C., "New Contributions to the Problem of Capture", *Icarus*, Vol. 30, pp. 385-401, 1977

[9] Horedt G., "Capture of Planetary Satellites", *Astronautical Journal*, Vol. 81, pp. 675-678, 1976

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Lunar Soft Landing Trajectory Optimization in a 6DoF Dynamical Model

R. Armellin, D. Dowlat, M. Lavagna
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano,
Via La Masa 34, 20156, Milano, Italy

Keywords: Lunar Landing, Trajectory Optimization, 6DoF Dynamics

Abstract

An algorithm for the optimization of a lunar soft-landing trajectory is presented. A 6DoF modeling of the dynamics is adopted together with an accurate description of the Moon gravity field. The problem is faced as a direct optimization problem with the goal of obtaining a vertical landing whilst minimizing the overall fuel consumption. The descent trajectory is supposed to start from the periselenium of a low Moon orbit. Four optimization phases are considered. Each phase is characterized by a different set of optimization variables, constraints, and increasing level of complexity. In the first phase the thrust direction is optimized considering the translational motion of the lander only. Furthermore, no throttle capability is considered. In the second phase the thrust direction is fixed in the spacecraft body reference frame. The proper thrust orientation is obtained by optimizing the control torques supplied to the lander by the attitude subsystem. In the third phase the thrust magnitude is optimized too, and the constraint of landing on specific site is added. Furthermore, more restrictive constraints on the final velocities (linear and angular) are set. Finally, in the fourth phase a more accurate gravitational model of Moon that includes the main harmonics is considered. The algorithm is tested on a lunar landing trajectory in an area close to the Moon's equator, with a candidate spacecraft for Google Lunar X-prize.

1 Introduction

During the 60's, the challenge for conquering the Moon gave a large impetus for the research and development of guidance algorithms for the soft landing on the lunar surface. Especially the Apollo missions made the majority contribution to the studies.
Akridge and Harlin [1] varied parameters such as initial parking orbit altitudes, thrust-to-weight ratio and specific impulse of the engine to define the best descent trajectory. Bennett and Price [2] worked on the Apollo landing trajectory and they divided the powered descent phase into three phases: a fuel-optimum braking phase, a landing approach transition phase and a terminal descent phase. The Apollo missions used an explicit guidance algorithm that was derived by Cherry [3].
Beyond Apollo there were also the Surveyor vehicles that used a guidance scheme based on the work done by Cheng [4]. They used an attitude control law that forced the flight path towards the vertical as time progressed.
During the same period other guidance algorithms were proposed such as Hall, Dietrich and Tiernan [5], Kriegsman and Reiss [6], Citron, Dunin and Meissinger [7], Feng and Wasynczuk [8] and, more recently, Sostaric and Rea [9].
However, a guidance algorithm needs a parallel trajectory optimization to obtain more efficient

solution. During the Apollo period the trajectory optimization capabilities were limited by computing power, but, in the last 15 years, with the increase in computing power, new studies were developed. Vasile and Floberghaghen [10] applied a Spectral Elements in Time (SET) approach to the lunar soft landing problem. Their work was developed by Ramanan and Madan Lal [11] that solved a two-point boundary problem with a controlled random search. Bei and Zhang [12] defined a precise-landing mission for small lunar probes. Liu, Duan and Teo [13] formulated an optimal control problem to ensure the soft landing with the least fuel consumption. Then Shan [14] studied a two-dimensional approach to find an optimal solution for the descent phase. More recently Cho, Jeong, Lee and Bang [15] considered the energy of de-orbit burning as a parameter to be optimized for the descent trajectory. Besides, in the last few years, NASA resumed the work of Bauer, Cornick and Stevenson [16], and Striepe, with the help of many other collaborators, treated it in several works. Firstly, they reviewed the previous work [17], and then, they developed a more efficient optimization tool [18]. Finally, they used the results to study the Autonomous Landing and Hazard Avoidance Technology (ALHAT) project [19] and then Davis [20] improved it.

The contribution of this paper is to define a method to compute an optimized descent trajectory from a circular parking orbit that could be used as the guidance for the landing of an autonomous spacecraft. The objectives of the current research are to:

1. Determine feasible landing trajectories, which minimize fuel expenditure, while considering translational motion and attitude dynamics.

2. Analyze the fuel usage due to the constraints considered as terminal attitude, terminal velocity and control histories.

The entire descent trajectory is included in the analysis and is treated as a single problem. This is done in order to generate a trajectory which is optimal over the entire time span of the landing process.

2 Equations of Motion

The 6DoF dynamics of the vehicle is written in body-fixed reference frame as

$$\begin{cases} \dot{\mathbf{r}} = \mathbf{v} - \boldsymbol{\omega} \times \mathbf{r} \\ \dot{\mathbf{q}} = \boldsymbol{\Omega} \cdot \boldsymbol{q} \\ \dot{\mathbf{v}} = \dfrac{\mathbf{F}}{m} - \boldsymbol{\omega} \times \mathbf{v} + \dfrac{\mathbf{T}}{m} \\ \dot{\boldsymbol{\omega}} = \mathbf{I}^{-1} \cdot (\mathbf{M}_c - \boldsymbol{\omega} \times \mathbf{I} \cdot \boldsymbol{\omega} - \dot{\mathbf{I}} \cdot \boldsymbol{\omega}) \\ \dot{m} = \dfrac{T}{I_{sp} g_0} \end{cases} \tag{1}$$

where $\mathbf{r}$ and $\mathbf{v}$ are the position and velocity vectors, $\mathbf{q}$ is the vector of quaternions used for attitude representation, $\boldsymbol{\omega}$ is the angular velocity. The external forces (only gravitational) acting of the body are indicated by $\mathbf{F} = \{F_x, F_y, F_z\}^T$, whereas the thrust is $\mathbf{T}$. It is assumed that no external torques act the on vehicle, and that a control authority is available only on the z axis, i.e., $\mathbf{M}_c = \{0, 0, M_{c_z}\}^T$. The mass variation is dependent on the propellant consumption: T is the thrust module, I_{sp} is the specific impulse, and g_0 is the Earth gravity acceleration at zero altitude. The time variation of the inertia tensor $\mathbf{I}$ is computed by assuming a linear variation with spacecraft mass. The dynamics of Eq. (1) are implemented in the *Custom Variable Mass 6DoF (quaternion)* block as part of the *Aerospace Blockset* of *Simulink*. This block asks as inputs the initial values for the position, linear and angular velocities (P, Q, R), Euler angles (ϕ, θ, ψ), inertia tensor, and mass. Note that the initial position vector must be supplied in Moon centered inertial reference frame. The numerical integration is performed with a fixed-step forth-order Runge-Kutta scheme.

3 Optimization Setup

The goal is to obtain a Moon vertical soft landing with the minimum propellant mass consumption. If m_0 is the initial mass of the spacecraft and m_F is mass at landing, the objective function to be minimized is:

$$\mathbf{g}(\mathbf{x}) = -\frac{m_F}{m_0}, \tag{2}$$

where $\mathbf{x}$ is the vector of optimization variables, scaled in such a way that $0 \leq \mathbf{x} \leq 1$. The nonlinear constraints can be written as $\boldsymbol{\psi}_l \leq \boldsymbol{\psi}(\mathbf{x}) \leq \boldsymbol{\psi}_u$. Note that upper and lower bounds must be equal in the previous expression to obtain equality constraints. The optimization problem is solved with the function *fmincon* of the *Matlab Optimization Toolbox*, which implements a sequential quadratic programming algorithm for the solution of a nonlinear parametric optimization problems. At every iteration of the optimization the block of the dynamics is called to propagate the trajectory, which is necessary for the evaluation of the objective and constraint functions. The optimization process is divided into phases. In each phase a different vector of variables is optimized and different set constraints are considered. The first two phases do not examine the problem of landing on a specific point, whereas the third and the fourth one consider the landing site location and a tridimensional landing trajectory. Only in the last phase the main harmonics of the Moon gravitational field are accounted for.

3.1 First Trajectory Optimization

In the first phase (FTO, First Trajectory Optimization) the attitude dynamics of the vehicle is neglected and the optimization vector is defined by

$$\mathbf{x} = \{\boldsymbol{\beta}, T, m_F\}^T, \tag{3}$$

where $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_k\}^T$ is the control angle measured clockwise from the local horizon to the thrust direction. How this angle is realized in practice is not addressed during this phase. The elements of $\boldsymbol{\beta}$ are defined over a time grid, whose values are collected in the vector $\mathbb{t}$

$$\mathbb{t} = \left[t_0 : \frac{t_F}{k-1} : t_F\right]^T. \tag{4}$$

The initial values of $\boldsymbol{\beta}$ are those of an optimal thrust direction profile, computed by solving planar optimization problem via calculus of variations [11]. The values at a generic instant of the trajectory are obtained by interpolating $\boldsymbol{\beta}$ with cubic splines. Note that the final mass is included in the optimization vector to reduce the optimization problem in Meyer form [22] and that a single thrust level is considered. The set of nonlinear constraints are

$$\begin{cases} |\mathbf{r}(t_F)| = R_M \\ |\mathbf{v}(t_F)| \leq w_V \\ m\left(t_F\right) = m_F, \end{cases} \tag{5}$$

where R_M is the Moon radius and w_V is a scalar value that limits the velocity modulus. These constraints force the satellite to achieve the Moon surface with the required landing velocity. The objective function is represented by Eq. 2 and the dynamical system by Eq. (1). The spacecraft is subjected to one control force given by the thrust engine $\mathbf{T} = \{T_x(\beta), T_y(\beta), 0\}^T$, and no external torques are applied. When the optimization process ends the optimal values of the control angle, the thrust module, and final mass are obtained, which are used for the subsequent optimization phase.

3.2 Second Trajectory Optimization

From the FTO analysis the profile of the thrust direction in local horizon reference frame is known. In the Second Trajectory Optimization (STO) it is assumed to have $\mathbf{T}$ always aligned with x body-axis direction, but with opposite direction. If $\alpha(t)$ is the angle between $\mathbf{T}$ and x direction of the inertial frame, it is possible to compute the control torque $\mathbf{M}_c(t)$ necessary to obtain the desired attitude by:

$$M_{c_z} \simeq \dot{I}_{zz}\dot{\alpha} + I_{zz}\ddot{\alpha}, \tag{6}$$

where $\mathbf{M}_c(t) = \{0, 0, M_{c_z}(t)\}^T$, and $\dot{\alpha}$ and $\ddot{\alpha}$ are obtained with a five-point differentiation formula [23]. The values of M_{c_z} are in body-fixed frame. The optimization vector for the STO problem is:

$$\mathbf{x} = \{\mathbb{M}, T_x, m_F, \psi_0, R_0\}^T, \tag{7}$$

where $\mathbb{M} = \{M_{c_z}(t_0), \ldots, M_{c_z}(t_F)\}^T$ is the vector that contains the values of M_{c_z} at $\mathbb{t}$, ψ_0 and R_0 are the yaw and the corresponding angular rate at time t_0, respectively. The set of nonlinear

constraints is the same of the FTO, with the only exception of reducing the admitted final velocity. The controls acting on the spacecraft in the body fixed frame are the thrust $\mathbf{T} = \{-T_x, 0, 0\}^T$ and torques $\mathbf{M}_c = \{0, 0, M_{c_z}(t)\}^T$. Thus, in the STO problem the attitude dynamics of the vehicle is analyzed with the goal of computing the control torque required for the optimal orientation thrust.

3.3 Third Trajectory Optimization

In the TTO (Third Trajectory Optimization) thrust modulation is included, and the final time is added to the optimization vector

$$\mathbf{x} = \{\mathbb{M}, \mathbb{T}, t_F, m_F, \psi_0, R_0\}^T, \tag{8}$$

where $\mathbb{T}$ is the vector of thrust defined on $\mathbb{t}$. The first guess for the $\mathbb{T}$ vector is built by replicating the single value computed in the STO for all the time points. The goal the third optimization is to compute a vertical soft landing trajectory. This goal is achieved by setting the linear constraints

$$\begin{cases} |\mathbf{r}(t_F)| = R_{cb} \\ \dfrac{\mathbf{r}(t_F)}{|\mathbf{r}(t_F)|} \times \mathbf{x} = \mathbf{0} \\ r_x(t_F) \leq 0 \\ 0 \leq v_x^{rel}(t_F) \leq w_A \\ \left|\left[v_y^{rel}(t_F), v_z^{rel}(t_F)\right]\right| \leq w_B \\ R(t_F) = 0 \\ m(t_F) = m_F, \end{cases} \tag{9}$$

where w_A are w_B are two scalar values. The first three equations impose constraints on the spacecraft position. The product vector enforces the vertical landing. When the product vector is zero then $\mathbf{r}$ and $\mathbf{x}$ are aligned but, since r_x is negative, they have opposite direction. The fourth and fifth inequalities represent constraints on the velocity relative to the landing site. In particular, the velocity has to be confined inside a narrow cone aligned with the local vertical. The second-last equation forces the satellite to approach the surface with no spin around the z axis.

The control actions system in the body reference system is:

$$\begin{cases} \mathbf{F}_c = \{-T_x(t), 0, 0\}^T \\ \mathbf{M}_c = \{0, 0, M_{c_z}(t)\}^T. \end{cases} \tag{10}$$

In the TTO the constraint on the landing site location is set and the tridimensional geometry of the orbit is computed. This part of the problem is analyzed after the optimization has successfully ended. The argument of periapsis of the transfer orbit ω is computed by

$$\omega = \text{asin}\left(\frac{\sin(\phi_{LS})}{\cos(i)}\right) - \vartheta, \tag{11}$$

where i is its inclination, ϕ_{LS} is the landing site latitude, and ϑ is the angle spanned by the landing trajectory, computed as

$$\vartheta = \text{acos}\left(\frac{\mathbf{r}(t_0) \cdot \mathbf{r}(t_F)}{|\mathbf{r}(t_0)||\mathbf{r}(t_F)|}\right). \tag{12}$$

The sidereal time of the landing site at landing θ_{LS} is given by

$$\theta_{LS} = \text{atan}\left(\frac{r_J(t_F)}{r_I(t_F)}\right), \tag{13}$$

where λ_{LS} is its longitude, and r_I and r_J are the first and second component of the position vector in Moon inertial reference frame. The sidereal time of the prime meridian at the beginning of the landing maneuver is

$$\theta_{PM} = \theta_{LS} - \lambda_{LS} - \omega_M t_F, \tag{14}$$

where $\omega_M t_f$ accounts for Moon rotation during the landing phase. With these information it is possible to compute both the point on the orbit where to start the landing maneuver and and its initial time. In the last landing trajectory optimization, called Real Trajectory Optimization (RTO), the main lunar gravitational harmonics are considered to refine the optimal landing path.

3.4 Real Trajectory Optimization

The main spherical harmonics are considered to better describe the Moon gravitational field. In

order to compute the exact value of the gravitational field during the optimization it is necessary to know the three-dimensional position vector of the spacecraft at every time step; thus, the value of ϑ has to be known during the optimization process. For this reason ϑ is introduced in the optimization vector

$$\mathbf{x} = \{\mathbb{M}, \mathbb{T}, t_F, m_F, \psi_0, R_0, \vartheta\}^T, \qquad (15)$$

The constraints adopted are those of system Eq. (9) with an additional condition

$$\vartheta(t_F) = \vartheta_F. \qquad (16)$$

The control forces and torques remain unchanged with respect to TTO.

4 Numerical Example

The landing problem starts from a circular lunar parking orbit of 350 km of altitude, inclined of 4.59 deg with respect to the Moon equator, and with a right ascension of ascending node of -13.06 deg. An impulsive maneuver injects the vehicle in transfer orbit that lowers the periselenium height to 15 km, from which the landing trajectory begins. The landing site is characterized by a longitude $\phi_{LS} = -2.99$ deg and a latitude $\lambda_{LS} = -23.34$ deg. The vehicle has an initial mass of 661 kg. The mass distribution is such that the initial tensor of inertia is

$$\mathbf{I}_0 = \begin{bmatrix} 385.58 & 0 & 0 \\ 0 & 385.58 & 0 \\ 0 & 0 & 330.5 \end{bmatrix} \text{ Nm}^2.$$

A maximum thrust of 1680 N is considered with an specific impulse of 318 seconds. The attitude control is guaranteed by momentum wheels with maximum torque capability of 0.1 Nm. The test case is representative of a preliminary scenario for an automatic mission in the framework of Google Lunar X-prize.

4.1 FTO and STO

The FTO optimization problem consider a first guess vector $\mathbf{x}_0$ of

$$\begin{cases} \boldsymbol{\beta}_g &= [175.99, 173.68, 169.01, \ldots \\ & \ldots\ 163.76, 157.84, 151.19] \text{ deg} \\ T_g &= 1400 \text{ N} \\ m_g &= 350 \text{ kg} \end{cases}$$

where the values of $\boldsymbol{\beta}$ are defined at $\mathbb{t} = [0, 110, 220, 330, 440, 550]$ s, and the subscript g stands for guess. The value of the constraint constant of Eq. 5 is $w_V = 5$ m/s. The optimization process takes 25 iteration to compute the optimal solution

$$\begin{cases} \boldsymbol{\beta} &= [180.0, 174.69, 168.64, \ldots \\ & \ldots\ 162.09, 154.95, 147.13] \text{ deg} \\ T &= 1675.7 \text{ N} \\ m_F &= 365.64 \text{ kg} \end{cases}$$

Starting from the values of $\boldsymbol{\beta}$ the control torque to guarantee the proper orientation of the thrust is numerically computed applying Eq. (6). The numerical values corresponding to $\mathbb{t}$ are

$$\mathbb{M}_g = [-0.157, -0.199, -0.408, -0.386, \ldots \\ \ldots\ -0.365, -0.412] \times 10^{-3} \text{ Nm},$$

plotted in Figure 1 in solid line. These values, together with $\psi_g = 90$ deg and $R = 0$ rad/s are used as first guesses for the STO. Referring to Eq. (5), a value of $w_V = 4$ m/s is set to limit the final velocity.
The optimization process ends in 13 iterations leading to the optimal solution

$$\begin{cases} \mathbb{M} &= [-0.083, -0.241, -0.392, -0.379, \ldots \\ & \ldots\ -0.399, -0.413] \times 10^{-3} \text{ Nm} \\ T &= 1673.5 \text{ N} \\ m_F &= 366.03 \text{ kg} \\ \psi_0 &= 88.21 \text{ deg} \\ R_0 &= 0.1015 \text{ deg/s} \end{cases}$$

Note that the optimal solution found is very close to the first guess. Figure 2 shows the altitude profile during the landing phase and Fig.

3 a zoom on the final part of the trajectory. It can be noticed that the problem formulation produces a landing trajectory almost tangential to the Moon surface with a non null final angular velocity, as shown in Fig. 4.

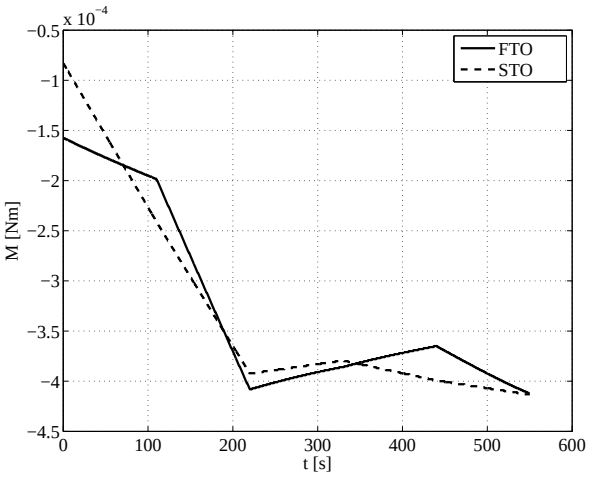


Figure 1: Comparison between the torque profile associated to the FTO and the optimized one in the STO

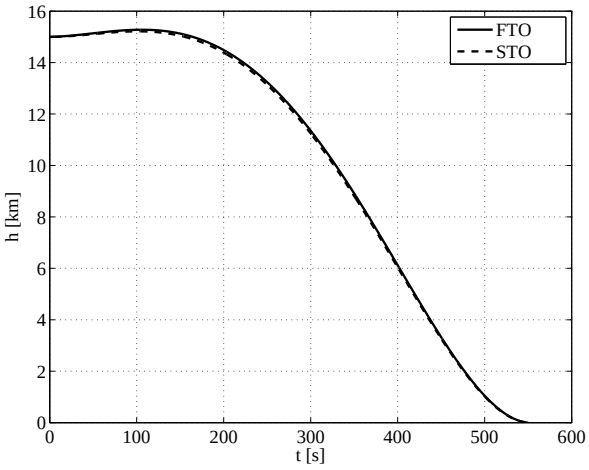


Figure 2: FTO and STO height profiles

4.2 TTO and RTO

The optimal values of the STO are used to setup the TTO. The final time $t_F = 550$ s is used as first guess and the vector $\mathbb{t}$ is used to build the initial thrust vector $\mathbb{T}_g = [1673.5, \ldots, 1673.5]$ N. The values $w_A = 3$ m/s and $w_B = 0.5$ m/s are set for the constants appearing in of Eq. 9. After 53 iterations the optimizer delivers the optimal solution

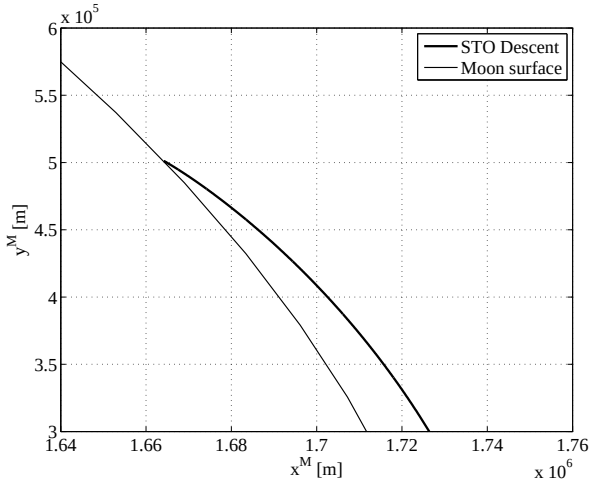


Figure 3: Zoom on the final part of the STO trajectory in Moon-centered Moon-fixed reference frame

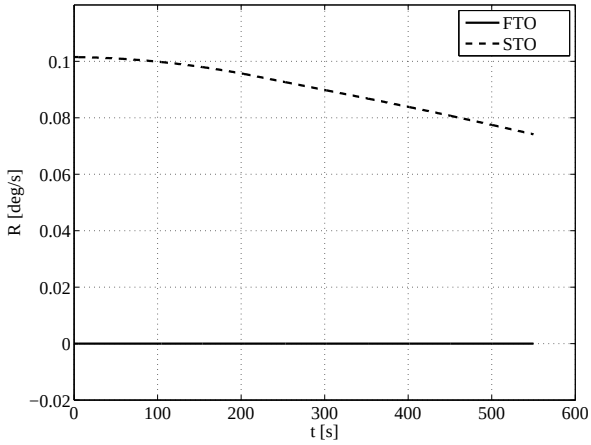


Figure 4: FTO and STO angular velocity profiles

$$
\begin{cases}
\mathbb{M} &= [0.544, -0.12, -0.02, 0.003, \ldots \\
&\quad \ldots\ 0.480, -0.982] \ \times 10^{-1} \text{ Nm} \\
\mathbb{T} &= [1.6762, 1.6778, 1.6697, 1.6800, \ldots \\
&\quad \ldots\ 1.3899, 1.6286] \ \times 10^{3} \text{ N} \\
m_F &= 346.41 \text{ kg} \\
t_F &= 608.23 \text{ s} \\
\psi_0 &= 84.01 \text{ deg} \\
R_0 &= -0.2698 \text{ deg/s}
\end{cases}
$$

Note that the imposition of the vertical landing produces a large increase in the required control torque as visible in Fig. 5. Furthermore, the maneuver duration sensibly increases causing a slight decrease in the landing mass. The negative effect of a longer duration on the propellant consumption is partially mitigated by the use of thrust modulation as shown in Fig. 6. Figure 7 shows final portion of the landing trajectory, in which the vertical arrival is achieved. Note that an increase in the altitude over the transfer orbit periselenium is necessary to accomplish this goal (see Fig. 8). In this case, increasing the transfer orbit periselenium could be beneficial for propellant mass reduction. Finally, figures 9 and 10 shows the behavior of Euler angles and the satisfaction on the constraint on the final angular velocity.

After the trajectory is optimized, relevant quantities to design the entire landing phase are computed by Eq. (11)-(13)

$$\begin{cases} \vartheta = 16.08 \text{ deg} \\ \omega = -56.76 \text{ deg} \\ \theta_{LS} = -53.64 \text{ deg} \\ \theta_{PM} = -30.40 \text{ deg}. \end{cases}$$

Once that ω is computed the transfer trajectory is fully designed. The value of the sidereal time of the prime meridian completely defines the starting time of the descent trajectory.

The *Spherical Harmonic Gravity Model* block implemented in *Simulink* is used to better describe the gravitational field of the Moon in the RTO. Spherical harmonics up to degree 4 are considered. The results of the third optimization are used as first guesses. In particular the time grid is defined by

$$\mathbb{t} = [0, 121.64, 243.29, 364.94, 486.59, 608.23] \text{ s},$$

and the guess of $\vartheta = 16.08$ deg is appplied. The values of w_A and w_B remain unchanged. The optimization process converges in 9 iterations to optimal solution

$$\begin{cases} \mathbb{M} = [0.457, -0.106, 0.016, -0.070, \ldots \\ \quad \ldots\ 0.508, -0.961] \times 10^{-1} \text{ Nm} \\ \mathbb{T} = [1.6274, 1.6725, 1.6795, 1.6775, \ldots \\ \quad \ldots\ 1.4921, 1.1835] \times 10^{3} \text{ N} \\ m_F = 349.16 \text{ kg} \\ t_F = 613.70 \text{ s} \\ \psi_0 = 85.68 \text{ deg} \\ R_0 = -0.2131 \text{ deg/s} \\ \vartheta_F = 16.37 \text{ deg}. \end{cases}$$

Finally, the updated value of anomaly of the pericenter is $\omega = -57.05$ deg. A little difference in the initial orientation of the satellite leads to a lower maximum height (see Fig. 8). The time span is similar, but RTO has a better thrust profile (see Fig. 6) that permits a lower mass consumption. Furthermore RTO final velocity has a lower value with respect to TTO. Finally, it is verified that the landing site point is unchanged (see Fig. 11). Figure 12 shows the complete lunar landing trajectory: the red line defines the RTO descent, the blue line the transfer orbit, the black line the parking orbit, the green line the lunar equator, and the dash-dotted line the prime meridian.

5 Conclusions

In this work a method for the optimization of the Moon soft-landing problem has been presented. The goal is to optimize both linear and angular dynamics of a spacecraft landing on the Moon surface. The objective was to implement an optimization method in a working environment useful for the successive implementation of descent and landing control laws. For this reason *Matlab Simulink* has been chosen. The spacecraft has been modeled as a rigid body with variable inertia tensor and mass. The purpose of the optimization process is to minimize the fuel consumption and to obtain a vertical landing with final zero angular velocity. The problem has been analyzed in four consecutive phases. The FTO and STO phases do not consider the constraint on the landing site and soft constraints

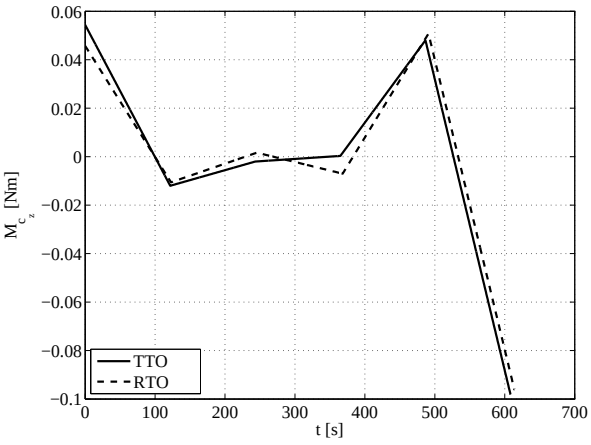


Figure 5: Torque profiles for TTO and RTO

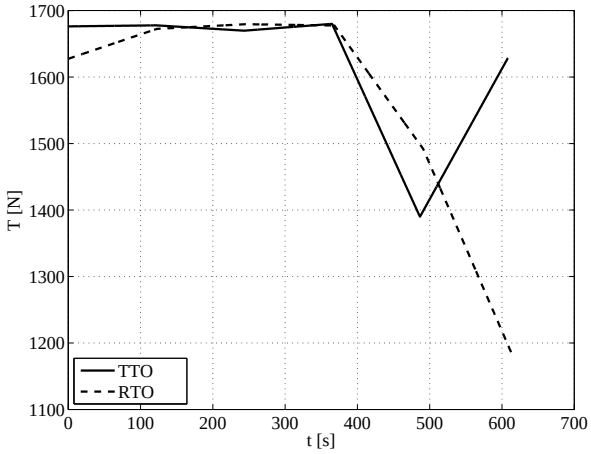


Figure 6: Thrust profiles for TTO and RTO

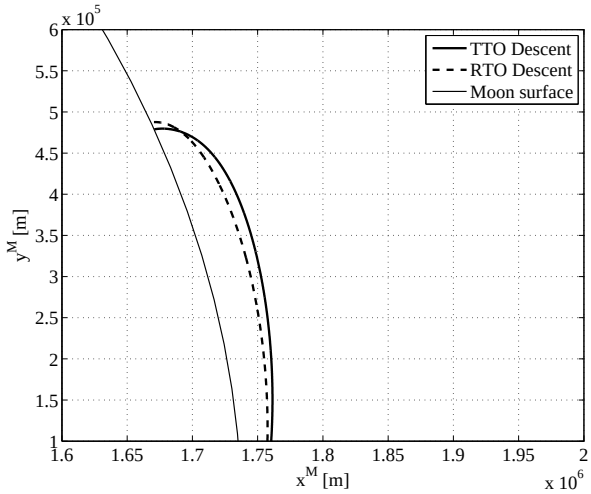


Figure 7: TTO and RTO trajectories in Moon-centered Moon-fixed frame

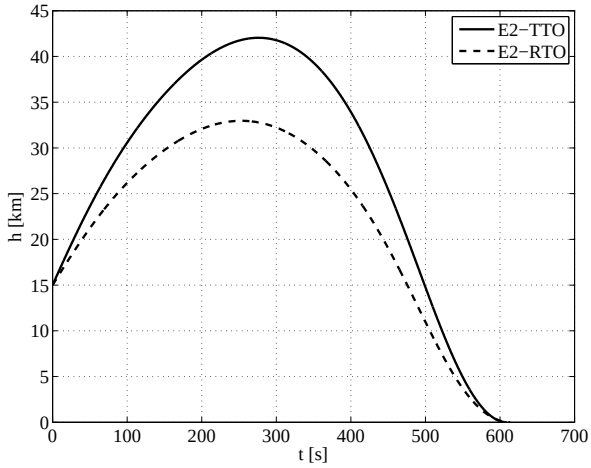


Figure 8: TTO and RTO height profiles

on the final velocity are imposed. Furthermore, the thrust is kept constant along the time span. In the TTO and RTO phases the final velocity is enforced to be close to the local vertical axis with null angular rates. Finally, thrust modulation is considered. A realistic lunar landing scenario has been analyzed to show the performance of the optimization algorithm.

Several areas of future work can be identified:

1. The three-dimensional approach can be extended. A landing site out of the orbital plane can be considered allowing control torques in all directions.

2. The optimization problem can include both the parking orbit parameters and the descent starting height as optimization variables.

3. A more refined model of external perturbations can be included. Firstly, the number of spherical harmonics considered can be incremented. Then, other external perturbations, such as solar radiation pressure or third-body effect, may be considered.

References

[1] Akridge, C.M., and Harlin, S.H., "Parametric performanec analysis of lunar missions, Part II, Lunar descent", Tech. Rep. MTP-P & VE-F-63-8, NASA, George C. Marshall Flight Center, Huntsville, AL, May 1963.

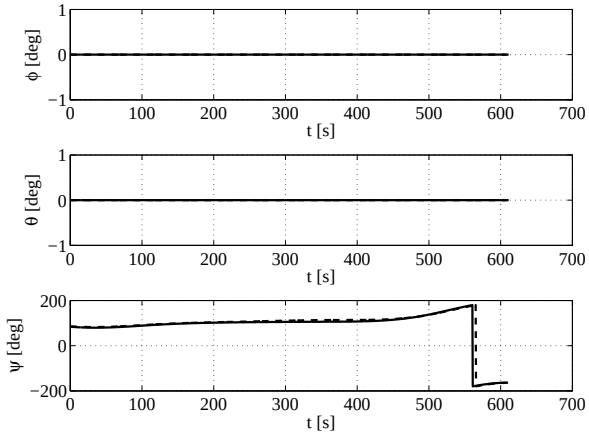


Figure 9: E2-RTO euler angles profile

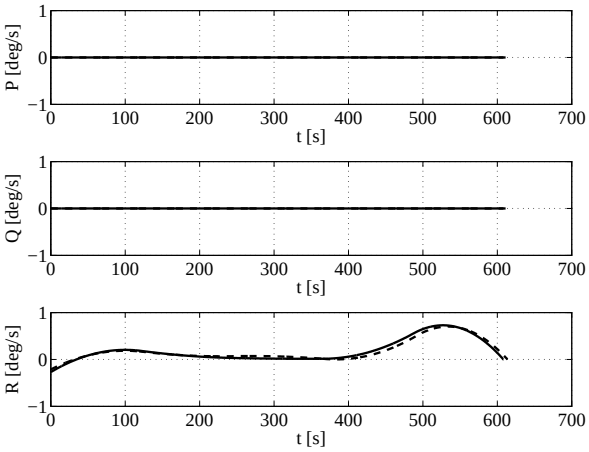


Figure 10: E2-RTO angular velocity profile

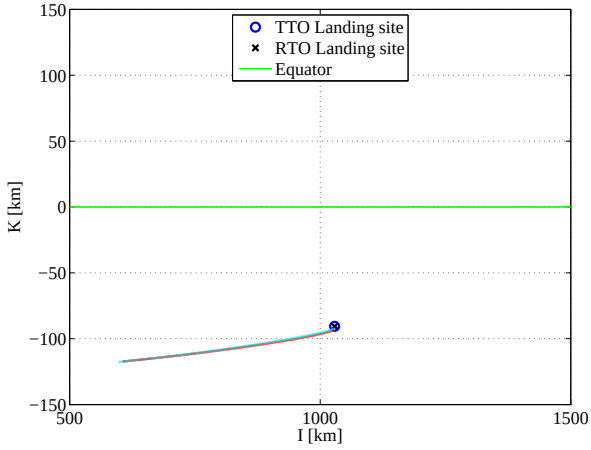


Figure 11: RTO landing site location

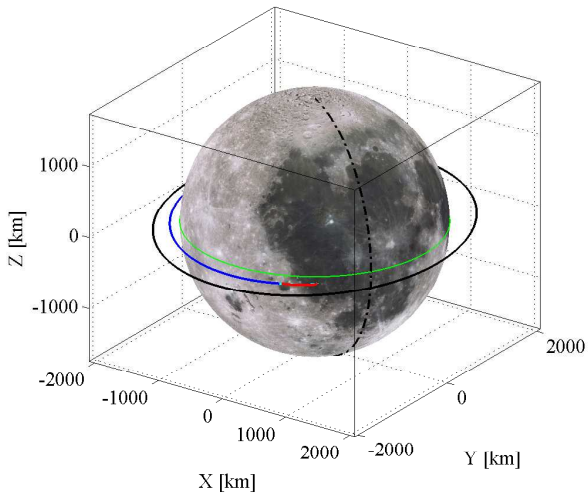


Figure 12: RTO trajectory three-dimensional representation

[2] Bennett, F.V., and Price, T.G., "Study of powered descent trajectories for manned lunar landings", Project Apollo Working Paper 1084, NASA Manned Spacecraft Center, Houston, TX, August 1963.

[3] Cherry, G.W., "E Guidance - A general explicit optimizing guidance law for rocket-propelled spacecraft", Tech. Rep. R-456, MIT Instrumentation Laboratory, Cambridge, MA, 1964.

[4] Cheng, R.K., "Surveyor terminal guidance", *IFAC Symposium on automatic control in the peaceful uses of space*, Stavanger, Norway, 1965, pp. 499-514.

[5] Hall, B.A., Dietrich, R.G., and Tiernan, K.E., "A minimum fuel vertical touchdown lunar landing guidance technique", *AIAA Guidance and Control Conference*, Cambridge, MA, 1963.

[6] Kriegsman, B.A., and Reiss, M.H., "Terminal guidance and control techniques for soft lunar landings", *ARS Journal*, 1963.

[7] Citron, S.J., Dunin, S.E., and Meissinger, H.F., "A terminal guidance technique for lunar landing", Hughes Aircraft Company, El Segundo, Cal., AIAA Journal, Vol. 2, No. 3, 1964.

[8] Feng, T.Y., and Wasynczuk, C.A., "Terminal guidance for soft and accurate lunar landing for unmanned spacecraft",

Hughes Aircraft Company, El Segundo, Cal., AIAA Journal, Vol. 5, No. 6, 1968.

[9] Sostaric, R.,R., and Rea, J. R., "Powered Descent Guidance Methods for the Moon and Mars", *AIAA 2005-6287*, AIAA Guidance, Navigation, and Control Conference and Exhibit, San Francisco, CA, 15-18, August 2005.

[10] Vasile, M., and Floberghaghen, R., "Optimal trajectories for lunar landing missions", *AAS/GSFC International Symposium*, Advances in the Astronautical Sciences, Greenbelt, MD, 1998, pp. 281-295.

[11] Ramanan, R.V., Madan Lal, "Analysis of optimal strategies for soft landing on the Moon from lunar parking orbits", *Journal of Earth System Science*, Vol. 114, No. 6, pp. 807-813, 2005.

[12] Bei, C., and Zhang, W., "A guidance and control solution for small lunar probe precise-landing mission", *Acta Astronautica*, Vol. 62, No. 1, January 2008, pp. 44-47.

[13] Liu, X.L., Duan, G.R., and Teo, K.L., "Optimal soft landing control for moon lander", *Automatica*, Vol. 44, No. 4, April 2008, pp. 1097-1103.

[14] Shan, Y., and Duan, G., "Study on the optimal fuel consumption of the singularity condition for lunar soft landing", *Proc. of the 2007 IEEE International Conference on Robotics and Biomimetics*, 2007, pp. 2254-2258.

[15] Cho, D.H., Jeong, B., Lee, D. and Bang, H., "Optimal perilune altitude of lunar landing trajectory", *International Journal of Aeronautical and Space Sciences*, Vol. 10, No. 1, May 2009, pp. 67-74.

[16] Bauer, G.L., Cornick, D.E., and Stevenson, R., "Capabilities and Applications of the Program to Optimize Simulated Trajectories (POST)", NASA CR-2770, February 1977.

[17] Powell, R., Striepe, S., Desai, P., and Braun, R., "Program to Optimize Simulated Trajectories (POST)", NASA, NAS1-13611, October 1997.

[18] Striepe, S.A., et al, "Program to Optimize Simulated Trajectories (POSTII)", Vol. II Utilization Manual, Version 1.1.6.G, January 2004, NASA Langley Research Center, Hampton, VA.

[19] Striepe, S.A., Fisher, J.L., "POST2 end-to-end Descent and Landing Simulation for the Autonomous Landing and Hazard Avoidance Technology Project", *Advanced in Astronautical Sciences*, 07-119, Langley Research Center, 2007.

[20] Davis, J., et. al., "POST2 End-to-End Descent and Landing Simulation for ALHAT Design Analysis Cycle 2", *International Planetary Probe Workshop 2010*, Paper 434, Barcelona, Spain, Jun 2010.

[21] Cook, G.E., "Perturbations of satellite orbits by tesseral harmonics in the Earth's gravitational potential ", *Planetary and Space Science*, 1963, Vol. 11, No. 7, pp. 797 to 815, Royal Aircraft Establishment, Farnborough UK. doi:10.1016/j.physletb.2003.10.071.

[22] Betts, J.T., "Practical Methods for Optimal Control Using Nonlinear Programming ", *Advances in Design and Control*, SIAM, Philadelphia, PA, 2001.

[23] Wu, C.H., "Formulas and tables of coefficients for numerical differentiation with function values given at unequally spaced points and application to solution of partial differential equations", *National Advisor Committee for Aeronautics*, Technical note 2214, Lewis Flight Propulsion Laboratory Cleveland, Ohio, 1950.

3rd CEAS Air&Space Conference
21st AIDAA Congress

High Order Optimal Feedback Control of Lunar Landing and Rendezvous Trajectories

P. Di Lizia, R. Armellin, and F. Bernelli-Zazzera
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Milano, Italy

M. Berz
Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA

Keywords: Optimal control; Lunar landing; Rendezvous; High-order methods; Differential Algebra

Abstract

Optimal feedback control is classically based on linear approximations, whose accuracy drops off rapidly in highly nonlinear dynamics. A high order optimal control strategy is proposed in this work, based on the use of differential algebraic techniques. In the frame of orbital mechanics, differential algebra allows the dependency of the spacecraft state on initial conditions and environmental parameters to be represented by high order Taylor polynomials. The resulting polynomials can be manipulated to obtain the high order expansion of the solution of two-point boundary value problems. Based on the reduction of the optimal control problem to an equivalent two-point boundary value problem, differential algebra is used in this work to compute the high order expansion of the solution of the optimal control problem about a reference trajectory. New optimal control laws for displaced initial states are then obtained by the mere evaluation of polynomials.

1 Introduction

Nominal space trajectories are usually designed by solving optimal control problems that maximize the payload launch-mass ratio while achieving the primary mission goals. However, uncertainties and disturbances affect the spacecraft dynamics in real scenarios. Moreover, state identification is influenced by navigation errors; consequently, the spacecraft state is only known with a given accuracy. Thus, after the nominal solution is computed, an optimal control strategy that assures the satisfaction of mission goals in the real scenario must be implemented. More specifically, given an initial deviation of the spacecraft state from its nominal value, the optimal control strategy aims at canceling the effects of such a deviation on the satisfaction of the mission requirements by correcting the nominal control law, while minimizing propellant consumption.

Classical optimal feedback control strategies are based on linear approximations, whose main advantage is the simplification of the problem. However, their accuracy drops off rapidly with increasing errors and decreasing control frequencies in highly nonlinear dynamics. Thus, nonlinear optimal feedback control has gained particular interest in recent years, and several strategies have appeared to tackle nonlinearities. One of the highly promising and rapidly emerging methodologies for designing nonlinear controllers is the state-dependent Riccati equation (SDRE) approach, which was originally proposed by Pearson and Burghart and then described in details by Cloutier, Hammett and

Beeler [1]. This approach involves manipulating the governing dynamic equations into a pseudo-linear non-unique form in which system matrices are given as a function of the current state and minimizing a quadratic-like performance index. An algebraic Riccati equation using the system matrices is then solved repetitivily online to give the optimal control law. Thus, the SDRE approach might turn out to be computationally expensive when the solution of the Riccati equation is not properly managed. This can prevent its use for real-time optimal control. A sub-optimal solution can be obtained using the approximating sequence of Riccati equations (ASRE) method [6]. Based on an approximate solution to the Hamilton-Jacobi-Bellman equation, the ASRE method avoids solving the Riccati equation repetitively at every instant and provides a closed-form feedback controller.

An alternative approach was recently introduced by Park and Scheeres, which relies on the theory of canonical transformations and their generating functions for Hamiltonian systems [8]. More specifically, canonical transformations are able to solve boundary value problems between Hamiltonian coordinates and momenta for a single flow field. Thus, based on the reduction of the optimal control problem to an equivalent boundary value problem, they can be effectively used to solve the optimal control problem analytically as a function of the boundary conditions, which is instrumental to optimal feedback control. The main difficulty of this approach is finding the generating functions via the solution of the Hamilton-Jacobi equation. This problem was solved by Park and Scheeres by expanding the generating function in power series of its arguments.

Differential algebraic (DA) techniques [3] are used in this work to develop an alternative approach to the generating function method. Differential algebra serves the purpose of computing the derivatives of functions in a computer environment. More specifically, by substituting the classical implementation of real algebra with the implementation of a new algebra of Taylor polynomials, it expands any function f of v variables into its Taylor series up to an arbitrary order n. DA techniques are used in this work to represent the dependency of the spacecraft state on the initial conditions by means of high order Taylor polynomials. Then, the resulting Taylor polynomials are manipulated to impose the boundary and optimality conditions of the optimal control problem. This enables the expansion of the solution of the optimal control problem with respect to the initial conditions about an available reference trajectory. The resulting Taylor polynomials can be evaluated for new solutions of the original optimal control problem, so avoiding repetitive runs of classical iterative procedures.

The paper is organized as follows. A brief introduction to differential algebra is given in Sect. 2. Being at the basis of the proposed methods, the possibility of expanding the flow of ODEs is presented in Sect. 3. The optimal control problem and the algorithm for the high order expansion of its solution are illustrated in Sect. 4 and 5, respectively. The application of the algorithm to lunar landing and rendezvous maneuvers is addressed in Sect. 6 and 7, respectively.

2 Differential Algebra

DA techniques find their origin in the attempt to solve analytical problems by an algebraic approach [3]. Historically, the treatment of functions in numerics has been based on the treatment of numbers, and the classical numerical algorithms are based on the mere evaluation of functions at specific points. DA techniques are based on the observation that it is possible to extract more information on a function rather than its mere values. The basic idea is to bring the treatment of functions and the operations on them to the computer environment in a similar way as the treatment of real numbers. Referring to Fig. 1, consider two real numbers a and b. Their transformation into the floating point representation, $\overline{a}$ and $\overline{b}$ respectively, is performed to operate on them in a computer environment. Then, given any operation $\times$ in the set of real numbers, an adjoint operation $\otimes$ is defined in the set of FP numbers such that the diagram in figure commutes. (The diagram commutes approximately in practice, due to truncation errors.) Consequently, transforming the real num-

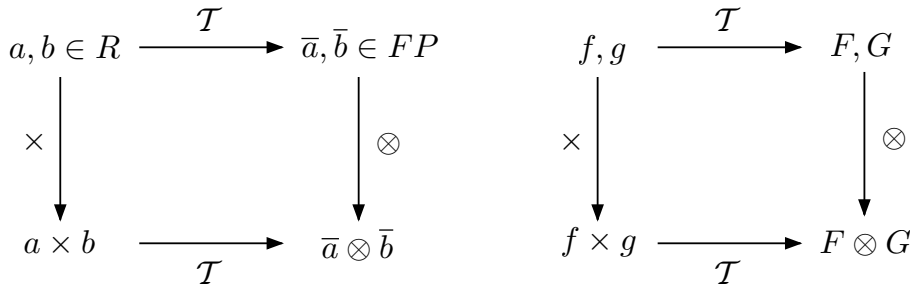


Figure 1: Analogy between the floating point representation of real numbers in a computer environment (left figure) and the introduction of the algebra of Taylor polynomials in the differential algebraic framework (right figure).

bers a and b in their FP representation and operating on them in the set of FP numbers returns the same result as carrying out the operation in the set of real numbers and then transforming the achieved result in its FP representation. In a similar way, suppose two sufficiently regular functions f and g are given. In the framework of differential algebra, the computer operates on them using their Taylor series expansions, F and G respectively. Therefore, the transformation of real numbers in their FP representation is now substituted by the extraction of the Taylor expansions of f and g. For each operation in the function space, an adjoint operation in the space of Taylor polynomials is defined such that the corresponding diagram commutes; i.e., extracting the Taylor expansions of f and g and operating on them in the function space returns the same result as operating on f and g in the original space and then computing the Taylor expansion of the resulting function. The straightforward implementation of differential algebra in a computer allows to compute the Taylor coefficients of a function up to a specified order n, along with the function evaluation, with a fixed amount of effort. The Taylor coefficients of order n for sums and product of functions, as well as scalar products with reals, can be computed from those of summands and factors; therefore, the set of equivalence classes of functions can be endowed with well-defined operations, leading to the so-called truncated power series algebra [2].

Similarly to the algorithms for floating point arithmetic, the algorithm for functions followed, including methods to perform composition of functions, to invert them, to solve nonlinear systems explicitly, and to treat common elementary functions [3]. In addition to these algebraic operations, also the analytic operations of differentiation and integration are introduced, so finalizing the definition of the DA structure. The differential algebra sketched in this section was implemented by Berz and Makino in the software COSY-Infinity [4].

3 High Order Expansion of ODE Flow

The differential algebra introduced in the previous section allows to compute the derivatives of any function f of v variables up to an arbitrary order n, along with the function evaluation. This has an important consequence when the numerical integration of an ODE is performed by means of an arbitrary integration scheme. Any explicit integration scheme is based on algebraic operations, involving the evaluations of the ODE right hand side at several integration points. Therefore, carrying out all the evaluations in the DA framework allows differential algebra to compute the arbitrary order expansion of the flow of a general ODE initial value problem.

Without loss of generality, consider the scalar initial value problem

$$\begin{cases} \dot{x} = f(x) \\ x(t_i) = x_i. \end{cases} \tag{1}$$

Replace the point initial condition x_i with the DA representative of its identity function, $[x_i] = x_i^0 + \delta x_i$, where x_i^0 is the reference point for the

expansion. If all the operations of the numerical integration scheme are carried out in the framework of differential algebra, the Taylor expansion of the solution with respect to the initial condition is obtained at each step. As an example, consider the forward Euler scheme

$$x_k = x_{k-1} + \Delta t \cdot f(x_{k-1}) \tag{2}$$

and analyze the first integration step; i.e.,

$$x_1 = x_0 + \Delta t \cdot f(x_0), \tag{3}$$

where $x_0 = x_i$. Substitute the initial value with $[x_0] = [x_i] = x_i^0 + \delta x_i$ in Eq. 3 for

$$[x_1] = [x_0] + \Delta t \cdot f([x_0]). \tag{4}$$

If the function f is evaluated in the DA framework, the output of the first step, $[x_1]$, is the Taylor expansion of the solution x_1 at t_1 with respect to the initial condition about the reference point x_i^0. The previous procedure can be inferred through the subsequent steps until the last integration step is reached. The result at the final step is the n-th order Taylor expansion of the flow of the initial value problem of Eq. 1 at the final time t_f. Thus, the expansion of the flow of a dynamical system can be computed up to order n with a fixed amount of effort.

4 Optimal Control Problem

Suppose the spacecraft moves under the general dynamics

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t), \tag{5}$$

where $\boldsymbol{x} = \{x_1, \ldots, x_v\}$ is the state vector and $\boldsymbol{u} = \{x_1, \ldots, x_m\}$ is the control vector ($m \leq v$). The optimal control problem aims at finding the m control functions $\boldsymbol{u}(t)$ that minimize the performance index

$$J = \varphi(\boldsymbol{x}_f, t_f) + \int_{t_i}^{t_f} L(\boldsymbol{x}(t), \boldsymbol{u}(t), t)\, dt. \tag{6}$$

The initial state vector, $\boldsymbol{x}_i$, and the final state vector, $\boldsymbol{x}_f$, are not necessarily fixed, as well as the final time t_f. In addition to the previous statements, boundary constraints

$$\boldsymbol{\psi}(\boldsymbol{x}_f, t_f) = 0, \tag{7}$$

where $\boldsymbol{\psi} = \{\psi_1, \ldots, \psi_p\}$, and path constraints

$$\boldsymbol{C}(\boldsymbol{u}(t), t) \leq 0, \tag{8}$$

where $\boldsymbol{C} = \{C_1, \ldots, C_q\}$, can be imposed.

The above problem can be solved by reducing it to a boundary value problem on a set of differential algebraic equations (DAEs) [5]. To this aim, the dynamics and constraints are added to the performance index J to form the so-called augmented performance index

$$\begin{aligned} \bar{J} \;=\; & \varphi(\boldsymbol{x}_f, t_f) + \boldsymbol{\nu}^T \boldsymbol{\psi}(\boldsymbol{x}_f, t_f) + \\ & + \textstyle\int_{t_i}^{t_f} [L(\boldsymbol{x}, \boldsymbol{u}, t) + \boldsymbol{\lambda}^T (\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t) + \\ & - \dot{\boldsymbol{x}}) + \boldsymbol{\mu}^T \boldsymbol{C}(\boldsymbol{u}, t)] dt, \end{aligned} \tag{9}$$

where two kind of Lagrange multipliers are introduced:

- a p-dimensional vector of constants, $\boldsymbol{\nu}$, for the final constraints of Eq. (7);
- an n-dimensional and a q-dimensional vector of functions $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ for the dynamics of Eq. (5) and the path constraints of Eq. (8), which are usually referred to as adjoint or costate variables.

The optimal control problem is then reduced to identifying a stationary point of the augmented performance index $\bar{J}$. This is achieved by imposing the gradient of $\bar{J}$ to be zero with respect to all optimization variables; specifically, the state vector $\boldsymbol{x}$ and the control vector $\boldsymbol{u}$, the Lagrange multipliers $\boldsymbol{\nu}$ and the costate variables $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, the unknown components of the initial state $\boldsymbol{x}_i$ and the final state $\boldsymbol{x}_f$, and the final time t_f. In particular, the optimality with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{x}$ leads to the following relations:

$$\begin{aligned} & \frac{\partial \bar{J}}{\partial \boldsymbol{\lambda}} = 0 \;\Rightarrow \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t) \\ & \frac{\partial \bar{J}}{\partial \boldsymbol{x}} = 0 \;\Rightarrow \dot{\boldsymbol{\lambda}} = -\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right)^T \boldsymbol{\lambda} - \left(\frac{\partial L}{\partial \boldsymbol{x}}\right)^T, \end{aligned} \tag{10}$$

whereas $\partial \bar{J}/\partial \boldsymbol{u} = 0$ yields

$$\left(\frac{\partial L}{\partial \boldsymbol{u}}\right)^T + \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}\right)^T \boldsymbol{\lambda} + \left(\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{u}}\right)^T \boldsymbol{\mu} = \boldsymbol{0}. \tag{11}$$

Equations (10) and (11) together are usually referred to as Euler-Lagrange equations. It is

worth observing that the Euler-Lagrange equations form a system of DAEs: the differential part is represented by Eq. (10), which defines the dynamics for the state variables $\boldsymbol{x}$ and the costate variables $\boldsymbol{\lambda}$; the role of the algebraic constraint is played by Eq. (11). The previous system must be coupled with the boundary conditions ensuing from the optimality conditions with respect to the remaining optimization variables (see [5] for further details). The optimal control problem is eventually reduced to a boundary value problem on a system of DAEs.

A particular optimal control problem is addressed in this work. The dynamics is supposed to be affine in the control vector $\boldsymbol{u}$; i.e.

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t) = \tilde{\boldsymbol{f}}(\boldsymbol{x}, t) + \mathrm{B}(\boldsymbol{x})\, \boldsymbol{u}, \tag{12}$$

where $\mathrm{B}(\boldsymbol{x})$ is a $v \times m$ matrix, whose elements do not depend on the controls. Moreover the control functions are sought to minimize the performance index

$$J = \frac{1}{2} \int_{t_i}^{t_f} \boldsymbol{u}^T \boldsymbol{u}\, dt \tag{13}$$

and no path constraints are imposed. Based on the previous hypotheses, Eq. (11) assumes the simpler form

$$\boldsymbol{u} + \mathrm{B}^T(\boldsymbol{x})\, \boldsymbol{\lambda} = 0. \tag{14}$$

Equation (14) supplies an explicit relation between the control functions $\boldsymbol{u}$ and the costate variables $\boldsymbol{\lambda}$, which can be substituted in Eq. (10). The original system of DAEs of the Euler-Lagrange equations translates into the system of ODEs

$$\begin{aligned} \dot{\boldsymbol{x}} &= \tilde{\boldsymbol{f}}(\boldsymbol{x}, t) - \mathrm{B}(\boldsymbol{x})\, \mathrm{B}^T(\boldsymbol{x}) \boldsymbol{\lambda} \\ \dot{\boldsymbol{\lambda}} &= -\left(\frac{\partial \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\lambda}, t)}{\partial \boldsymbol{x}} \right)^T \boldsymbol{\lambda}. \end{aligned} \tag{15}$$

Therefore, the original optimal control problem reduces to a two-point boundary value problem (TPBVP) on the set of ODEs in Eq. (15), where boundary conditions are imposed on the initial and final values of the state and costate variables, depending on the optimal control problem at hand.

5 High Order Optimal Feedback

Suppose the problem of transferring a spacecraft from a fixed initial state to a fixed final state with fixed t_i and t_f is of interest; i.e., boundary conditions assume the simpler form

$$\begin{cases} \boldsymbol{x}_i &= \overline{\boldsymbol{x}}_i \\ \boldsymbol{x}_f &= \overline{\boldsymbol{x}}_f. \end{cases} \tag{16}$$

The optimal control problem is then reduced to the problem of solving Eq. (15) subject to the boundary conditions in Eq. (16).

Several techniques are available in the literature to solve the previous problem for assigned $\overline{\boldsymbol{x}}_i$ and $\overline{\boldsymbol{x}}_f$, like the simple and multiple shooting schemes or difference methods [9]. This means that, given $\overline{\boldsymbol{x}}_i$ and $\overline{\boldsymbol{x}}_f$, the previous techniques are applied to compute the initial values of the costate variables that solve the TPBVP, which will be indicated as $\boldsymbol{\lambda}_i^0$. The solution is then uniquely identified by the initial state and costate vectors, $\overline{\boldsymbol{x}}_i$ and $\boldsymbol{\lambda}_i^0$ respectively.

Assume now a reference solution $\boldsymbol{\lambda}_i^0$ is available and suppose the Taylor expansion of the solution of the optimal control problem with respect to the initial state $\boldsymbol{x}_i$ is of interest. Differential algebra can effectively serve this purpose. To this aim, initialize both the initial state $\boldsymbol{x}_i$ and the initial costate $\boldsymbol{\lambda}_i$ as DA variables. This means the variations

$$\begin{aligned} [\boldsymbol{x}_i] &= \overline{\boldsymbol{x}}_i + \delta \boldsymbol{x}_i \\ [\boldsymbol{\lambda}_i] &= \boldsymbol{\lambda}_i^0 + \delta \boldsymbol{\lambda}_i \end{aligned} \tag{17}$$

to the fixed initial state $\overline{\boldsymbol{x}}_i$ and the reference solution $\boldsymbol{\lambda}_i^0$ are considered.

Using the techniques introduced in Sect. 3, the solution of Eq. (15) is expanded with respect to the initial state and costate vectors. More specifically, the dependence of the final state and costate vectors on their initial values are obtained in terms of the high order polynomial map

$$\begin{aligned} \begin{pmatrix} [\boldsymbol{x}_f] \\ [\boldsymbol{\lambda}_f] \end{pmatrix} &= \begin{pmatrix} \overline{\boldsymbol{x}}_f + \delta \boldsymbol{x}_f \\ \boldsymbol{\lambda}_f^0 + \delta \boldsymbol{\lambda}_f \end{pmatrix} \\ &= \begin{pmatrix} \overline{\boldsymbol{x}}_f \\ \boldsymbol{\lambda}_f^0 \end{pmatrix} + \begin{pmatrix} \mathcal{M}_{\boldsymbol{x}_f} \\ \mathcal{M}_{\boldsymbol{\lambda}_f} \end{pmatrix} \begin{pmatrix} \delta \boldsymbol{x}_i \\ \delta \boldsymbol{\lambda}_i \end{pmatrix}, \end{aligned} \tag{18}$$

where $\overline{\boldsymbol{x}}_f$ and $\boldsymbol{\lambda}_f^0$ are the constant part of the map (i.e., the reference solution flowing from $\overline{\boldsymbol{x}}_i$

and $\boldsymbol{\lambda}_i^0$ under the ODEs in Eq. (15)), whereas all higher order terms are included in $\mathcal{M}_{\boldsymbol{x}_f}$ and $\mathcal{M}_{\boldsymbol{\lambda}_f}$.

Subtract now the constant part from Eq. (18) for

$$\begin{pmatrix} \delta\boldsymbol{x}_f \\ \delta\boldsymbol{\lambda}_f \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{\boldsymbol{x}_f} \\ \mathcal{M}_{\boldsymbol{\lambda}_f} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{x}_i \\ \delta\boldsymbol{\lambda}_i \end{pmatrix}. \tag{19}$$

Then, extract $\mathcal{M}_{\boldsymbol{x}_f}$ from Eq. (19) and consider the map

$$\begin{pmatrix} \delta\boldsymbol{x}_f \\ \delta\boldsymbol{x}_i \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{\boldsymbol{x}_f} \\ \mathcal{I}_{\boldsymbol{x}_i} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{x}_i \\ \delta\boldsymbol{\lambda}_i \end{pmatrix}, \tag{20}$$

which is built by concatenating $\mathcal{M}_{\boldsymbol{x}_f}$ with the identity map for $\delta\boldsymbol{x}_i$, $\mathcal{I}_{\boldsymbol{x}_i}$.

Using suitable inversion techniques for high order polynomials [3], the map in Eq. (20) can be inverted to obtain

$$\begin{pmatrix} \delta\boldsymbol{x}_i \\ \delta\boldsymbol{\lambda}_i \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{\boldsymbol{x}_f} \\ \mathcal{I}_{\boldsymbol{x}_i} \end{pmatrix}^{-1} \begin{pmatrix} \delta\boldsymbol{x}_f \\ \delta\boldsymbol{x}_i \end{pmatrix}. \tag{21}$$

The high order polynomial map in Eq. (21) relates the displacements of the initial state and costate vectors from their reference values $\overline{\boldsymbol{x}}_i$ and $\boldsymbol{\lambda}_i^0$, $\delta\boldsymbol{x}_i$ and $\delta\boldsymbol{\lambda}_i$ respectively, to the displacement of the final state vector from its fixed value $\overline{\boldsymbol{x}}_f$, $\delta\boldsymbol{x}_f$, and again $\delta\boldsymbol{x}_i$.

Suppose now the problem of reaching the fixed final state $\overline{\boldsymbol{x}}_f$ regardless of any error on the initial state $\delta\boldsymbol{x}_i$ is of interest. This means the new solution $\boldsymbol{\lambda}_i$ of the optimal control problem corresponding to the new initial condition $\boldsymbol{x}_i = \overline{\boldsymbol{x}}_i + \delta\boldsymbol{x}_i$ is to be computed. The map in Eq. (21) can effectively serve this purpose. More specifically, the boundary condition

$$\boldsymbol{x}_f = \overline{\boldsymbol{x}}_f \tag{22}$$

must be imposed. To this aim, note that $\boldsymbol{x}_f = \overline{\boldsymbol{x}}_f + \delta\boldsymbol{x}_f$. Consequently, Eq. (22) reduces to

$$\delta\boldsymbol{x}_f = 0. \tag{23}$$

Substituting Eq. (23) into the high order polynomial map of Eq. (21) yields

$$\begin{pmatrix} \delta\boldsymbol{x}_i \\ \delta\boldsymbol{\lambda}_i \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{\boldsymbol{x}_f} \\ \mathcal{I}_{\boldsymbol{x}_i} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \delta\boldsymbol{x}_i \end{pmatrix}. \tag{24}$$

Extract the last v components of the map in Eq. (24), which will be indicated as

$$\delta\boldsymbol{\lambda}_i = \mathcal{M}_{\boldsymbol{\lambda}_i}(\delta\boldsymbol{x}_i). \tag{25}$$

The polynomial map in Eq. (25) is the high order Taylor expansion of the solution of the optimal control problem with respect to the initial state $\boldsymbol{x}_i$: given any displacement $\delta\boldsymbol{x}_i$ of $\boldsymbol{x}_i$ from the reference value $\overline{\boldsymbol{x}}_i$, the mere evaluation of the polynomials in Eq. (25) delivers the high order correction $\delta\boldsymbol{\lambda}_i$ to $\boldsymbol{\lambda}_i^0$ to obtain the corresponding solution $\boldsymbol{\lambda}_i$ of the optimal control problem.

It is worth observing that a possible alternative approach to solve the previous problem could have consisted in solving the TPBVP for the new solution $\boldsymbol{\lambda}_i$ using classical techniques. However, the main disadvantage of this approach is that a new TPBVP must be solved for each displaced initial condition. This involves running through the iterative procedures of the classical TPBVP solvers. Each iterative procedure is able to deliver one solution, whose validity is limited to the corresponding $\delta\boldsymbol{x}_i$. Consequently, the classical TPBVP solvers should be applied for each new $\delta\boldsymbol{x}_i$. The Taylor expansion of the optimal control problem supplies an effective alternative method to overcome this issue. First of all, analytical information are gained, which can supply a valuable insight on the underlying dynamics. Moreover, for any error $\delta\boldsymbol{x}_i$, the mere evaluation of polynomials suffices to obtain the new optimal control law to reach $\overline{\boldsymbol{x}}_f$, so avoiding the use of iterative algorithms. Nevertheless, the polynomial relation between $\delta\boldsymbol{\lambda}_i$ and $\delta\boldsymbol{x}_i$ given by Eq. (25) is accurate up to the order of the DA-based computation.

6 Lunar Landing

The technique introduced in Sect. 5 is used here for the optimal feedback control of a probe landing on Moon's south pole. The control profile is designed in the frame of the controlled two-body problem. Referring to Figure 2, the lander is supposed to originally move on an elliptical polar descent orbit, taking it from an altitude of 100 km (apocenter) to an altitude of 20 km (pericenter). The landing phase is supposed to start at the pericenter of the descent orbit. Final conditions are imposed to position the lander over Moon's south pole at an altitude of 2 m, with a downward velocity of 3 m/s, from which the final phase of the landing maneu-

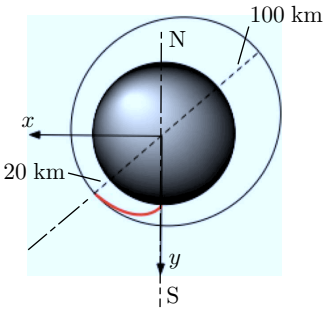


Figure 2: Lunar landing problem.

ver is supposed to start. A cartesian reference frame is selected to describe the dynamics: the y-axis is aligned with Moon's south pole; the x-axis lies on Moon's equatorial plane, pointing towards the orbital descending node; the z-axis is selected to form a right-handed reference system. The landing dynamics is described by the set of ODEs:

$$\begin{aligned} \dot{\boldsymbol{r}} &= \boldsymbol{v} \\ \dot{\boldsymbol{v}} &= -\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{u}, \end{aligned} \tag{26}$$

in which $\boldsymbol{r}$ and $\boldsymbol{v}$ are the probe position and velocity, respectively; $r = ||\boldsymbol{r}||$; μ is Moon's gravitational parameter; and $\boldsymbol{u}$ is the control vector. As from Eq. (26), the dynamics is affine in the control vector $\boldsymbol{u}$. Thus, Eq. (15) holds for the problem at hand and the optimal control problem is then reduced to a TPBVP with fixed initial and final states for the landing probe.

A reference solution of the optimal control problem is first identified by solving the resulting TPBVP. The initial time is chosen to be zero, whereas the landing duration is set to 31 min. A simple shooting technique is used to solve the TPBVP and to compute the reference trajectory reported in Fig. 3. Figure 4 illustrates the corresponding reference control profile in terms of histories of its components. Due to the symmetry of the problem, the reference trajectory lies completely on the x-y plane.

The initial probe position and velocity, $\boldsymbol{r}_i$ and $\boldsymbol{v}_i$ respectively, are now supposed to be affected by errors. The high order optimal feedback control algorithm introduced in Sect. 5 is applied to optimally correct the reference control law. More specifically, the reference trajectory in Fig. 3 is used as reference solution for the Taylor expansions. The algorithm is then applied to compute the high order polynomial map of Eq. (25) for the problem at hand. Thus, given any displacement $\delta\boldsymbol{x}_i = (\delta\boldsymbol{r}_i, \delta\boldsymbol{v}_i)$ of the initial state vector from its nominal value, the polynomial map is readily evaluated to correct the nominal $\boldsymbol{\lambda}_i^0$ for the new initial value of the costate vector.

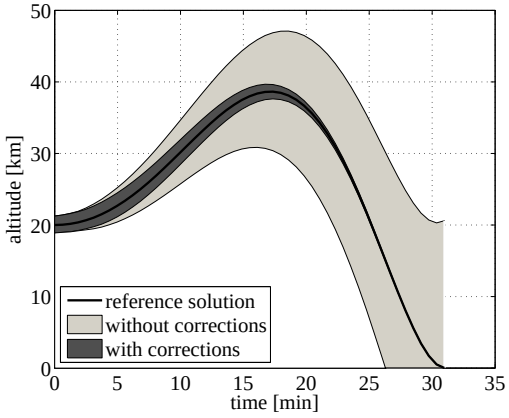


Figure 3: Lunar landing: altitude dispersion.

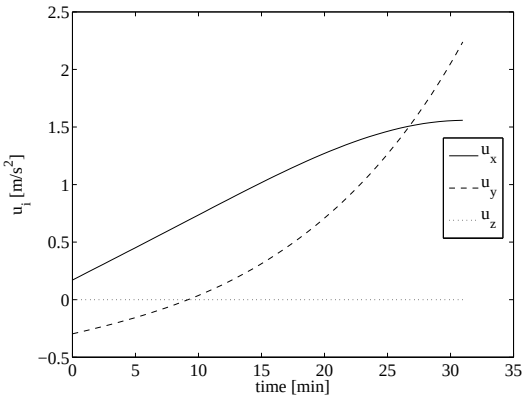


Figure 4: Lunar landing: reference control.

The performances of the procedure are studied in the followings. A maximum position error of 1 km and a maximum velocity error of 5 m/s are supposed to affect each component of the initial lander position and velocity, respectively. The final dispersion at landing is then investigated. First of all, for the sake of a more complete analysis, given any displaced initial conditions, no corrections to the nominal costate variables (and, consequently, to the controls) are supplied. In particular, 100 samples

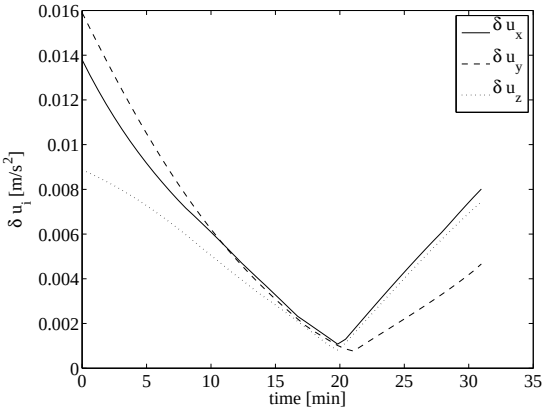


Figure 5: Lunar landing: control corrections.

are randomly generated within the initial uncertainty box with uniform distribution. Each sample is then propagated using the nominal guidance law; once all integrations have been performed, the maximum and minimum lander altitudes at each integration time are computed. Figure 3 shows the resulting altitude dispersion throughout landing. The figure illustrates how initial conditions corresponding to both impacts on Moon's surface (lower area of the strip) and trajectories moving away from the landing site (higher area of the strip) are included in the initial error box.

The DA-based high order algorithm described in Sect. 5 is then applied. In particular, third order corrections are computed using Eq. (25): for the same random samples of Fig. 3, the errors on the initial state are computed and the map is evaluated to correct the reference $\boldsymbol{\lambda}_i^0$. The resulting set of trajectories is reported again in Fig. 3 for the sake of comparison. The corrected optimal control laws take the probe to the final desired conditions and the resulting final dispersion is drastically reduced. The necessary control corrections are analyzed in Fig. 5. More specifically, for each component of the control vector $\boldsymbol{u}$, the maximum control correction is evaluated among the random samples, and the resulting curves are reported in figure. A maximum control correction of about 0.016 m/s^2 is required for the given error box.

7 Rendezvous Maneuver

A rendezvous maneuver is here analyzed as a further test case for the high order optimal feedback technique introduced in Sect. 5. The study of this problem is motivated by the work of Park, Guibout and Scheeres based on the alternative approach of generating functions [7, 8]. The space rendezvous is a maneuver which takes two spacecraft, originally moving on different orbits, to the same final reference orbit, matching their positions and velocities. Referring to Figure 6, this rather general case can be focused on the problem of a spacecraft (referred to as *chaser*) targeting an object (referred to as *target*) on its orbit.

A continuously propelled rendezvous maneuver is considered. The target is supposed to move on a circular orbit of radius R, whereas the chaser is assumed to be subject to a controlled two-body dynamics. In this framework, the rendezvous maneuver is classically designed in a non-inertial reference frame that is centered at the target position, with x-axis constantly aligned with the orbital radius, y-axis directed towards the target orbital velocity, and z-axis chosen to form a right-handed coordinate system with x and y (see Figure 6). Thus, the non-inertial reference frame rotates along the circular target orbit with constant angular velocity ω and the chaser is subject to the relative dynamics

$$\begin{aligned} \dot{x} &= v_x, \quad \dot{y} = v_y, \quad \dot{z} = v_z \\ \dot{v}_x &= 2\dot{y} - (1+x)(\frac{1}{r^3} - 1) + u_x \\ \dot{v}_y &= -2\dot{x} - y(\frac{1}{r^3} - 1) + u_y \\ \dot{v}_z &= -\frac{1}{r^3} z + u_z \,, \end{aligned} \tag{27}$$

where lengths and time are normalized using R and $1/\omega$ respectively; $\boldsymbol{u} = (u_x, u_y, u_z)$ is the control vector; and $r = \sqrt{(1+x)^2 + y^2 + z^2}$.

The chaser is supposed to have initial offsets from the target in both position and velocity, which are denoted by $\delta\boldsymbol{r}_i$ and $\delta\boldsymbol{v}_i$ respectively. The optimal control problem is solved to design the control functions $\boldsymbol{u}$ that take the chaser from its initial displaced state to the target state (i.e.,

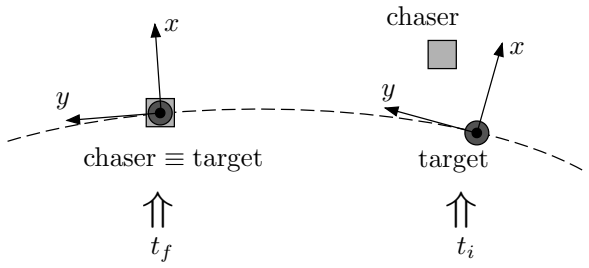


Figure 6: Rendezvous maneuver.

to the origin of the rotating frame with zero velocity) in a given time $t_f - t_i$. The relative dynamics in Eq. (27) is affine in the control vector $\boldsymbol{u}$. Thus, the optimal control problem can be reduced to a TPBVP with fixed initial and final states for the chaser.

Similarly to the previous test case, a nominal solution of the optimal control problem must be identified before applying the high order DA-based technique. To this aim, it is worth observing that the relative dynamics in Eq. (27) satisfies $\boldsymbol{f}(\,\boldsymbol{x} = 0, \boldsymbol{u} = 0, t\,) = 0$, with $\boldsymbol{x} = (x, y, z, v_x, v_y, v_z)$. This means that $\boldsymbol{x}(t) = 0$ and $\boldsymbol{u}(t) = 0$ for any t is a trivial solution of the optimal control problem that is used as reference solution for the high order expansion.

The performances of the high order optimal feedback control algorithm are now investigated. The chaser is supposed to have a displaced initial position $\delta\boldsymbol{r}_i = (0.2, 0.2, 0)$ and a displaced initial velocity $\delta\boldsymbol{v}_i = (0.1, 0.1, 0)$. The rendezvous maneuver is designed to take the chaser to the target state in 1 time unit. The exact solution of the optimal control problem is first identified by solving the associated TPBVP using a simple shooting technique. The result is reported in Fig. 7, Fig. 8, and Fig. 9 in terms of position, velocity, and control profile, respectively. As can be seen from Fig. 7 and Fig. 8, the exact solution (solid lines) takes the chaser to the target state in the assigned time. The exact solution is then compared with those achieved by the DA-based optimal feedback control algorithm introduced in Sect. 5 using different expansion orders. As can be seen, the low accuracy of the 1-st order correction is significantly improved using 4-th and 6-th order expansions.

The main advantage of the high order opti-

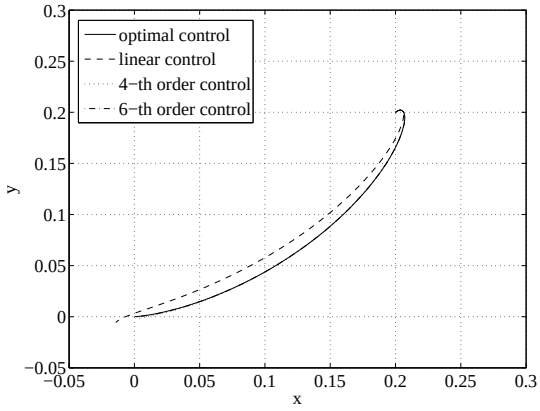


Figure 7: Rendezvous: position.

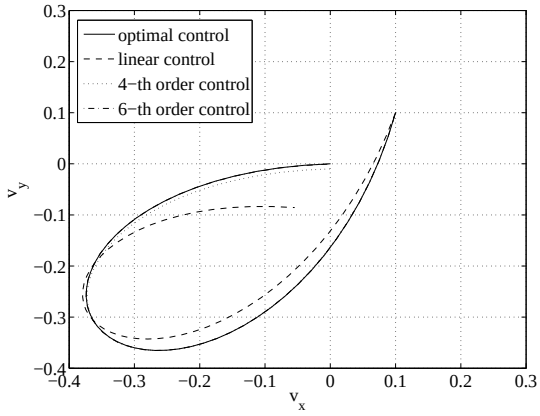


Figure 8: Rendezvous: velocity.

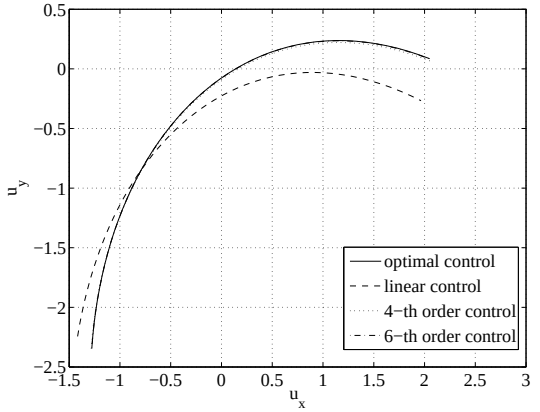


Figure 9: Rendezvous: control.

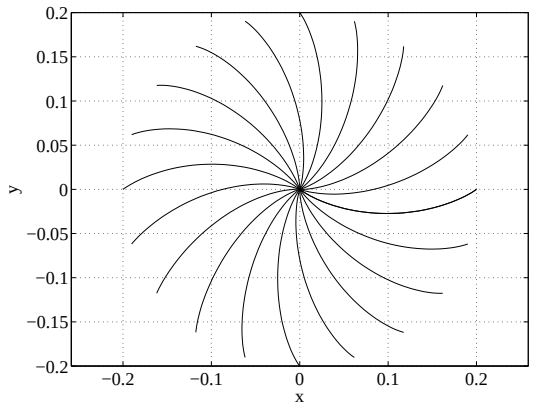


Figure 10: Rendezvous: trajectories corresponding to displaced initial positions lying on a circle of radius 0.2.

mal feedback control algorithm is that, for any initial offset of the chaser with respect to the target, the same polynomial map is evaluated to compute the corresponding optimal control law. This means that the high order map in Eq. (25) must be computed only once for all possible offsets, and the optimal control laws are then obtained through the mere evaluation of polynomials. This feature is exploited in Fig. 10: a set of displaced positions distributed over a circle of radius 0.2 in the rotating frame is selected. For each sample, a 6-th order correction is computed using the polynomial map of Eq. (25). As can be seen, the chaser is always moved to the origin of the reference frame.

8 Conclusion

A method for the computation of optimal feedback control laws based on differential algebra has been introduced, with applications to lunar landing and rendezvous maneuvers. The method relies on the high order expansion of the solution of the optimal control problem about a reference trajectory. Thus, it improves the results of classical techniques based on the linearization of the dynamics. Moreover, the method reduces the computation of new optimal control laws to the mere evaluation of polynomials. This is a valuable advantage over the conventional nonlinear optimal control strategies, which are mainly based on iterative procedures. This work focused on the problem of transferring a spacecraft from an initial fixed state to a final fixed state. Further developments will investigate the imposition of soft constraints, as well as the minimization of alternative performance criteria. Moreover, path constraints on the controls will be included.

References

[1] Beeler, S.C., "State-Dependent Riccati Equation Regulation of Systems with State and Control Nonlinearities", NASA, 2004.

[2] Berz, M., "The new method of TPSA algebra for the description of beam dynamics to high orders", Los Alamos National Laboratory, 1986.

[3] Berz, M., *Modern Map Methods in Particle Beam Physics*, Academic Press, 1999.

[4] Berz, M., and Makino, K., "COSY INFINITY version 9 reference manual", Michigan State University, 2006.

[5] Bryson, A.E., Ho, Y.C., *Applied Optimal Control*, Hemisphere Publishing Co., Washington, 1975.

[6] Cimen, T., and Banks, S. P., "Nonlinear Optimal Tracking Control with Application to Super-Tankers for Autopilot Design", *Automatica*, Vol. 40, 2004, pp. 1845-1863.

[7] Park, C., Guibout, V., and Scheeres, D., "Solving Optimal Continuous Thrust Rendezvous Problems with Generating Functions", *Journal of Guidance, Control, and Dynamics*, Vol. 29, 2006, pp. 321-331.

[8] Park, C., and Scheeres, D., "Solution of Optimal Feedback Control Problems with General Boundary Conditions Using Hamiltonian Dynamics and Generating Functions", *Automatica*, Vol. 42, 2006, pp. 869-875.

[9] Stoer, J., and Bulirsch, R., *Introduction to Numerical Analysis*, Springer, New York, 1993.

CEAS 2011

Renewable Energy and Energy Harvesting

Green design: How to assess environmental benefits in air transport?

R. Ilg and H. Krieg
Dept. Life Cycle Engineering (GaBi), Chair of Building Physics (LBP), University of Stuttgart

Hauptstrasse 113, 70771 Echterdingen, Germany, Telephone: +49-(0)711-48 99 99 22

Robert.Ilg@lbp.uni-stuttgart.de

Keywords: *Life Cycle Assessment (LCA), Carbon Footprint, Environmental Assessment, Green Design, Aviation Supply Chain Management*

Abstract

Life Cycle Assessment (LCA) is a useful technique for assessing the environmental aspects and potential impacts associated with a product by first compiling an inventory of relevant inputs and outputs of a product system; second to evaluating the potential environmental impacts associated with those inputs and outputs and third to interpret the results of the inventory analysis and impact assessment phases in relation to the objectives of the study.

LCA studies, according to the EN ISO 14044 methodology studies the environmental aspects and potential impacts throughout the product's life (i.e. cradle to grave) from raw materials acquisition through production, use and disposal. Figure 1 shows an overall life cycle of an aircraft, taking into account all phases and supply chains, starting from the raw material extraction until the end-of-life.

The paper deals with the methodological basis and shows an example of an on-going EU project, how the approach is applied and environmental impacts in the aviation sector are quantified.

1 General Introduction

1.1 Life Cycle Assessment (LCA) methodology

The environmental evaluation is based on the methodology of Life Cycle Assessment (LCA). An overview is given on the methodology behind LCA according to EN ISO 14040 [1] and EN ISO 14044 [2]. The LCA framework is set by the 4 core pillars "Goal and Scope definition", "Life Cycle Inventory Analysis (LCI), "Life Cycle Impact Assessment (LCIA)" and the "Interpretation" (see Fig. 1).

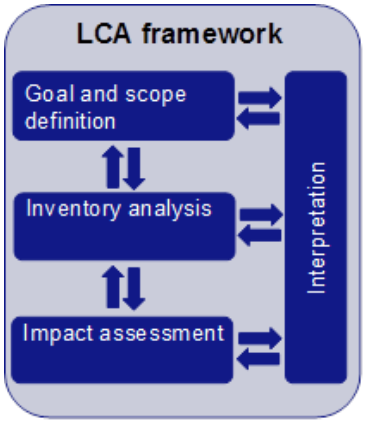


Fig. 1 Main steps of an LCA according to [2]

Furthermore a LCA includes, according to the ISO 14044 [2] the environmental aspects and potential impacts throughout the product's life (i.e. cradle to grave) from raw materials extraction through production, use and disposal.

Concerning the entire life cycle, optimization potentials, by shifting problems into another life cycle phase, can be detected and avoided.

LCA) is a suitable tool for analyzing and assessing environmental impacts caused through production, use and disposal of products or product systems for specific applications. Depending on the goal and scope of a LCA study, LCA does not produce clear-cut straightforward assertions, but rather it gives diverse and complex results. It supports the process of decision-making by making complex issues transparent [1].

1.2 Life Cycle Assessment (LCA) software tools and databases

The availability of an appropriate set of tools, practical and adopted to a day to day application, meeting the individual requirements is indispensable to strengthen and support the use, benefit and acceptance of LCA as well as LCI in the aviation industry. The goal is to empower the applicability and feasibility of LCA/LCI information in practice by addressing the specific questions to be answered within industry – simplified and focused in application, but without restriction in quality of data and results. As mentioned in the chapter before, the environmental evaluation is based on various data coming from various sectors and calculation procedures for the results (e.g. various LCIA methods), therefore the use of an appropriate database and tool to conduct a LCA is indispensable. There exist various databases and tools to conduct an LCA (e.g. GaBi, Simapro, etc.).

To meet the variety in application of LCA-/LCI-based information, an integrated tool was used – the so-called GaBi 4 software and databases [3]. The software system GaBi 4 was developed in close cooperation with experts from industry and also uses datasets provided by the European Commission on their platform (further information can be found on: [4])

1.3 The EU-Project NICETRIP

The NICETRIP project specifically addresses the acquisition of new knowledge and technology validation concerning the tilt-rotor and is fully relevant to the strategic objective 1.3.2.1: "Integration of technologies towards the future tilt-rotor aircraft", of the work programs of call 3 "Thematic Priority Aeronautics and Space". The main project objectives are:

- To validate the European civil tilt-rotor concept based on the ERICA architecture,
- To validate critical technologies and systems through the development, integration and testing of components of a tilt-rotor aircraft on full-scale dedicated rigs,
- To acquire new knowledge on tilt-rotor through the development and testing of several wind tunnel models, including a large-scale full-span powered model,
- To investigate and evaluate the introduction of tilt-rotors in the European Air Traffic Management System,
- To assess the sustainability of the tilt-rotor product with respect to environmental issues and to define the path towards a future tilt-rotor flying demonstrator.

Especially the assessment of the sustainability of the tilt-rotor system focusing on the environmental performance is considering future perspectives and develops recommendations, strategies and plans for future work on the way to a flying prototype. Achieving these objectives is the projects driving goal.

The aims of the overall sustainability task are to provide a holistic picture and a reasonable environmental evaluation of the rotorcraft system that includes total life cycle considerations. The tilt-rotor concept and the production phase of the rotorcraft system will be analyzed and environmental key parameters identified. A sustainability assessment will be done, based on the environmental key

parameters, e.g. emissions or energy consumption, to identify the environmental processes of relevance ("hot spots") over the life cycle of the considered rotorcraft system.

2 Systematic Analysis of the air transport system, identification of optimization potentials

As explained in chapter 1, the life cycle contains the 3 phases production, use and end-of-life of the air transport system, in this case the tilt-rotor system given by the EU project NICETRIP. This paper focuses on the production phase and excludes therefore the use and end-of-life phase, as it first will serve as an overview on the general methodology to quantify environmental impacts in the aviation sector (see Fig. 3). Further papers are planned to get a holistic LCA picture for the tilt-rotor system by including all life cycle phases.

2.1 Definition of the air transport system

Within this section, the production of the tilt-rotor system, including the raw material process chain, is given by highlighting the data collection and modeling steps resulting in the overall impacts related to the production of the tilt-rotor system. It also deals with assumptions made, as for some components only limited information was available.

To assess the environmental impacts of the production of the NICETRIP aircraft, all information on production processes was modelled. In a first step, data on materials, processes and technologies was gathered from companies participating in the EU-funded project. The information from those questionnaires was used to create a model in the GaBi LCA software. Based on this model, the environmental impacts of the production of the NICETRIP tilt-rotor were evaluated.

The preliminary design of the vertical take-off and landing aircraft is similar to Fig. 2.



Fig. 2 Preliminary design of the tilt-rotor aircraft [5]

The specifications and aircraft characteristic used for the design are summarized in Table 1. The design empty gross weight is given by 7070 kg [5]. The design values of the tilt-rotor aircraft are separated into the STOL (Short Take-off or Landing Aircraft) and VTOL performance values (Vertical Take-off or Landing Aircraft) (see Table 1).

Table 1: Main design characteristics of the tilt-rotor aircraft [5]

Range	1200 km
Cruise altitude	7500 m
Empty gross weight	7070 kg
Gross weight-VTOL	10870 kg
Alternate gross weight STOL	11570 kg
Payload-VTOL	19 passengers with luggage
Payload-STOL	22 passengers with luggage
Cruise speed-VTOL mission	300 knots (around 556km/h)
Cruise speed-STOL mission	330 knots (around 610km/h)
Operative range-VTOL	650 Nautical Miles (NM)
Operative range-STOL	820 Nautical Miles (NM)

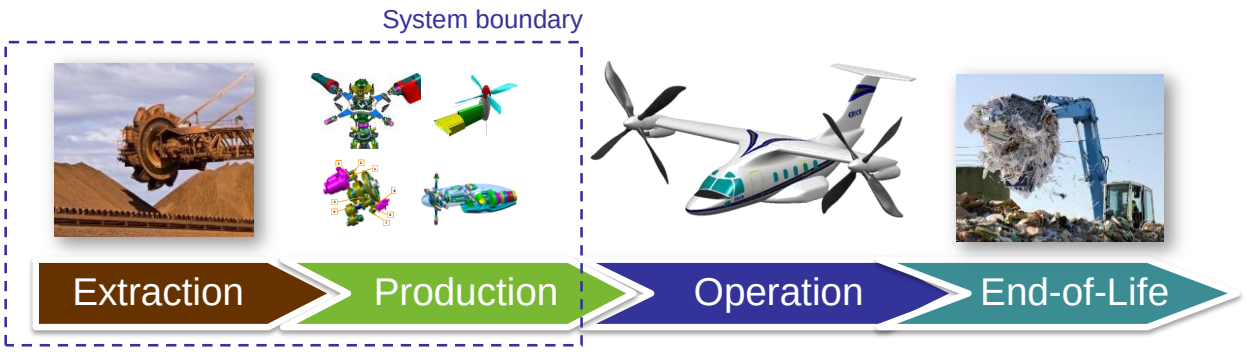


Fig. 3 Life cycle of a tilt-rotor aircraft

Due to the complexity of the tilt-rotor system the production was characterized into 3 main modules:

- Airframe,
- Powerplant and
- Auxiliary System.

The single components are modeled as detailed as possible, taking into account the manufacturing processes used to produce the respective part of the tilt-rotor aircraft.

The functional unit used for the study was the production of the tilt-rotor aircraft with the empty gross weight of 7337kg, which differs from the given weight of 7070kg in Table 2, as during the NICETRIP project, various parts were updated and recalculated.

2.2 Life Cycle Inventory of the tilt rotor system

The relative weight breakdown of the individual modules and sub-modules based on the whole tilt-rotor aircraft is shown in Table 2.

The airframe accounts for a share of 37% to the total weight of a tilt-rotor aircraft as well as the powerplant of 40% and the auxiliary systems of 24%.

It can be seen that the major sub-modules of the module 1 “Airframe”, considering the weight of the sub-modules, are the fuselage and the wing structure. The main sub-modules within the module 2 “Powerplant” module are the drive system and the engines, followed by the rotor hubs. Roughly spoken an almost equal distribution is given for the sub-modules of the module 3 “Auxiliary Systems”.

The data on weight, materials and processes for 1497 kg of the tilt-rotor aircraft, which means for 20% of the tilt-rotor aircraft, was given by the companies. The data on weight and materials for 2930 kg (40%) was also collected by the companies. For the remaining 40%, which is 2910 kg, only the data on weight was given. Finishing of the aircraft was not regarded, as no information on the composition was available.

Table 2: Data collection of the tilt-rotor aircraft

Module	Sub-module	Weight [kg]	Data availability
1. Airframe	1.1 Fuselage	1150	Data on weight and materials
	1.2 Wing	717	Data on weight, materials & processes
	1.3 Tail plane surfaces	200	Data on weight and materials
	1.4 Nacelles	320	Data on weight and materials
	1.5 Landing gear	300	Data on weight
2. Power-plant	2.1 Rotor hubs	780	Data on weight, materials & processes
	2.2 Rotor blades	360	Data on weight and materials
	2.3 Rotating controls	200	Data on weight and materials
	2.4 Drive system	700	Data on weight and materials
	2.5 Engines	700	Data on weight
	2.6 Fuel system	120	Data on weight
	2.7 APU	60	Data on weight
3. Auxiliary Systems	3.1 Flight control system	200	Data on weight
	3.2 Hydraulic system	300	Data on weight
	3.3 Electrical system	400	Data on weight
	3.4 Avionics	200	Data on weight
	3.5 Furnishing & equipment.	450	Data on weight
	3.6 Seats	180	Data on weight
Sum	-	**7337**	-

2.3 Life Cycle modeling of the tilt rotor system

Fig. 4 gives an overview of the modeled tilt-rotor system. It contains three plans of the production of module 1 “Airframe”, module 2 “Powerplant” and module 3 “Auxiliary System”. The individual modules are here connected with the final product “Tilt-Rotor System”. Such a connection is called “flow” and can be a material flow as well as an energy flow. These flows combine the individual processes, which can be constituted of many processing steps. The main material and energy flows related to the production phase of the tilt-rotor aircraft as well as the assumptions, which are made for the modeling in GaBi 4, are determined and listed in the following sub chapters.

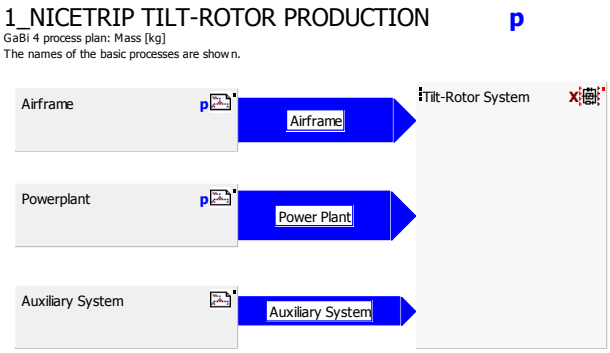


Fig. 4 Modeling of the production phase of a tilt-rotor aircraft in GaBi 4

2.4 Life Cycle Impact Assessment (LCIA) of the tilt rotor system

The following impact categories were analyzed:

- Primary Energy Consumption (PE),
- Global Warming Potential (GWP),
- Acidification Potential (AP),
- Eutrophication Potential (EP),
- Photochemical Ozone Creation Potential (POCP).

The relative share of the three modules of the production (namely airframe, powerplant and auxiliary system) to the impact categories as well as the PE demand caused by the production of the whole tilt-rotor aircraft are illustrated in Fig. 5.

The overall PE demand of the tilt-rotor aircraft production is 3,27E+6 MJ. The main part is related to the production of the airframe (55%), which is mainly made of CFRP and aluminum. The production of the powerplant accounts for a significant share of 28% to the PE demand. The production of the auxiliary system contributes with a share of 17%.

The GWP from the production of the tilt-rotor aircraft is 1,97E+05 kg CO_2-Equivalent. It follows mainly the PE consumption since this impact category is mainly determined by the CO_2 emissions due to fossil fuel combustion. Therefore the biggest part of the CO_2 emissions is caused by the production of the airframe (55%), which is mainly made of aluminum and CFRP. The contributions of the remaining parts to the GWP are 29% from the powerplant and 16% from the auxiliary system.

The production of the tilt-rotor aircraft contributes 609 kg SO_2-Equivalent to AP. The AP is primarily caused by sulphur dioxide emissions; the rest of the acidification potential is almost completely coming from nitrogen oxide emissions. About 47% of the AP is caused by producing the powerplant.

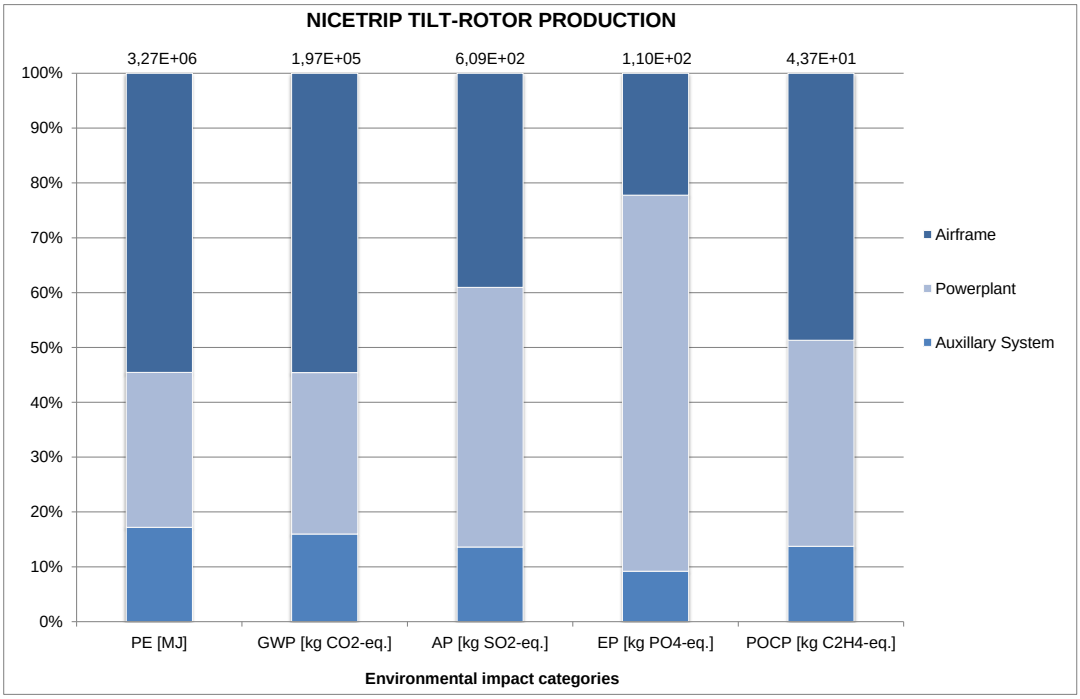


Fig. 5 Environmental impact categories of the NICETRIP aircraft production

The main share of the AP within the powerplant is coming from the drive system and engine, which is mainly made of titanium, aluminum, but also tantalum. 39% of emissions are caused by manufacturing the airframe and most of them are related to the production of the fuselage and the wing. The remaining about 16% of AP is related to the production of the auxiliary system.

The EP from the production of the tilt-rotor aircraft sums up to 110 kg PO_4^{3-} -Equivalents. Most of the EP originates from the emission of nitrogen oxides. Thus 69% are related to the fabrication of the powerplant and more than half of the EP from the powerplant is due to the production of the drive system, which is mainly made of stainless steel. The production of the airframe accounts for a share of 22% to the EP. Producing the auxiliary system contributes with a share of 9% to the AP.

The POCP from the production of the tilt-rotor aircraft accounts for 43,7 kg C_2H_2-Equivalent in total. It is primarily caused by manufacturing the airframe (49%) and almost half of the emissions are related to the production of the fuselage. The production of the powerplant accounts for a share of 38%, 14% are caused by producing the auxiliary system.

The results will show, how a detailed analysis of a system, in this case the tilt-rotor system, analyzed from an environmental point of view, could quantify the hot-spots of a system and recommendations could be given for optimization potentials. Further, more in depth results could be extracted, but this would overdraw the paper.

3 Conclusion

The paper dealt with the production and modeling of the NICETRIP tilt-rotor aircraft. The modeling of the tilt-rotor system was carried out by using the software GaBi 4. The data for the modeling was collected from the participating companies of the NICETRIP project. For some components (around 20% of the total weight), the data on weight, materials and processing technologies were available. For other components (around 40%), data on weight and materials were available. For the remaining components (about 40%), only limited data was available. Therefore, the assumption on materials or processes for these components was carried out based on a literature research.

Due to the flexible and modular setup of the tilt-rotor system in GaBi, it first offers the opportunity to analyze the specific components and modules of a tilt-rotor aircraft. Secondly, several parameters and modules within the tilt-rotor aircraft can be exchanged, for example by replacing individual components with different materials on weight or processing technologies to calculate scenarios or design alternatives.

The tilt-rotor aircraft in this study was mainly made of aluminum (about 30%), CFRP (about 21%), stainless steel (about 17%), titanium (about 8%) and other materials (about 24%, Copper, Rubber etc.). Standard processing technologies for the materials were used, for example casting, die-casting, machining and prepreg-autoclave molding. Since the data on the specific processing technologies of the materials in the aviation sector were not available, profound estimations were used and to some extent, standard processes from other industry sectors in the GaBi database were used.

LCA captures environmental impacts across the entire life cycle, and thus is the ideal approach to have a quantified and sound basis to calculate environmental costs and implement effective solutions on a future path to a sustainable and green aviation transport system.

Acknowledgement

This paper is based on work conducted in the scope of the EU-Project “NICETRIP - Novel Innovative Competitive Effective Tilt Rotor Integrated Project”, co-financed by the European Commission under the contract number AIP5-CT-2006-030944. This contribution is highly appreciated.

References

[1] EN ISO 14040, Environmental management - Life cycle assessment - Principles and framework, 2006.

[2] EN ISO 14044, Environmental management - Life cycle assessment - Requirements and guidelines, 2006.

[3] LBP, PE, GaBi 4 Software System and Databases for Life Cycle Engineering. Copyright, TM. Stuttgart, Echterdingen, Online URL http://www.gabi-software.com/, last access 2011-08-10, 1992-2011.

[4] European Commission - Joint Research Centre (JRC) - Institute for Environment and Sustainability, Documentation and exchange format for the documentation and publication of Life Cycle Inventory (LCI) and Life Cycle Impact Assessment (LCIA) data sets, Online URL http://lca.jrc.ec.europa.eu/lcainfohub/datasetArea.vm, last access 2011-08-16, 1995-2010.

[5] EU-Project NICETRIP, ERICA - General specification and design data, Technical report WP1, Agusta Westland, 2007.

Experiences of energy harvesting using piezoelectric components

Fabio Capece[(1)], Serena Gabriele[(2)], Paolo Gaudenzi[(3)]
[(1)]Ph.D. Student, Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italia
Email: fabio.capece@uniroma1.it

[(2)]Graduated student in Aerospace Engineering at Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italia
Email: serena_gabriele@libero.it

[(3)]Full Professor, Dipartimento di Ingegneria Meccanica e Aerospaziale – Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italia
Phone number: +39 0644585304 Fax: +39 0644585670
Email: paolo.gaudenzi@uniroma1.it

Keywords: *Piezoelectricity, Energy Harvesting*

Abstract

Piezoelectric elements are energy transducers. For their capability to convert mechanical energy into electrical energy, they can be attached on a vibrating structure and used both for active and passive vibration control. Instead of damping the vibration down by closing the electrodes of the piezoelectric element on a resistive load, a circuit capable to accumulate the electrical energy produced can be created.

The activities illustrated in the present paper have a double nature, theoretical and experimental, showing the results obtained with an energy harvesting system which captures energy from the vibrations of a cantilever beam with PZT patches attached in appropriate positions. From the theoretical point of view, mathematical models have been developed to study the energy harvesting system for both the piezoelectric element and the hosting structures.

From the experimental point of view, an extensive campaign has been conducted in order to validate the theoretical models developed in the frame of the work.

1 Introduction

Energy harvesting is the process of acquiring ambient energy from the operational environment and to convert it into electrical energy; this energy can be used to charge storage devices by means of a proper electrical circuit.[1]

The energy accumulated could then be used to provide energy to electrical equipments connected with the storage device or for other purposes. A typical destination for such an energy is to power low and ultra-low power electronics such as MEMS devices. There are many advantages in the employment of this technology: first of all the possibility to recover energy preexisting in the surrounding of the system and to solve the problem of continuously powering electronics in situations where battery substitution is not feasible or, at least, extremely complicated. Furthermore, an energy harvesting system could eventually be an endless source of power.

The conversion of ambient energy can be realized using different kind of transducers, depending on the type of energy sources; within the range of possible opportunities, energy harvesting from structural vibrations can be

successfully accomplished through piezoelectric elements, that can be clearly identified as the major method of power harvesting in such a situation.

Many researchers concentrated on piezoelectric power harvesting technology. Umeda et al.[2] showed and quantified the capacity of power generation from the impact of a steel ball on a piezoelectric element. Elvin et al. [3] verified the ability of accumulating the electrical power produced by the energy harvesting system in order to power a telemetry circuit. Ottman et al. [4] created an adaptive piezoelectric energy harvesting circuit capable of extracting the maximum amount of power from the piezoelectric component to charge a battery. Because of the large amount of power required to the adaptive circuit respect to that generated by the piezoelectric component, this type of energy harvesting circuit can be successfully used for values of open circuit voltage grater than 25V; for values of open circuit voltage lower than 25V, a direct charge circuit has to be implemented. Sodano et al. [5] made several energy harvesting experiments using a piezoelectric plate excited by a mechanical exciter at resonant and random frequencies; using a direct charge energy harvesting system, they reached electrical powers of order of magnitude of milliwatts, and charged low capacity batteries in some hours, depending on the capacity of the battery and on the exciting frequency. Lesieutre et al. [6] demonstrated that a piezoelectric energy harvesting system has the advantage to naturally produce damping. Finally, recent studies have also been conducted [7],[8],[9],[10] to improve piezoelectric energy harvesting systems.

On the base of the previous literature, the aim of the present study is to provide some analytical models and several experiences of power generation from the vibrations of an aluminum cantilever beam with attached piezoelectric patches. The generated energy is accumulated on a rechargeable battery and the flow of power from the PZT component to the storage device is regulated by the characteristics of the energy harvesting circuit. After having shown the evidence of the physical phenomenon of PZT power harvesting and the effectiveness of the mathematical models presented, some parametrical analyses are provided in order to choose the optimal characteristics of each energy harvesting component. All the design choices and the correspondences between the theoretical analysis and the experimental results are reported.

2 Experimental Setup

The experimental setup is made up of three different configurations of cantilever beam, (being L, b and t the length, width and thickness respectively) with two piezoelectric Quickpack© (QP10N) (Fig.1) patches attached symmetrically on both sides of the beam, close to the clamped edge, as shown in Fig.2. An A/C voltage is generated when the piezoelectric vibrates perpendicular to its middle plane. One of the piezoelectric patch is utilized as a sensor to monitor the structural response of the cantilever beam, while the other one is connected to the energy harvesting circuit. The piezoelectric QP10N characteristics, as well as geometric and mechanical properties of the three beam configurations, are shown in the table 1 and figures 1 and 2.

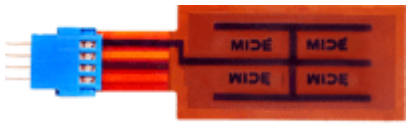


Fig. 1 – Piezo QP10N

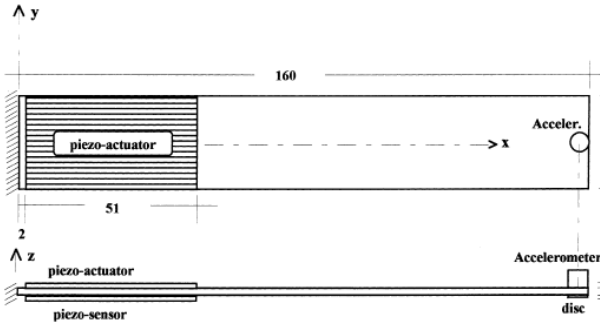


Fig. 2 – Structural configuration, with piezo patches bonded close to clamped edge

Each cantilever beam is excited at its tip by a mechanical exciter, Vibration Exciter – Type 4809, driven at the desired amplitude and

frequency by the signal generated by the workstation, Scada LMS III, and produced by a FFT Analyzer; To drive the exciter, the voltage of the signal produced by the FFT Analyzer needs to be properly amplified. To this end, a power amplifier is used. The experimental setup is shown in figure 3. The signal produced to drive the exciter is sent as feedback to the analyzer in order to perform the Fourier analysis together with the signal coming from the piezoelectric acting as sensor.

	PZT QP10N	BEAM1	BEAM2	BEAM 3
L [mm]	50,8	303	160	100
b [mm]	25,4	25,4	25,4	25,4
t [mm]	0,381	2	2	1.1
ρ [kg/m^3]	5670	2637	2637	1644
E_{11} [GPa]	18,95	59,7	59,7	250
C [μF]	0,06			
e_{31} [C/m^2]	12,73			

Table 1 – Beam and Piezo geometric characteristics

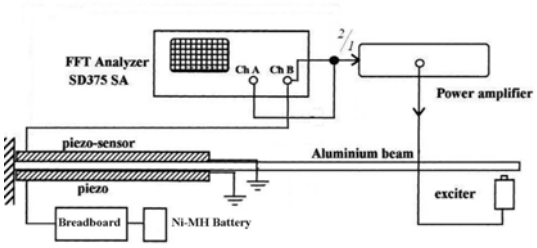


Fig. 3 – Experimental Setup

3 The model of the piezoelectric energy harvesting system

Ideally, a piezoelectric energy harvesting system can be divided into three subsystems:

- The piezoelectric generator
- The energy harvesting circuit
- The storage device.

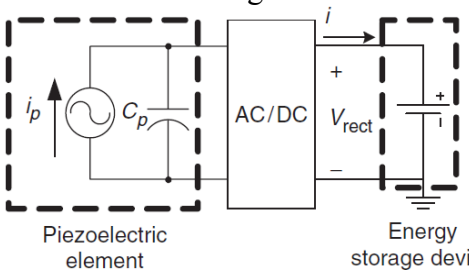


Fig. 4 - Electric schematization of the energy harvesting processing

In this section mathematical models are developed in order to provide an analytical description of each subsystem.

3.1 The piezoelectric generator

The characteristic of a piezoelectric element is to generate an electric current if a mechanical stress is applied on it. Let ε_1 be the strain in the x_1 direction produced by the action of a bender force applied in the x_3 direction; if no electric voltage is applied on PZT electrodes, the electric displacement D acting in the x_3 direction can be written as:

$$D_3 = e_{31}\varepsilon_1$$

being e_{31} the piezoelectric electromechanical coupling coefficient.

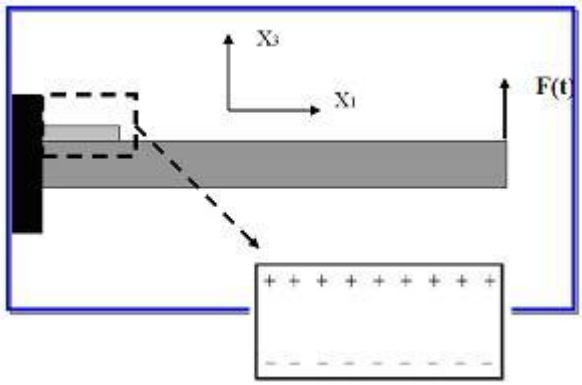


Fig. 5 - Illustration of the structural setup, and piezoelectric direction of polarization

Denoting with A the piezoelectric area, the charge produced by the piezoelectric element is:

$$Q = \int_A D_3\, dA \Rightarrow Q = b\int_0^L e_{31}\,\varepsilon_1\, dx_1 \Rightarrow$$

$$Q = \frac{bt_b}{2} e_{31}\left[\left.\frac{\partial w}{\partial x_1}\right|_{x_1=\xi_1} - \left.\frac{\partial w}{\partial x_1}\right|_{x_1=\xi_2}\right]$$

where w is the displacement in *the* x_3 direction, b is the width and t_b is the thickness of the beam.

In the illustrated configuration (piezo patches attached to the top surface of a beam) the generation of an electric charge Q is due to the fact that the bending of the beam generates a field of strain (and stress) in the piezo patches that induce, due to direct piezoelectric effect, a charge generation that can be collected on the

electrodes, placed at the top and bottom of the piezo patches.

Since the piezoelectric patch is attached to a cantilever beam, let us consider all the system (beam and piezo) as a vibrating cantilever beam where vibrations are produced by a sinusoidal force applied at the tip of beam. The equation of the cantilever beam is:

$$\rho A \frac{\partial^2 w(x,t)}{\partial t^2} + E_b I_b \frac{\partial^4 w(x,t)}{\partial x^4} = 0$$

with the following boundary condition:

$$\begin{cases} w(0,t) = 0 \\ w'(0,t) = 0 \\ w''(L,t) = 0 \\ EI \cdot w'''(L,t) = F \sin(\omega t) \end{cases}$$

ρ being the mass per unit volume and A the area of the section of the beam.

The thickness of the piezoelectric element is one order of magnitude smaller than that of the beam so, in the expression of Q, the thickness t_b has been considered as negligible. The distribution of the ε along the piezo has been assumed as uniform.

The dynamics of the beam can be then expressed with a modal approach. The modal shapes can then be obtained as usual in terms of displacements and from them the curvature shape modes are easily obtained. The curvature modes are direct expression of the strain (and stress) at the top of the beam, where the piezoelectric harvester is placed.

Once the dynamics of the beam is solved in terms of modal responses and the displacements response $w(x,t)$ is reconstructed, the displacement w can be introduced in the expression of Q. The dynamics of a beam with attached piezo is described in ref. [11] where piezos are considered as actuators.

3.2 The piezoelectric energy harvesting circuit

Recent studies demonstrated the possibility of implementing an adaptive energy harvesting circuit: it has been demonstrated [4] that it is extremely advantageous to have a complex circuit capable of measuring the instantaneous value of the current produced by the piezoelectric element, and to drive the choice of using either the direct charge circuit or the adaptive one. In particular, an increment in the current produced of order of 4 times has been observed. Unfortunately, the electric components needed to perform this kind of adaptive circuit require a level of open circuit voltage (i.e. the voltage measured between the electrodes of the piezoelectric generator without any kind of additional electric components) of 25V, which is extremely high if compared with the ones that a piezoelectric element, integrated in structures acting in typical working environment, can generate. This value, furthermore, reveals to be at least one order of magnitude higher than those found both in literature, and in the present work, whose values are extremely similar

Since the piezoelectric open circuit voltage value is lower than 25V, according with the results obtained by Lesieutre et al. [4], a direct charge piezoelectric energy harvesting circuit has been implemented.

This circuit is made up of an AC-DC voltage rectifier, a capacitor and a storage device.

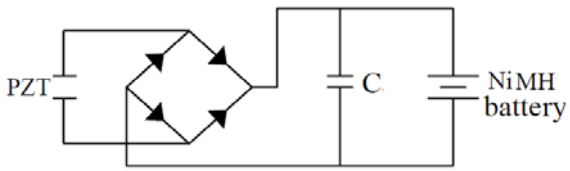


Fig. 6 Electric schematization of the energy harvesting circuit

Firstly, a way to create an electrical model of the piezoelectric generator has to be found. Lesieutre et al. [4] showed that a piezoelectric generator can be seen as an ac current generator in parallel with a capacitor. The piezoelectric element has thus an internal capacitive impedance. In the figure 7 a model of the electric circuit made up of the piezoelectric component and the resistive load is illustrated.

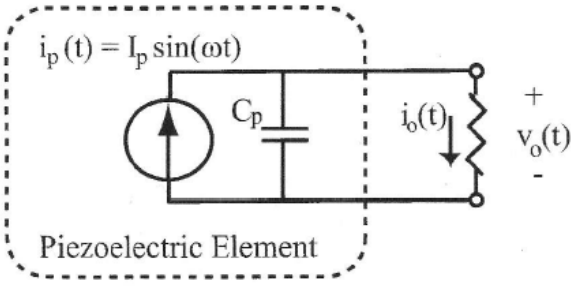


Fig. 7 - Circuital model of piezoelectric element

Considering the impedances of both the piezoelectric element and the load, the relationship between the impedances that provides the maximum power flow can be easily obtained from the model illustrated in fig. 8:

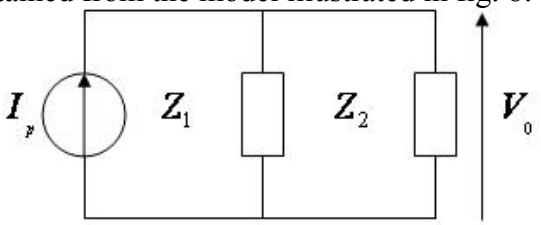


Fig. 8 - Circuital representation of piezoelectric element

The output voltage can be expressed as:

$$V_0 = \frac{Z_1 \cdot Z_2}{Z_1 + Z_2} \cdot I_p$$

In the expression the subscript 1 is used for the piezoelectric while the subscript 2 is used for the load. The piezoelectric impedance can be expressed as follows:

$$Z_1 = \frac{V_p}{I_p} = -\frac{j}{\omega \cdot C_p}$$

So, the amplitude of the voltage is then:

$$V_0 = I_p \frac{R}{\sqrt{1 + (\omega C_p R)^2}}$$

The maximum flow of power is reached when the two impedances are matched, thus corresponding to consider an output voltage of one half of the value of the piezoelectric voltage $Z_1 = Z_2$.

As it has been explained, the battery has to be charged with a DC current so an AC-DC rectifier has to be used for the purpose. A mathematical representation of the circuit is illustrated in fig. 9:

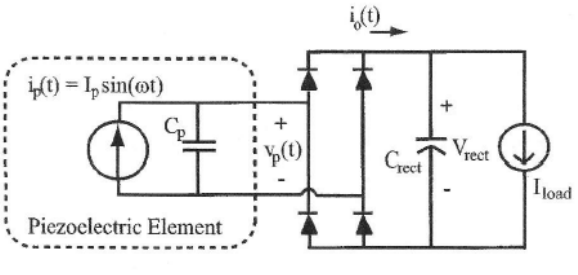


Fig. 9 - Circuital representation of energy harvesting circuit

Let us indicate with *u* the time necessary to the piezoelectric voltage to reach from zero the value V_{rect} . The value of the output current is then:

$$i_0(t) = \begin{cases} 0 & 0 \le \omega t \le u \\ \dfrac{C_{rect}}{C_{rect} + C_p} I_p \left| sen(\omega t) \right| & u \le \omega t \le \pi \end{cases}$$

If the capacity of the rectificator is significantly larger than the piezoelectric capacity, we can assume that the output current is equal to the current generated by the piezoelectric element, in fact we have:

$$C_{rect} >> C_p \Rightarrow i_0(t) = i_p(t)$$

4 Experimental validation of the model

In the previous paragraphs, we introduced several simple mathematical models to be used to predict the results expected from experiments. These models are derived analyzing the behavior of the piezoelectric, the equation of dynamic of a cantilever beam, and equation of electric circuit. To verify the reliability of the provided models, some basic experiments have been performed.

In the first experiment, the electrical model of the piezoelectric element, shown in figure 7, was considered.

The experimental setup used the cantilever beam described as the third configuration in table 1. The geometrical and electrical characteristics are shown in the figure 10. The beam is loaded at its extremity by a sinusoidal force with a frequency of 50 Hz, using the experimental setup shown in the previous pages.

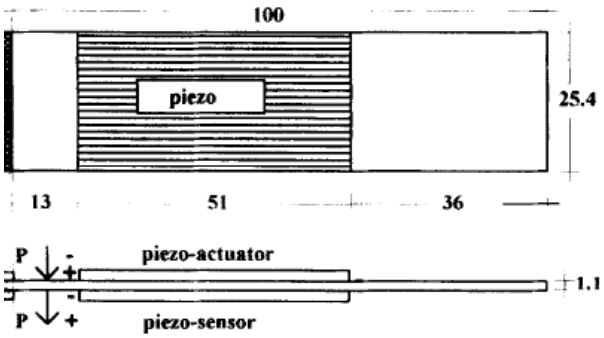


Fig. 10 – Third experimental structural configuration

The electrodes of the piezoelectric component have been closed on a resistance of varying magnitude and the value of the output voltage has been measured. In Fig. 11, the theoretical and experimental values of the output voltage for different magnitudes of the resistance, have been reported. The experimental values are represented by the dots while the theoretical values are represented by the continuous line.

A good correspondence between the experimental and theoretical values of the output voltage is observed.

In the second experiment the optimal condition of the load impedance has been verified. As it has been discussed before, this condition is reached when the impedance of the load matches the impedance of the piezoelectric. Using the same experimental setup of the previous experiment, the experimental value of the output power has been measured varying the magnitude of the resistance, from 5 KΩ to 33MΩ, as shown in the Fig. 12.

We can notice in the figure that as the power has a peak value of about 2MΩ for a piezoelectric current of 216μA, which corresponds to 2.45 μW of power. Another important aspect to notice is that, for high values of the magnitude of the resistance, the output power tends asymptotically to a fixed value, in accordance with theoretical results.

In the third experiment the dependence between the piezoelectric current and the frequency of the sinusoidal force has been shown.

The third configuration of the cantilever beam has been excited with sinusoidal forces for value of the frequency. The intensity of the current generated by the piezoelectric element has been measured with a digital multimeter. As it can be seen from Fig. 13, the intensity of the current increases of an order of magnitude in proximity of the resonant frequency equal to 50 Hz for the beam under concern.

5 Optimal placement of the piezoelectric harvester

In order to maximize the power output from structural vibrations of mechanical devices by means of piezoelectric patches attached to the structures, one has to establish the optimal placement of piezo device.

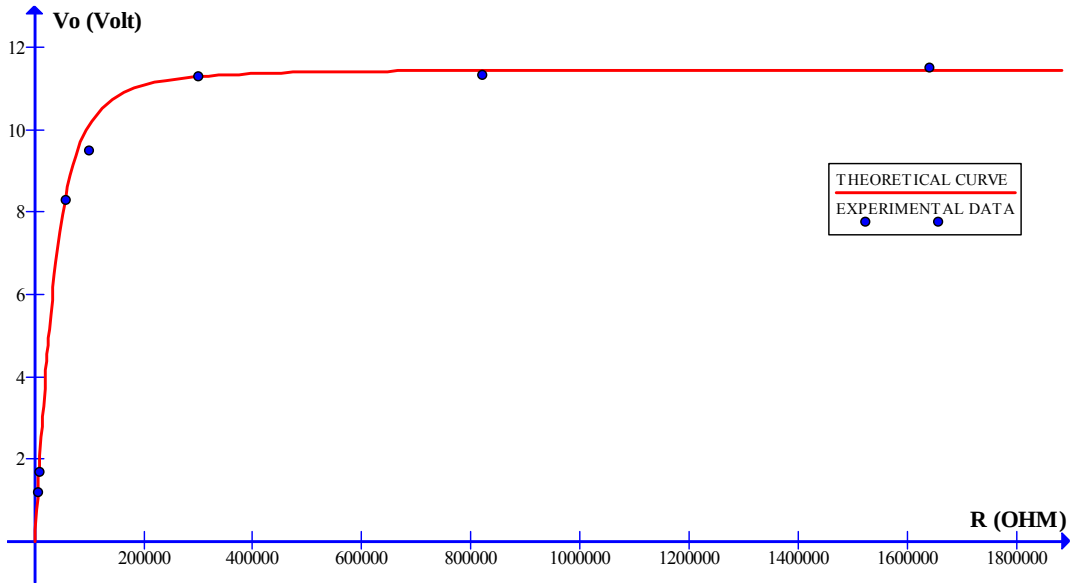


Fig. 11 – Comparison of relationship between V_0 and R and the experimental data

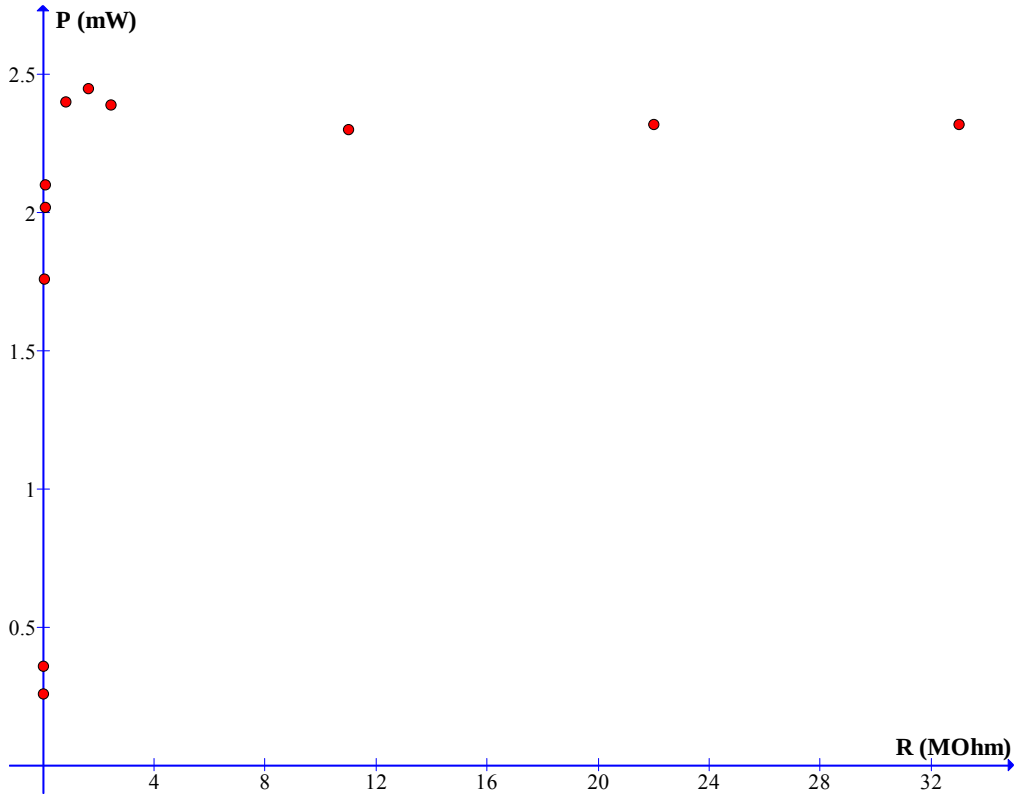


Fig. 12 - Experimental value of power produced by piezoelectric as a function of the load resistance, represented by the circuit in fig. 7

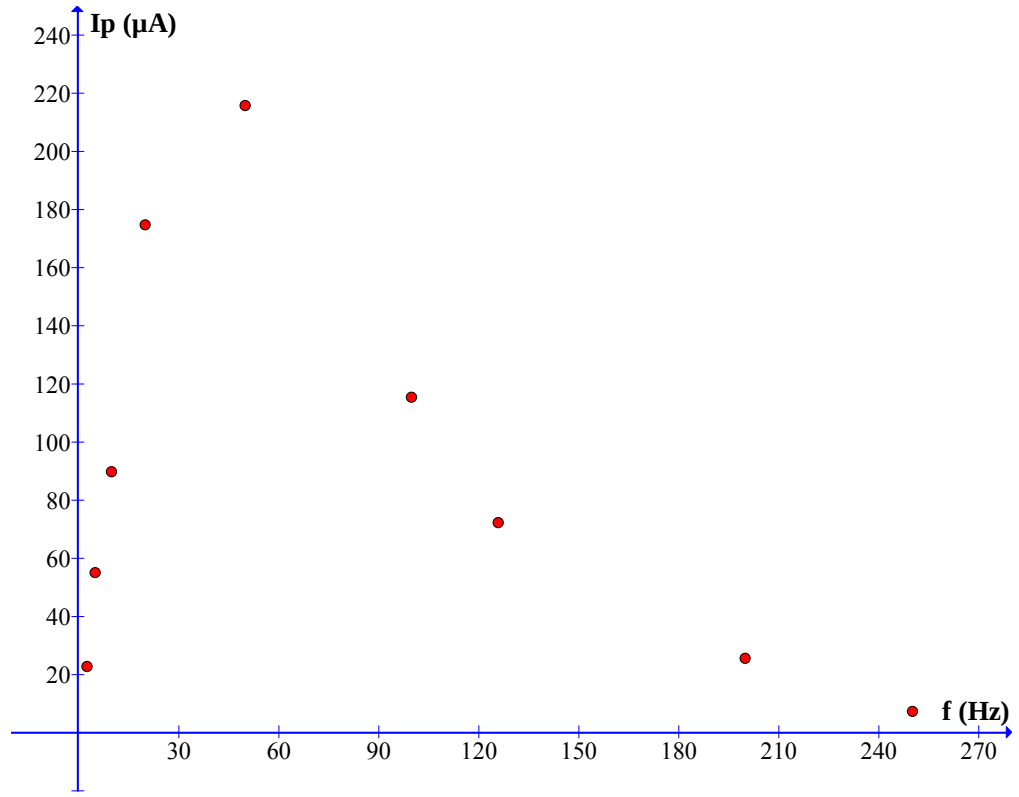


Fig. 13 - Experimental value of current produced by piezoelectric as a function of the exciting frequency

In most cases the structural vibrations are generated by mechanical components subjected to random loading force that mostly excite the first mode of vibration. The shape of the first vibration mode, especially the strain mode of the system, needs to be considered. It can be shown that in the case of a cantilever beam the maximum value of the power output can be reached with the piezoelectric element attached near the clamped edge of the beam and with the same piezoelectric covering the most of part of the beam. This happens because in this way the piezoelectric element undergoes the maximum deformation of the cantilever (expressed as the modal curvature of the first mode of a cantilever

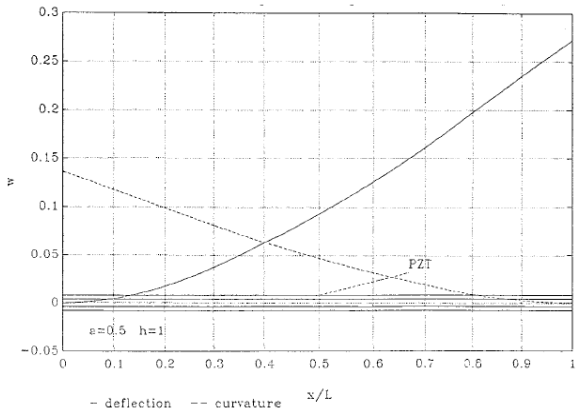


Fig. 14 - Cantilever beam — optimal PZT patch to excite mode 1.

in fig.14) and so is able to absorb the greatest value of the elastic deformation energy stored in the beam [11].

Some parametrical analyses for the three configurations of the cantilever beams have been conducted with piezos covering only part of the available surface. The cantilever beams have been excited at the trailing edge with a sinusoidal force at the first resonant frequency and the magnitude of the intensity of the current generated by the piezoelectric component has been measured. In table 2 the results obtained are shown. The ratio between the piezoelectric area and beam area has been indicated with Ap/Ab.

From the obtained results it can be stated that the maximum value of the current occurs for the piezoelectric element being attached next to the clamped edge and increasing the Ap/Ab ratio, the magnitude of the intensity of the current generated by the piezoelectric element also increases, as expected.

	Ap/Ab	***f_1(Hz)***	***Ip(μA)***
Configuration 1	0,17	13	75
Configuration 2	0,32	37	270
Configuration 3	0,51	50	515

Table 2 – Ratio of piezoelectric patches coverage

The fifth experiment has been conducted to confirm that, as previously stated, the maximum value of the output power occurs when the rectifier voltage is equal to:

$$V_{rect} = \frac{I_p}{2\omega C_p}$$

The experiment has been conducted varying the magnitude of the load resistance, considering the configuration of the cantilever beam excited at the free edge with a sinusoidal force with a frequency of 50 Hz. The Fig. 15 shows the experimental values of the output power obtained varying the magnitude of the resistance.

According to the theoretical value, the maximum of the power output is obtained for a load resistance of about 11MΩ.

Finally, an energy harvesting system with a NiMH battery of a capacity of 120mAh and a voltage of 1,2V has been created. The experimental setup adopted was the same of the previous experiments: the third configuration of the cantilever beam excited with a sinusoidal force close to the first natural frequency. The output voltage has been measured at regular time intervals, and the figure 16 shows the voltage measured during the charging of the battery

The charging process lasted about 15 hours, even if the voltage across the voltage across the electrodes reached the operational value of 1.2 Volts after only approximately one hour.

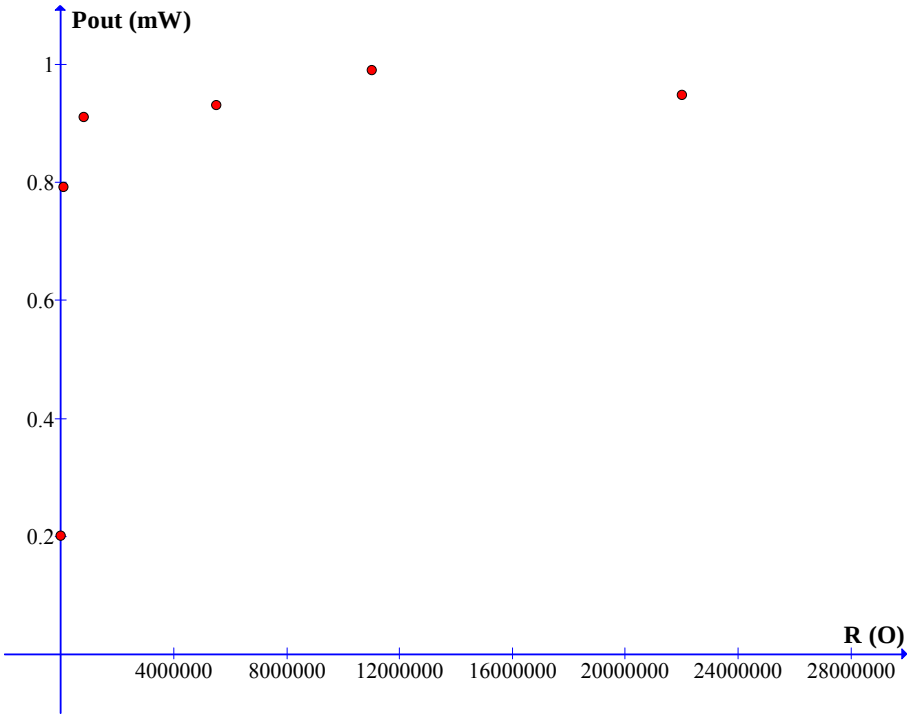


Fig. 15 - Experimental value of power dissipated in a resistive load, in the circuit of fig.9, as a function of the load resistance

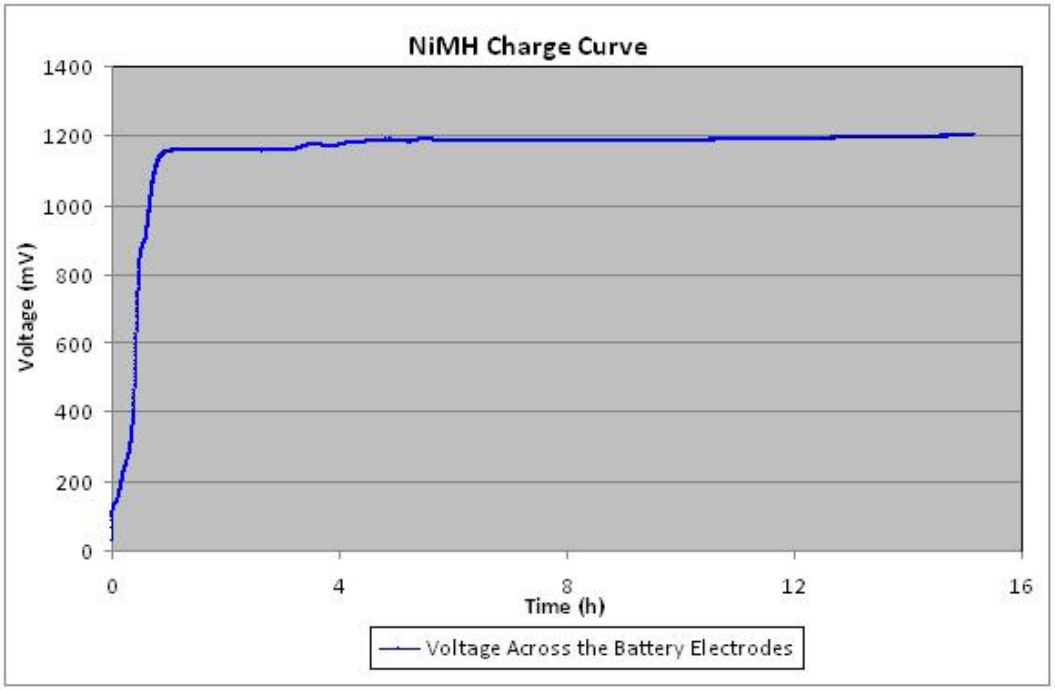


Fig. 16 – NiMH battery charging curve

6 Conclusion

In this work, a NiMH battery has been charged by means of piezoelectric element bonded on a vibrating structures. Piezoelectric materials are capable to convert the hosting structures mechanical vibration into electrical energy, conveyed into the storage device through an electric circuit. This allow to gather electrical energy from a vibrating environment in an almost endless way (at least until the vibration is present), using this energy for powering several devices with a duty cycle compatible with the energy collected.

A criteria for the optimal placement of patched piezoelectric element used as energy harvester has been also illustrated

Future developments of the presented activities will see investigation on the use of different kinds of storage devices, such as rechargeable battery of different technology, e.g. Li-ion, or other devices, such as Electric Double Layer Capacitor (EDLC), commonly known as supercapacitor.

References

[1] Paolo Gaudenzi, **"**Smart Structures: Physical Behaviour, Mathematical Modelling and Applications**",** John Wiley 2009

[2] Umeda M., Nakamura K. and Ueha S., "Energy storage characteristics of a piezo-generator using impact induced vibration", *Journal of Applied Physics*, Vol. 36, pp. 3146–3151

[3] Niell G. Elvin, Alex A. Elvin and Myron Spector, **"**A self-powered mechanical strain Energy sensor**",** *Smart Material Strucures,* Vol 10, 2001, pp. 293-299

[4] Ottman G. K., Hofmann H., Bhatt A. C., and Lesieutre G. A., *"Adaptive Piezoelectric Energy Harvesting Circuit for Wireless, Remote Power Supply",* IEEE Transactions on Power Electronics, Vol. 17, No. 5, September 2002, pp.1-8

[5] Sodano H., Magliula E. A., Park, G., and Inman D. J., "Electric Power Generation using Piezoelectric Devices*", 13th International Conference on Adaptive Structure and Technologies*

[6] Lesieutre G. A., Ottman G.K. , Hoffmann H., "Damping as a result of piezoelectric Energy harvesting", *Journal of Sound and Vibration* 2004

[7] Kauffman J. L. , Lesieutre G.A., **"**A Low-order Model for the Design of Piezoelectric Energy Harvesting Devices**",** *Journal Of Intelligent Material Systems And Structures,* Vol. 20— March 2009

[8] Dietl J. M., Wickenheiser A. M. and Garcia E. **"**A Timoshenko beam model for cantilevered piezoelectric energy harvesters**",** *Smart Material and Structures,* 19 (2010)

[9] Sodano H., Liao, **"**Piezoelectric Damping of Resistively Shunted Beams and Optimal Parameters for Maximum Damping**",** *Journal of Vibration and Acoustics,* August 2010

[10] Junrui Liang and Wei-Hsin Liao, **"**Energy flow in piezoelectric Energy harvesting systems**", "**, *Smart Material and Structures,* 20 (2011)

[11] Gaudenzi P., Barboni R., Fantini E., Mannini A. **"**Optimal Placement of PZT actuators for the Control of Beam Dynamics**",** *Smart Material and Structures*, Vol. 9, 2000

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Renewable Energy Perspectives for Aviation

H. Kuhn, C. Falter, A. Sizmann
Bauhaus Luftfahrt e.V., Lyonel-Feininger Straße 28, 80807 Munich

Keywords: *alternative renewable energies, solar drop-in fuels, solar hydrogen, feasibility of electric flying*

Abstract

The transformation of the energy supply from fossil to renewable energy is the single most important challenge for the aviation industry's long-term future.

The substitution of fossil kerosene by suitable, sustainable and scalable alternative energy carriers is a key requirement for which strategies in three categories, i.e. the drop-in fuel, non-drop-in fuel and fully electric energy path, are presented. For long-term renewable energy perspectives, drop-in fuels from solar thermal reactors, the use of renewable hydrogen for either synthetic paraffinic kerosene or fuel cells, and the material advancements in electric energy storage promise substantial innovation potential. Based on fundamental principles for future energy technology assessment it is shown that the fully electric aircraft may be closer to realization that generally assumed.

1 Introduction

The future of aviation will be shaped by a growing mobility demand, limited fossil energy resources and by the need for climate protection. Today, individual transportation in general and aviation in particular are sectors that are most dependent on fossil fuels which are ideal energy carriers due to their high energy density and convenient handling and storage properties. The most significant challenge for the future of aviation is therefore the substitution of fossil fuels or the aircraft power system to implement entirely renewable carbon-neutral energy carriers and power systems in aviation. Selected alternative fuels are summarised in Table 1.

Table 1 Main benefits and limitations of alternative fuels for aviation. The color code indicates a qualitative aptitude of a fuel option in a 2050 time frame.

	Suitability	Sustainability	Scalability
Drop-in fuels			
CTL, GTL			
BTL			
HEFA			
STL			
Non-drop-in fuels			
LNG			
LH2			
Alcohols			

CTL/GTL/BTL: coal/gas/biomass-to-liquid; HEFA: hydroprocessed esthers and fatty acids; STL: sunlight-to-liquid; LNG: liquid natural gas; LH2: liquid hydrogen;

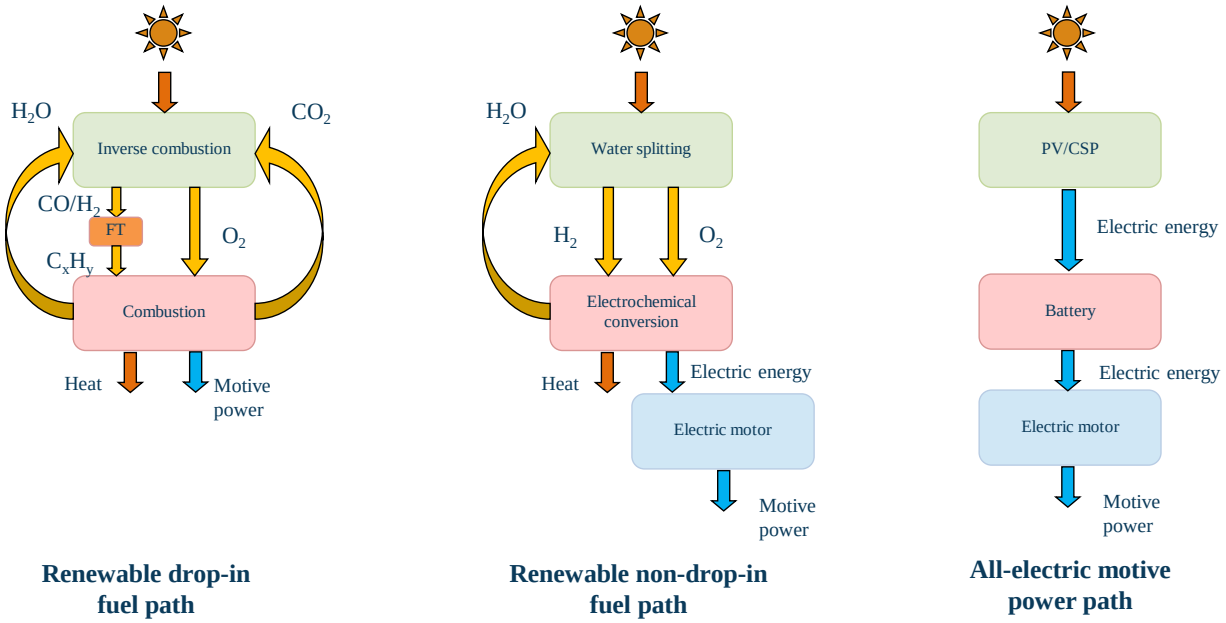


Fig. 1 Three paths for renewable energy perspectives in aviation. PV: photovoltaics; CSP: concentrated solar power.

Suitability, sustainability and scalability serve as major criteria in order to evaluate different alternative fuels. A low suitability score indicates that a re-design of the aircraft fuel system and changes in the infrastructure are necessary for implementation. The scalability criterion addresses the technical feedstock potential, the processing potential and challenges in high volume logistics.

Examples of three alternative energy strategies are shown in Fig. 1, (1) the drop-in-capable solar fuel option, using novel solar chemical reactors as discussed below or the well-known photosynthesis path, both of which present a renewable energy perspective with the advantage of no changes to infrastructure or aircraft fuel systems, (2) the non-drop-in solar fuel option which opens the perspective of highly efficient fuel cell power systems and electric propulsion in aviation, and (3) the all-electric option using solar, wind and other renewable power, the latter two require a complete re-design of the aircraft but offer the ultimate flexibility in the choice of primary energy. Figure 2 shows the energy and exergy densities of various energy carriers and the large "gap" between hydrocarbon fuels and batteries. For a fair comparison, the net exergy, i.e. the net useful work, and the physical potential have to be assessed (see also Section 4.1). This reduces the gap of 56 to a factor of approximately 8 and greatly enhances the feasibility of fully electric aviation.

In this paper the feasibility of strategies ranging from solar drop-in fuels to step-change design of the all-electrically powered aircraft are discussed.

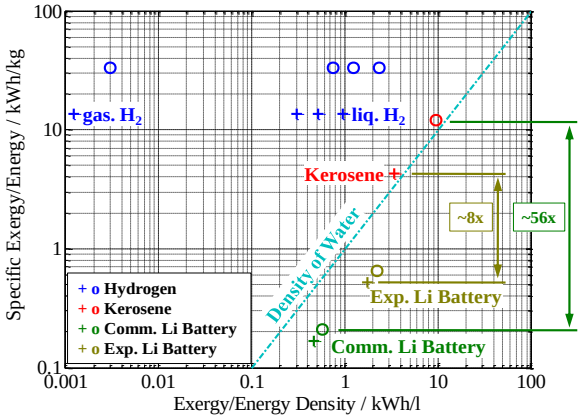


Fig. 2 Specific exergy vs. exergy density (+) as well as specific energy and energy density (o) of selected energy carriers.

2 Drop-in Fuels

The use of drop-in fuels which have equal characteristics as conventional kerosene requires in general no adaptation of the fuel distribution infrastructure, on-board fuel systems or combustion engines. One way to reduce the carbon footprint of aviation is to introduce sustainable, carbon-neutral drop-in fuels that can be used to substitute conventional kerosene in present motive power systems and could therefore

represent an elegant way to fulfill the International Air Transport Associations (IATA) target of carbon-neutral growth starting in the year 2020. Certified drop-in fuels based on different resources exist and enter long-term testing phases (Lufthansa, KLM).

2.1 Biofuels

The biomass-to-liquid (BTL) and the hydroprocessed-ester-and-fatty-acids (HEFA) processes convert biomass to liquid fuels by gasification/FT and hydrogenation, respectively, and provide a potential sustainable path. Their capability to supply an appreciable amount of liquid fuels is a topic of scientific research and remains to be quantified. The growing of biomass requires the allocation and tilling of farmland and therefore competes in general with the production of food plants which has already lead once to rising corn prices in Mexico, for example, and which might lead to a growing conflict of food versus fuel production. An analysis addressing land availability is published elsewhere [2].

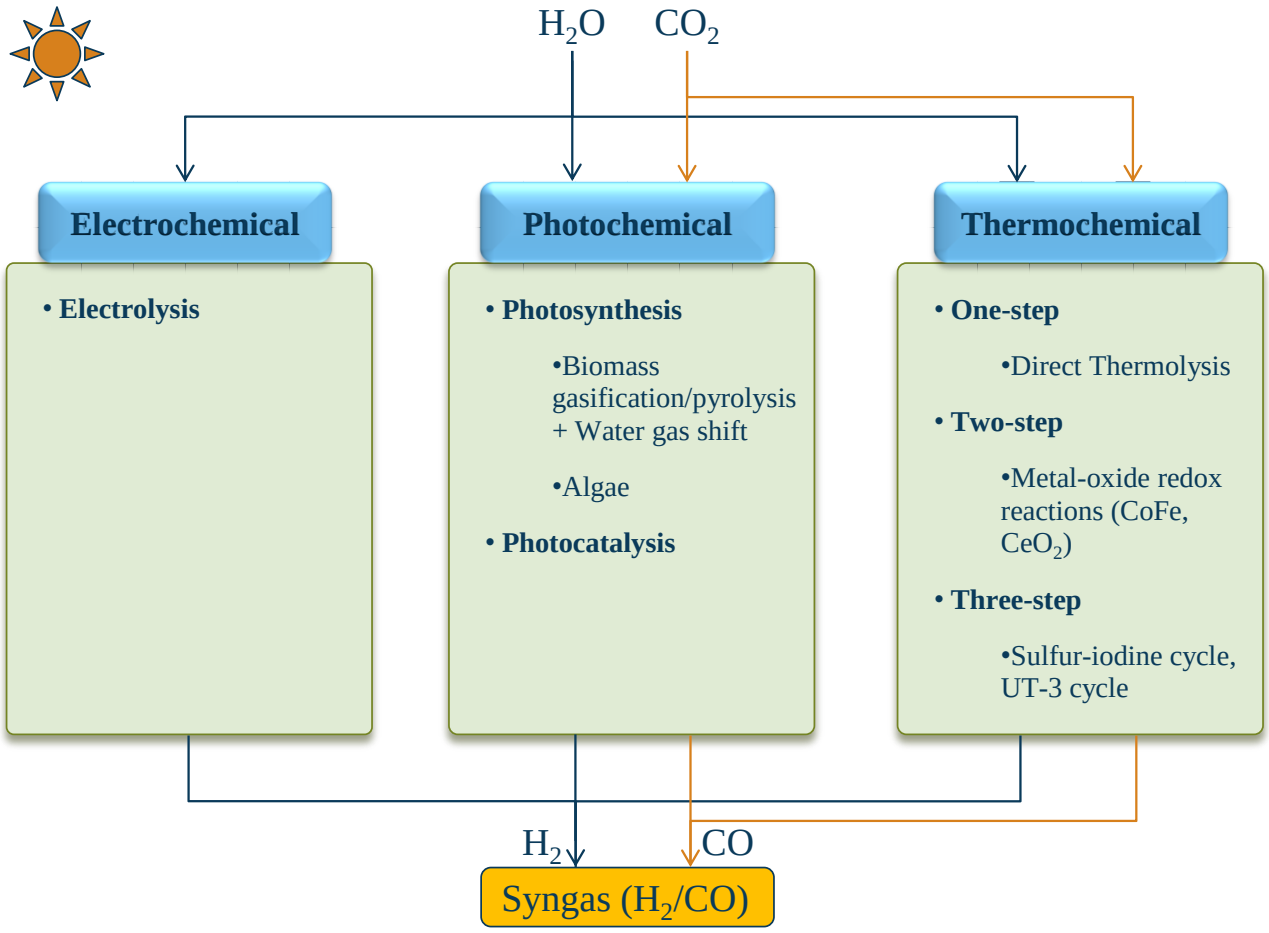


Fig. 3 Three paths to solar syngas (hydrogen and carbon monoxide) production, a precursor for solar fuels.

Different approaches exist for the production of sustainable fuels based on synthetic gas (syngas), see Fig. 3, all of which satisfy the requirement of suitability which is an important feature for quick implementation of renewable aviation fuel options. Gas-to-liquid (GTL) and coal-to-liquid (CTL) processes require fossil energy carriers that are gasified to produce syngas, a mixture of hydrogen and carbon monoxide, which is converted to liquid hydrocarbons in the Fischer-Tropsch (FT) process, invented in Germany in the 1920s. The use of non-sustainable coal and natural gas reduces finite reserves of energy carriers and adds significantly to CO_2 emissions [1].

A possible remedy for the problem of common farmland is the use of solar energy which is not confined to arable land but is, in varying concentrations, incident on every part of

the earth. Regions most interesting for the use of solar energy, due to their relatively high amount of accumulated irradiation, are found mainly in deserts near the equator where no replacement of food crops or settlements is to be expected.

2.2 Solar Drop-in Fuels

2.2.1 Paths to Solar Drop-in Fuels

Solar fuels are produced from syngas which is derived electrochemically, photochemically or thermochemically, see Fig. 2, and converted via FT-synthesis to liquid hydrocarbons.

In *electrochemistry*, hydrogen is produced via solar electrolysis, a well-known process, the efficiency of which is limited today to roughly 15%[1], partly due to the twofold energy conversion. *Photochemistry* converts photoenergy into chemical energy via photosynthesis or photocatalysis. Gasification of biomass allows for the direct production of syngas, making use of the natural photosynthetical process which includes CO_2 capture from the atmosphere. The conversion efficiency of solar energy to chemical energy stored in the liquid fuel is limited to less than two percent[2], mainly due to a rather low photosynthetical conversion efficiency.

Concentrated solar energy incident on a solar reactor can be used to drive *thermochemical* reverse combustion cycle.

2.2.2 Thermochemical Solar Fuels

In a two-step solar thermochemical cycle a metal oxide is first reduced at an elevated temperature which causes an oxygen deficiency in the material. This chemical potential can be used in the second step at a lower temperature of about 1000 K to split water and carbon dioxide into syngas. Implementations of this cycle comprise the use of the redox-pairs ZnO/Zn, Fe_3O_4/FeO [3,4] or CeO_2 (non-stoichiometric oxygen deficiency) [5]. The theoretical thermodynamic potential of the latter cycle is 19% without considering heat recuperation [5]. Provided the technical challenges, e.g. scale-up or CO_2 capture from air can be solved, the ceria cycle combines suitability, sustainability and scalability with a relatively high efficiency and is therefore a very promising way to produce liquid fuels from solar energy, carbon dioxide and water. The combustion of a solar fuel will produce water and CO_2, thereby completing the material balance.

The conversion efficiency of the thermochemical path can be subdivided in absorption efficiency $\eta_{\text{absoprtion}}$ of the solar reactor and conversion efficiency of heat to work $\eta_{\text{heat-to-work}}$.

$$\eta = \eta_{\text{absorption}} \cdot \eta_{\text{heat-to-work}} \qquad (1)$$

$\eta_{\text{absoprtion}}$ is the rate of useful heat absorbed divided by the incident solar power:

$$\eta_{\text{absorption}} = 1 - \frac{\sigma T^4}{IC} \qquad (2)$$

wherein σ is the Stefan-Boltzmann constant, T the nominal reactor temperature, I the solar irradiation and C the concentration ratio. The maximum amount of work that can be gained from a heat engine working between a high (T_H) and a low (T_L) temperature is defined by the Carnot efficiency:

$$\eta_{\text{heat-to-work}} = 1 - \frac{T_L}{T_H} \qquad (3)$$

In Fig. 4 the solar-to-chemical energy conversion efficiency is shown for different concentration ratios of the optical system and in comparison to the ideal Carnot efficiency. For increasing concentration ratios the attainable system efficiency rises, showing peak values at an optimum temperature that shifts towards higher values as the concentration ratio increases. This is due to adverse effects of heat radia-

[1] Electrolysis efficiency is assumed to be 70%, solar-to-electrical energy efficiency 20%. The latter value could represent the limit for industrial scale multi-cristalline silicon cells where further improvement is not precluded due to e.g. the use of mono-cristalline material or tandem cells.

[2] Photosynthetical efficiency (not accounting for plant growth) is assumed to be 5% [6], gasification efficiency 70% and FT-synthesis efficiency 50%.

tion losses off the reactor and increased efficiency of heat conversion.

The overall level of efficiency that can be reached in principle with thermochemical conversion is relatively high and well above 50% for concentration ratios of 1000 or more.

Recently, a lab-scale experiment has successfully proven the concept of solar thermochemical syngas production [5], whereas the conversion of syngas to liquid hydrocarbons is a well established process. While still in development and therefore showing low efficiency today, its potential makes this process highly interesting for a mid-term alternative fuel option.

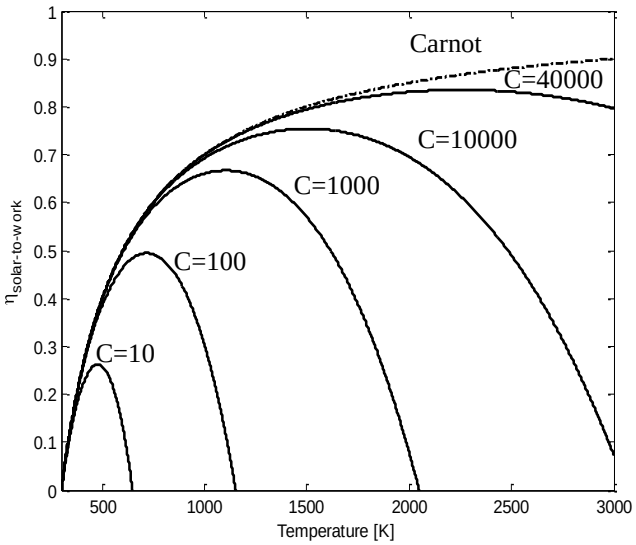


Fig. 4 Efficiency of thermochemical energy conversion for different solar concentration ratios [7,8].

3 Renewable Non-Drop-in Fuels

3.1 Alternative Non-Drop-in Fuels

As opposed to fossil or solar kerosene, *non-drop-in* fuels are not compatible with today's transportation fuel systems and therefore cannot be used without major adaptations in infrastructure including fuel production, distribution and storage, and motive power system of the aircraft. Renewably produced hydrogen and carbon dioxide captured from air can be methanized to produce *methane* that can be fed to motive power systems. Methane can be either stored at high pressure in its gaseous state or at low temperature in its liquid state. CO_2 capture from air is being investigated and has been shown to work at lab-scale. It is, however, not yet ready to be used in an industrial scale. *Alcohols* are produced by fermentation of biomass (e.g. sugar cane, corn or cellulosic material) and require in large concentrations modifications of the engine and storage system, because sealings are negatively affected. As mentioned before, the growing of energy crops may lead to a conflict with food production. *Electrical* energy gained from renewable primary energy sources is also considered to be a sustainable non-drop-in energy carrier which is stored in batteries. Benefits arise from the high overall efficiency of fully electric systems. However, the accommodation of batteries in an aircraft requires large redesign efforts. *Hydrogen* can in principle be produced from any renewable primary energy by electrolytical splitting of water, where solar energy additionally offers the use of photochemical and thermochemical paths.

3.2 Solar Hydrogen

A much discussed alternative to fossil kerosene is the implementation of hydrogen as a fuel for use in motive power systems such as aircrafts. Hydrogen can either be burned with oxygen from air within an internal combustion engine (ICE, i.e. turbine, engine, etc.) to produce shaft power or electrochemically converted to electricity within a fuel cell to produce shaft power by the use of an electric motor.

Several advantages promote the implementation of hydrogen as a fuel for aviation: given a sustainable production through by the use of renewable energy sources, the hydrogen cycle can be CO_2-free since water is the only emission. On the other hand, the combustion of fossil fuels releases CO, CO_2, SO_x, NO_x and soot, besides others, some of which have adverse effects on human health. As no carbon is used in the cycle carbon dioxide capture is not required which could be a considerable simplification of the overall process. Hydrogen production and its utilisation in fuel cells is a quite mature technology that has been demonstrated in stationary and mobile applications from mobile computers

to small aeroplanes (e.g. Boeing, DLR, ENFICAS). The foreseeable risk in its implementation is therefore rather low.

Although having been demonstrated as a working technology, the use of hydrogen as an energy carrier in the transportation sector faces technological challenges that need to be overcome for its widespread implementation. Primarily, storage of hydrogen still seems to present an issue regarding its use in mobile systems. In airplanes, a high volumetric energy density is needed to reduce the impact of fuel volume on the whole system. Hydrogen has its highest volumetric energy density in its liquid state at 10.1 MJ/l which is about one third of the kerosene value of about 33 MJ/l. The reduced energy density in connection with the given size of fuel storage tanks causes a reduction in mission length. A major safety issue may be a leakage of the hydrogen tank and hence the formation of an explosive gas mixture. Hydrogen being a non-drop-in fuel requires adaptations of on-board fuel infrastructure and motive power systems of an aircraft which demands financial efforts as well as time for its implementation. Hydrogen replacing a considerable share of fossil kerosene in aviation will therefore most certainly be a long-time perspective due to its implied structural changes in the current transportation system.

Paths to produce solar hydrogen are shown in Fig. 3 and are identical with the paths mentioned in Section 2 since hydrogen is a constituent of syngas.

4 Electric Motive Power Systems

The all-electric aircraft is the most radical approach in developing a new long-term energy perspective for aviation. Battery-based power systems provide in-flight zero-emission performance and the ultimate flexibility in the choice of primary energy (solar, wind, hydro, etc.) and potentially minimises the environmental footprint of aviation. Hybrid power systems that generate electricity by fuel cell or turbo-engine-generator systems may significantly reduce harmful in-flight emissions as well.

However, the feasibility of electric flying requires a careful analysis of future electric technologies because of the obvious energy “gap” as shown in Fig. 2.

4.1 Feasibility of Electric Flying

The feasibility perspectives and scaling properties of fully or hybrid electric motive power systems are governed by scientific developments outside the field of aviation and the following three fundamental principles as proposed by Bauhaus Luftfahrt [9,10]:

1. *Exergy concept*: the usefulness of an energy carrier for aviation is determined by its gravimetric and volumetric exergy density and not by its energy content
2. *Specific power and Ragone metrics*: the usefulness of an energy carrier in combination with a power converter is determined by its combined power density and exergy density (Ragone metrics)
3. *Hybridisation degree of freedom*: two or more energy storage and/or conversion devices that each are inadequate for electric flying may constitute an enabling power system when combined to a hybrid system.

Exergy is the part of the energy that can be entirely transformed into useful work. The net exergy of combustion engines is fundamentally limited in contrast to fully electric power systems. As shown in Fig. 2 the exergy gap between kerosene and batteries is greatly reduces compared to the energy gap. The next principle indicates that the feasibility of electric flying is also determined by the specific power requirements for take-off and climb, and not just by the exergy density. The hybridisation degree of freedom enables feasible solutions for electric flying by combining e.g. high power batteries and a fuel cell system.

At first, the system boundaries need to be defined for a meaningful comparison of different motive power systems, based on thermal and non-thermal energy conversion. The green boxes shown in Fig. 5 describe a combined energy storage and conversion system that delivers electric energy. The red dashed boxes describe

energy conversion subsystems with fuel input and electric power output, i.e. energy conversion devices such as fuel cells or turbo-engines. These devices are inherently characterised by specific power densities, whereas the combined systems are characterised by specific power *and* specific exergy densities as required for the feasibility analysis by the second principle. The performance scaling of these subsystems is one key to proper power system design.

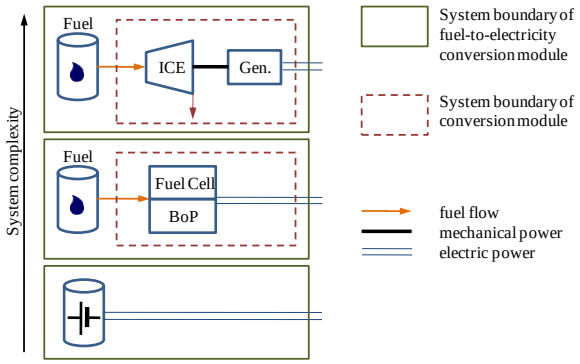


Fig. 5 System boundaries for comparison of fundamentally different energy storage and power conversion systems. (ICE: internal combustion engine; Gen.: el. generator; BoP: balance of plant)

Figure 5 shows that the ICE and fuel cell systems allow for an independent scaling of power level and energy content. Batteries incorporate both characteristics in one device. Hence either characteristic defines the necessary battery size. The system complexity increases with the number of components and may have negative impact on maintenance and overhaul. As shown in Fig. 5 and discussed in Section 4.3, the usefulness of the battery concept strongly depends on advances in material science and development.

The feasibility assessment of electric flying therefore requires an exergy and power metric that puts all energy carriers into a proper perspective with respect to conventional aircraft fuel, i.e. kerosene. Different motive power systems can be easily characterised and compared within the Ragone diagram, Fig. 6. Power systems in the quadrants A and B provide insufficient power for take-off. Only power systems in the quadrants D and C_2 enable electric flying over reasonable ranges. Subquadrant C_1 offers sufficient power however with insufficient range.

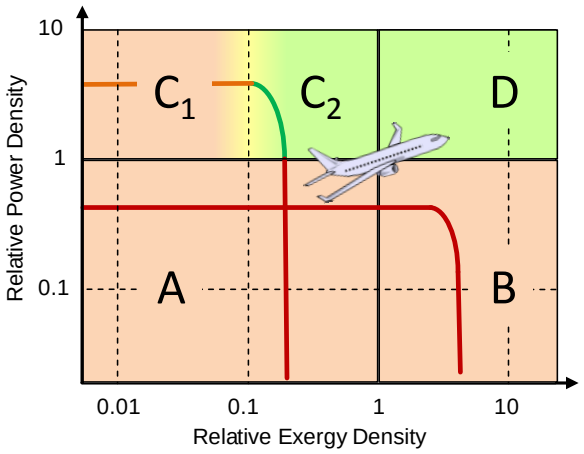


Fig. 6 Visualisation of the second principle for the feasibility assessment of electric flying. A hydrogen fuel cell system in the performance quadrants A and B provides insufficient power, whereas a future battery system provides the required power in quadrant C_2 with a reduced range due to lower exergy density.

Figure 7 shows a Ragone diagram for several battery technologies and representative data points for a typical conventional, an improved conventional long range business jet. Here the energy gap of Fig. 2 translates into a gap in operating range: the reduction in range is in first approximation proportional to the reduced exergy density of the batteries.

Applying the Breguet range equation for the battery-powered (R_{el}), Eq. 4, and the conventional powered aircraft (R_c), Eq. 5,

$$R_{el} = \frac{L/D}{g} \cdot \eta_P \cdot \eta_{el} \cdot \rho^{B}_{E,m} \cdot \frac{m_{Battery}}{MTOM} \qquad (4)$$

$$R_c = \frac{L/D}{g} \cdot \eta_P \cdot \eta_{TE} \cdot \rho^{K}_{E,m} \cdot \ln\left(\frac{1}{1 - m_{Fuel}/MTOM}\right) \qquad (5)$$

and replacing the fuel mass of the conventional aircraft by batteries in this preliminary assessment, the range of the battery-powered aircraft settles at below 555 km (300 NM) for an exergy density of 200 Wh/kg. This is roughly 5% of the original range according to a battery technology level of 2010.

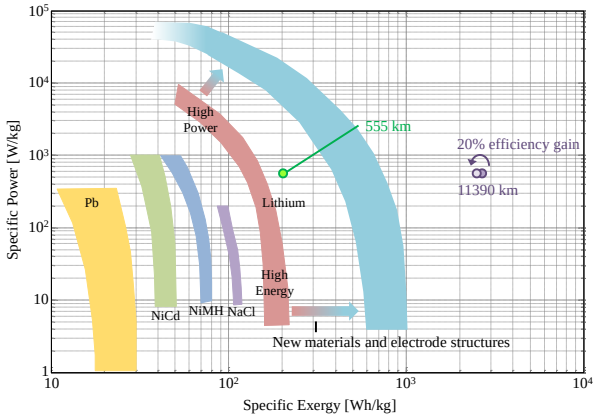


Fig. 7 Ragone diagram of selected battery technologies, prospects of lithium batteries and performance characteristics of a conventional (magenta dot; range of 11390 km (6150 NM)) and battery powered (green dot; range of less than 555 km (300 NM); technology level 2010) long-range business aircraft.

In the equations above, L/D is the lift-over-drag ratio, η_P is the propulsor efficiency, η_{el} is the efficiency of the electric system comprising the battery, the power management and distribution system (PMAD) and the electric motor, η_{TE} is the efficiency of the turbo-engine and $\rho^{B}_{E,m}$ and $\rho^{K}_{E,m}$ are the specific energies of the batteries and kerosene, respectively. $m_{Battery}$ and m_{Fuel} is the mass of the batteries and kerosene and MTOM is the maximum take-off mass.

Depicted is also an assumed efficiency gain of 20% for the conventional aircraft efficiency through improved aerodynamics. Due to the reduced demand of fuel for the fixed range of 11390 km (6150 NM) and the higher overall efficiency of 20%, the exergy density requirement for the motive power system of the conventional jet is slightly relaxed. However the most significant contribution to closing the range gap is expected to comes from advances in battery technology as shown by the light blue area in Fig. 7.

The Breguet range of a conventional and a battery-powered ATR72 is plotted versus the fuel mass-to-MTOM in Fig. 8. Different specific energies for the batteries are assumed and the corresponding ratio for the conventional ATR72 is given as well. As expected, the range increases with increasing specific energy of the battery. Interestingly to attain a range of 1000 km[3] a specific energy of 800 Wh/kg, i.e. four times the state-of-the-art commercial battery capacity, is required.

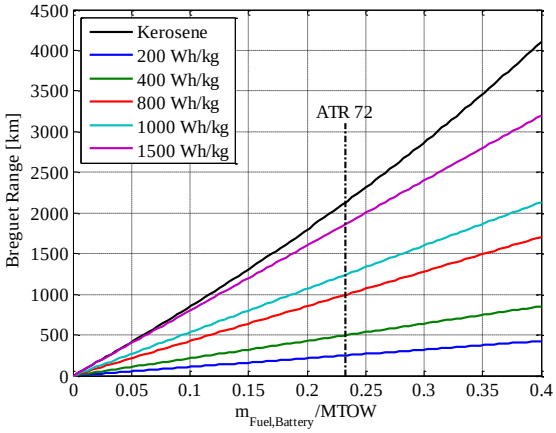


Fig. 8 Breguet range of a conventional turbo-engine powered ATR (black) and fully electric battery powered range estimates.

4.2 Motive Power System Architectures

The integration of fully or hybrid electric motive power systems allows for the optimisation of e.g. aerodynamic and propulsive efficiency, control dynamics, service accessibility as well as mass and load distribution. The new degrees of freedom arise due to the separation of power generation and power consumption compared to the traditional arrangement with turbo-engines under the wings. From a wide combinatorial variety of energy storage and power generation subsystems, Figure 9 gives an overview of a few selected motive power systems.

[3] 99% of all ATR72 flights are at great circles distances below 1050 km, whereas 90% are below 600 km. Most flights are in the great circle distances of 251-300 km (16% of all ATR72 flights) Source: Official Airline Guide (OAG).

4.3 Key Technologies

The key technologies for electric flying are batteries, fuel cells and electric motors.

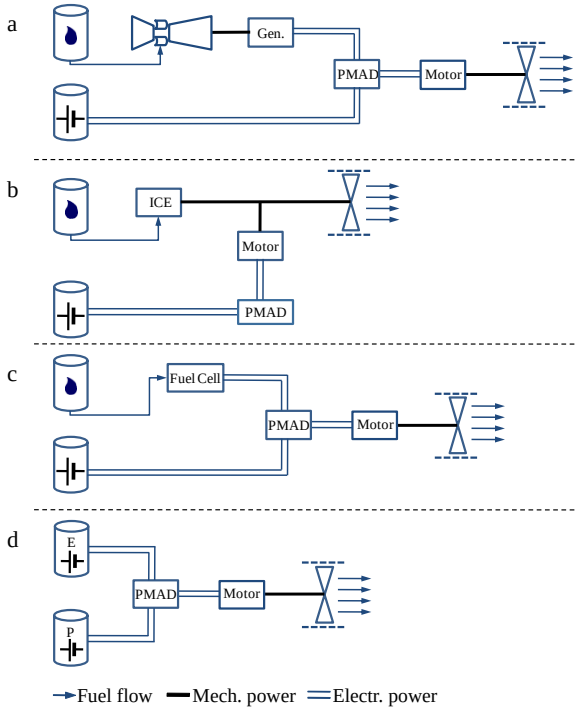


Fig. 9 Schematic depiction of selected hybrid motive power systems out of a combinatorial variety of power generator subsystems.

4.3.1 Lithium Batteries

Lithium batteries offer the highest available specific energy among all battery technologies known today and are therefore at the focus of current research. Crucial battery components with significant innovation potential are the electrodes, separators and electrolytes. In the following we discuss selected positive and negative electrode materials, see Tables 2 and 3, an elaborated evaluation of electrode materials will be published in a separate article as it would exceed the scope of this article [11]. $LiCoO_2$, $LiMn_2O_4$ and $LiFePO_4$ are commonly used as positive electrode materials for commercial lithium batteries and in combination with the primary negative electrode material, LiC_6, commercial batteries with a specific energy of up to 250-280 Wh/kg can be realised today. Li_2S and $LiTiS_2$ are proposed as new positive electrode materials as they offer very high capacities (Li_2S) [12,13,14] and discharging/charging rates ($LiTiS_2$) [15]. But these electrode materials especially have to overcome lifetime and mass production issues.

Due to safety and economical reasons LiC_6, basic material is graphite, is the primary negative electrode material for lithium batteries instead of metallic lithium although it exhibits a significant reduced capacity. LiC_6 offers reasonable lifetimes and acceptable capacities at low costs. Nevertheless, in order to increase a battery's capacity LiSi alloys are in the focus of research due to the multi-fold higher capacity prospects of this electrode material. Step-change advancements in nano-technology yielded in ten-fold higher capacities compared to graphite (LiC_6) and have been already demonstrated at lab-scale [16,17].

For a quick recharge of the batteries or temporary high power demands $LiFePO_4$ seems to be a proper positive electrode material. Charging rates of up to 400C have been demonstrated with this electrode material [18] allowing for a complete charge of the electrode in 400^{-1} hour.

These developments on the electrode materials level significantly reduce the exergy difference between today's battery technologies and kerosene fuelled propulsion systems, thus bringing electric flying closer to realization than generally assumed.

Table 2 Positive electrode materials for lithium batteries. Capacities refer to the charged state.

Material	Eq. Voltage vs. Li/Li^+	Rev. Range	Capacity Ah/kg
$LiCoO_2$	3.9	0.5	147
$LiMn_2O_4$	3.9	1	154
$LiFePO_4$	3.5	0.95	169
Li_2S	2.2	2	1671
$LiTiS_2$	1.96	1	239

Table 3 Negative electrode materials for lithium batteries. Capacities refer to the charged state.

Material	Eq. Voltage vs. Li/Li^+	Rev. Range	Capacity Ah/kg
LiC_6	0.1	1	339
Li_{solid}	0.1	1	3861
$Li_{4.4}Si$	0.3	4.4	2011

4.3.2 Electric Motors

Conventional high torque or high speed electric motors have typical specific powers around 3-8 kW/kg today. Hybrid or high-temperature superconducting electric motors and generators exhibit specific powers around 9 kW/kg and show prospects for specific powers of well above 15 kW/kg including the cryo-coolers, hence outperforming state-of-the-art turbo-engines [19]. Thus, these devices could theoretically replace turbo-engines without mass penalties.

4.3.3 Fuel Cells

Renewable hydrogen fed to polymer electrolyte fuel cells (PEFC) provide a sustainable way of supplying electric power and heat for on-board systems. PEFC exhibit the highest specific powers among fuel cells at 1.2 kW/kg on stack level today. The advantage of fuel cell systems over batteries is the independent scalability of the amount of energy stored in form of fuel, independent from the level of power generation in the stack. New catalysts still allow for performance enhancements, decreased costs as well as for higher impurity tolerances, which for example increases durability and stability [20,21]. Alternative membrane materials allow for altered operating conditions compared to state-of-the-art membrane materials [22,23], i.e. lower gas humidification needs which in turn reduces additional subsystems and power demand. Although solid oxide fuel cells (SOFC) may operate on kerosene, their low specific power implies significant mass penalty which is one of the main drawbacks of this technology.

5 Conclusion

The analysis of the energy options shows that (1) solar fuels are able to overcome known sustainability and/or scalability limitations of bio-to-liquid and other drop-in bio-fuels, that (2) the (non-drop-in) hydrogen fuel perspective has inherent limitations that are either solved by feeding into the solar hydrocarbon fuel process or by serving the electric propulsion paradigm shift via PEM fuel cells, and that (3) the all-electric regional aircraft is well within the physical regime of feasibility and potentially closer to realization than generally assumed, mainly driven by a few key advancements in material science and nanotechnology.

References

[1] Group on Int. Aviation and Climate Change, “Agende Item 2: Review of Aviation Emissions Related Activities Within ICAO and Internationally: U.S. Fuel Trend Analysis and Comparison to GIACC/4-IP/1”, GIACC/4-IP/12, 4th Meeting, Montreal, Canada, May 25-28, 2009.

[2] Riegel F. and Steinsdörfer J., “Bioenergy in Aviation: The Question of Land Availability, Yields and True Sustainability”, 3rd *CEAS Air&Space Conference*, Venice, 2011.

[3] Stamatiou A., Loutzenheiser P.G. and Steinfeld A., “Solar syngas production via H2O/CO2-splitting thermochemical cycles with Zn/ZnO and FeO/Fe3O4 redox reactions”, *Chem. Mat.*, Vol. 22, 2010, pp. 851-859.

[4] Loutzenheiser P.G., Galvez M.E., Hischier I., Graf A. and Steinfeld A., “CO2 splitting in an aerosol flow reactor via the two-step Zn/ZnO solar thermochemical cycle”, *Chem. Eng. Sci.*, Vol. 65, No. 5, 2010, pp. 1855-1864.

[5] Chueh W.C. *et al.*, “High-Flux Solar-Driven Thermochemical Dissociation of CO2 and H2O Using Nonstoichiometric Ceria”, *Science*, Vol. 330, No. 6012, 2010, pp. 1797-1801.

[6] Hall D.O. and Rao K.K., “Photosynthesis”, 6th ed., Cambridge University Press, Cambridge, 1999.

[7] Meier A. and Sattler C., “Solar fuels from concentrated sunlight”, IEA SolarPACES, 2009.

[8] Fletcher E.A. and Moen R.L., “Hydrogen and Oxygen from Water”, *Science*, Vol. 197, No. 4308, 1977, pp. 1050-1056.

[9] Sizmann A., “Neue Energieperspektiven der Luftfahrt”, *Fuelling the Climate 2010*, Hamburg, Germany, 2010, http://www.haw-hamburg.de/klimaschutz_in_der_luftfahrt.html

[10] Kuhn H., Sizmann A., “Progress and Perspectives of Electric Air Transport”, *ICAS 2012*, Brisbane, Australia, 2012, accepted

[11] Kuhn H., to be submitted

[12] Yang Y. *et al.*, “New Nanostructured Li2S/Silicon Rechargeable Battery with High Specific Energy”, *Nano Lett.*, Vol. 10, No. 4, 2010, pp. 1486-1491.

[13] Hassoun J. and Scrosati B., “A High-Performance Polymer Tin Sulfur Lithium Ion Battery”, *Angew. Chem. Int. Edit.*, Vol. 122, No. 13, 2010, pp. 2421-2424.

[14] Hassoun J., Panero S., Reale P. and Scrosati B., “A New Type of Lithium-ion Battery Based on Tin Electroplated Negative Electrodes”, *Int. J. Electrochem. Sci.*, Vol. 1, No. 3, 2006, pp. 110-121.

[15] Zhou S., Liu X. and Wang D., “Si/TiSi2 Heteronanostructures as High-Capacity Anode

Material for Li Ion Batteries", *Nano Lett.*, Vol. 10, No. 3, 2010, pp. 860-863.
[16] Cui L.F., Yang Y., Hsu C.M. and Cui Y., "Carbon-Silicon Core-Shell Nanowires as High Capacity Electrode for Lithium Ion Batteries", *Nano Lett.*, Vol. 9, No. 9, 2009, pp. 3370-3374.
[17] Chan C.K. *et al.*, "High-performance lithium battery anodes using silicon nanowires", *Nature Nanotechnology*, Vol. 3, No. 1, 2008, pp. 31-35.
[18] Kang B. and Ceder G., "Battery materials for ultrafast charging and discharging", *Nature*, Vol. 458, No. 7235, 2009, pp. 190-193.
[19] Sivasubramaniam K. *et al.*, "Development of a High Speed HTS Generator for Airborne Applications", *IEEE T. Appl. Supercon.*, Vol. 19, No. 3, 2009, pp. 1656-1661.
[20] Wang D. *et al.*, "Highly Stable and CO-tolerant Pt/Ti0.7W0.3O2 Electrocatalyst for Proton Exchange Membrane Fuel Cells", *J. Am. Chem. Soc.*, Vol. 132, No. 39, 2010, pp. 10218-10220.
[21] Mazumder V., Chi M., More K.L. and Sun S., "Core/Shell Pd/FePt Nanoparticles as Active and Durable Catalyst for Oxygen Reduction Reaction", *J. Am. Chem. Soc.*, Vol. 132, No. 23, 2010, pp. 7848–7849
[22] Burrato S.K., "Engineering the next generation", *Nature Nanotechnology*, Vol. 5, No. 3, 2010, 176.
[23] Moghaddam S. *et al.*, "An inorganic-organic proton exchange membrane for fuel cells with a controlled nanoscale pore structure", *Nature Nanotechnology*, Vol. 5, No. 3, 2010, pp. 230-236.

Design and testing of piezoelectric energy harvesting system from vibrations for wireless sensors

Amoroso F., Pecora R., Ferraro S., Lecce L.
University of Naples "Federico II", Italy

Girolami I., Antuofermo P.
MerMec SpA, Italy

Keywords: *energy harvesting, piezoelectric, bimorph*

Abstract

The concept of power harvesting works towards the development of self-powered devices not requiring replaceable power supplies. The increase in power used by the electronics has led to a reduction in battery life and has limited the functionality of the devices. In an effort to extend the life and to reduce the volume of the electronics, researchers have begun investigating methods to obtain energy from the ambient energy surrounding the device. In this work the authors focused on the characterization of a piezoelectric vibration harvesting device to be applied in a train derailment detector system, in which the harvesting device supplies the wireless transmission unit sending warning messages to the trains driver. In what follows the attempt to harvest energy by means of a specific architecture to be placed, according to the project specifications, on the train gear, is presented.

1 General Introduction

The concept of power harvesting works towards developing self-powered devices that do not require replaceable power supplies. The increase in power used by the electronics has led to a reduction in battery life and has limited the functionality of the devices. In an effort to extend the life and to reduce the volume of the electronics, researchers have begun investigating methods of obtaining energy from the ambient energy surrounding the device. Piezoelectric vibration harvesting devices attract more studies due to its simple design, easy fabrication, high efficiency, high electromechanical coupling, and no external voltage source requirement. Literature cover various topics in power harvesting based on piezoelectric materials, including the design of efficient harvesting geometries, improving efficiency through circuitry, implantable and wearable power supplies, etc. Flexible piezoelectric materials have received the most attention due to their ability to directly convert applied strain energy into usable electric energy and their easy integration into a system. A method to increase the amount of energy harvested from a piezoelectric is to utilize the most efficient coupling mode in a specific architecture configuration. Many conventional systems are characterized by a single piezoceramic in bending mode. In such mode, a force is applied perpendicularly to the poling direction, an example of which is a bending beam that is poled on its top and bottom

surfaces. In this work the authors focused on the characterization of a power harvesting device, using piezoelectric elements to convert vibrations into electricity. The device is designed to be applied in a train derailment detector system and to supply a wireless transmission unit, sending warning messages to the trains driver. The first part of the work is focused on the description of the basic analytical model used for the design, with the addition of a simplified model useful to visualize the correct way to wire piezoelectric elements. An experimental validation of the harvesting system has then been performed by using a slip table simulating the same acceleration and vibration amplitude recorded during the train service. In what follows the attempt to harvest energy by means of a specific architecture to be placed, according to the project specifications, on the train gear, is presented.

2 Analytical model

The analytical model, on which this work is is based, is the one developed by (Roundy). A model of the generator is important not only for estimating the amount of power provided by a given vibration source, but also for making explicit relationships that give to the system's designer some hints on how to improve its performance.

Because this analytical description deals with just one bimorph (i.e.: two piezoelectric elements wired in parallel) the next step in the characterization of a complete piezoelectric model is to study the best way to wire piezoelectric elements. With this goal in mind, a simplified circuit model has been developed (by considering the piezoelectric element like an ideal generator, neglecting the converse piezoelectric effect) and the results are shown in the following section.

2.1 Basic analytical model

The basic configuration is a cantilever beam which consists of two piezoelectric elements wired in parallel to form a bimorph with a central metallic layer. An analytic model can be developed based on beam theory and on the constitutive equations of piezoelectric materials. A convenient method is to model both the mechanical and electrical portions of the piezoelectric system as circuit elements (Figure 1).

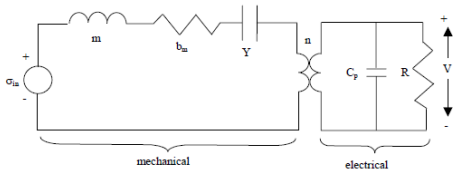


Figure 1: Circuit model of piezoelectric bimorph with resistive load.

The parameters represented in the previous figure summarize the physical characteristics of the model. "σ_{in}" is an equivalent input stress; in other words, it is the stress developed as a result of the input vibrations. "m", shown as an inductor, represents the effect of the mass, or inertial term. The resistive element represents damping, or mechanical loss; the damping coefficient, b_m, relates stress to tip displacement. Finally, the stiffness element is represented as a capacitor and labeled with the elastic constant, Y.

Assuming the model developed by (Roundy), the following expressions of the voltage and power generated can be reported:

$$V = \frac{-j\omega \frac{Y_c d_{tc} b''}{\varepsilon}}{\left[\frac{1}{RC_p}\omega_n^2 - \left(\frac{1}{RC_p} + 2\zeta\omega_n\right)\omega^2\right] + j\omega\left[\omega_n^2(1+k^2) + \frac{2\zeta\omega_n}{RC_p} - \omega^2\right]} A_{in} \quad (1)$$

$$P = \frac{1}{\omega^2} \frac{RC_p^2 \left(\frac{Y_c d_{tc} b''}{\varepsilon}\right)^2}{(4\zeta^2 + k^4)(RC_p\omega)^2 + 4\zeta k^2 (RC_p\omega) + 2\zeta^2} A_{in}^2 \quad (2)$$

The optimal load resistance can then be found by differentiating the power equation with respect to R, setting the result equal to zero, and solving for R. The resulting optimal load resistance is shown below.

$$R_{opt} = \frac{1}{\omega C_p} \frac{2\zeta}{\sqrt{4\zeta^2 + k^4}} \quad (3)$$

2.2 Simplified model useful to wire piezoelectric elements

There are different formulations of the constitutive equations of a piezoelectric element. By using these formulations it is possible to derive two different ways to schematize, in terms of circuit, a piezoelectric element in a very simplified way:

1) A series of a voltage generator and a capacitor with a capacity equal to the piezoelectric element;

2) A parallel of a current generator and a capacitor with a capacity equal to the piezoelectric element;

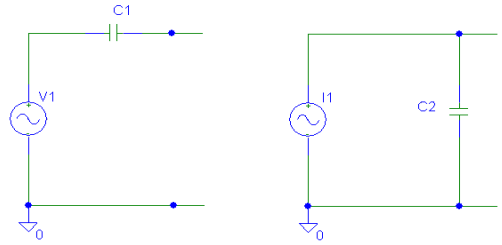


Figure 2: Circuit diagrams of a piezoelectric element.

These two circuits are equivalent, as it could be seen applying Thevenin-Norton theorem. In order to simplify the work it is possible to wire to one of these circuits just a resistive load and to study the behavior of a piezoelectric element under a mechanical vibration at a single frequency, with angular pulsation equal to ω.

The voltage on the resistive load can be found applying the voltage divider rule:

$$V_R = \frac{j\omega RC}{1+j\omega RC}V \Rightarrow \Rightarrow |V_R| = \frac{\sqrt{\omega^2R^2C^2}}{\sqrt{1+\omega^2R^2C^2}}|V| \quad (4)$$

In a first case it has been studied a series connection of $n=2$ piezoelectric elements. By sing (eq. 4) it is possible to write the voltage on the resistive load as follows:

$$V_{R\_S} = \frac{j\omega RC_{eq}}{1+j\omega RC_{eq}}(V_1+V_2) \quad (5)$$

where V_1 and V_2 are the voltages generated by each piezoelectric. Since $V_1 = V_2 = V$:

$$V_{R_S} = \frac{j\omega RC_{eq}}{1+j\omega RC_{eq}}(2V) \Rightarrow \Rightarrow |V_{R\_S}| = \frac{2\sqrt{\omega^2R^2C_{eq}^2}}{\sqrt{1+\omega^2R^2C_{eq}^2}}|V| \quad (6)$$

Generalizing this equation for n piezoelectric elements it is finally possible to write:

$$V_{R_S} = \frac{j\omega RC_{eq}}{1+j\omega RC_{eq}}(nV) \Rightarrow \Rightarrow |V_{R\_S}| = \frac{n\sqrt{\omega^2R^2C_{eq}^2}}{\sqrt{1+\omega^2R^2C_{eq}^2}}|V| \quad (7)$$

where $C_{eq} = C/n$.

It is now important to calculate the ratio between the voltage that could be generated on the resistive load using a series of n piezoelectric elements and the voltage that could be generated using just one:

$$K_S = \frac{|V_{R_S}|}{|V_R|} = \frac{\sqrt{1+\omega^2R^2C^2}}{\sqrt{1+\omega^2R^2C^2/n^2}} \quad (8)$$

In a second case it has been studied a parallel connection of $n=2$ piezoelectric elements. Using the current divider rule and applying Norton theorem ($I_1 = j\omega C_1V_1$ and $I_2 = j\omega C_2V_2$), the current on the resistive load can be written as:

$$I_R = \frac{j\omega(C_1V_1+C_2V_2)}{1+j\omega R(C_1+C_2)} \quad (9)$$

Therefore the voltage on the resistive load can be evaluated as follows.

$$V_{R\_P} = RI_R = \frac{j\omega R2C}{1+j\omega R2C}V \quad (10)$$

where: $V_1 = V_2 = V$ and $C_1 = C_2 = C$.

Generalizing (eq. 10) for n piezoelectric elements it is finally possible to write:

$$V_{R_P} = \frac{j\omega RnC}{1 + j\omega RnC}(V)$$
$$|V_{R\_P}| = \frac{\sqrt{n^2\omega^2R^2C^2}}{\sqrt{1 + n^2\omega^2R^2C^2}}|V| \quad (11)$$

As in the previous case, it has been calculated the ratio between the voltage that could be generated on the resistive load using a parallel of n piezoelectric elements and the voltage that could be generated using just one:

$$K_P = \frac{|V_{R\_P}|}{|V_R|} = n\frac{\sqrt{1 + \omega^2R^2C^2}}{\sqrt{1 + n^2\omega^2R^2C^2}} \quad (12)$$

In Fig. 3 and **Errore. L'origine riferimento non è stata trovata.** it has been represented how the coefficients K_S and K_P vary with the number of piezoelectric elements n. The following values have been used: $\omega = 2\pi 50\left[rad/s\right]$, $R = 2\cdot 10^3[Ohm]$, $C = 90\cdot 10^{-9}[F]$.

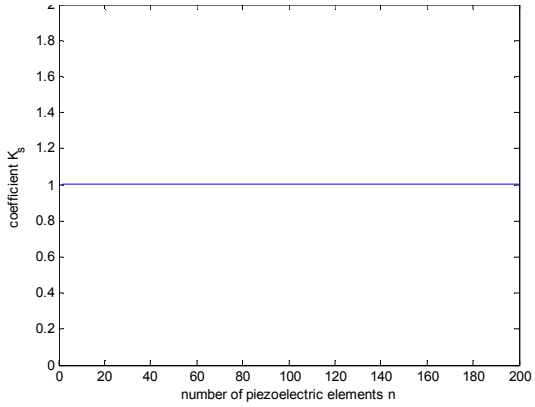


Figure 3: Coefficient Ks (series)

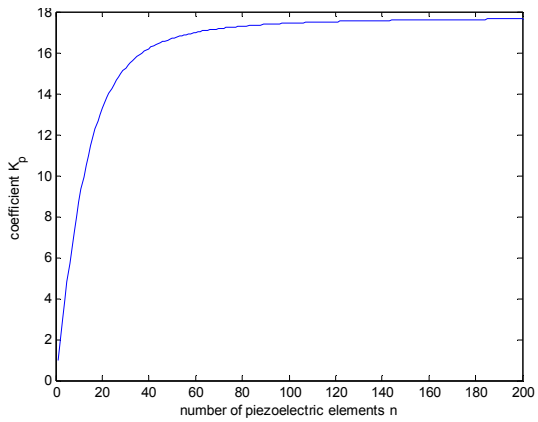


Figure 4: Coefficient Kp (parallel)

From these two graphs an important conclusion has been derived:

It's not convenient to wire piezoelectric elements in series because the voltage generated is about the same as the one generated with a single piezoelectric.

Moreover, if an expression for the optimal resistance has been already calculated in the previous analytical formulation (Roundy), it is possible to derive a simpler equation of the optimal resistance with this model too. The analytical term for the power transferred to the load in a parallel configuration of n piezoelectric elements is given by:

$$P = \frac{|V_{R\_P}|^2}{R} = \frac{n^2\omega^2RC^2}{1 + n^2\omega^2R^2C^2}|V|^2 \quad (13)$$

The optimal resistance can then be found by differentiating this equation with respect to R, setting the result equal to zero, and solving for R.

$$R_{opt} = \frac{1}{n\omega C} \quad (14)$$

Therefore, the optimal resistive load decreases with the number of piezoelectric elements placed in parallel. Fig. 5 and Fig. 6 show the graphs voltage-resistance and power-resistance obtained changing the number of piezoelectric elements in a cantilever beam configuration. The curves are plotted for a sinusoidal acceleration of amplitude 1.1760 m⁄s^2 and frequency corresponding to the first mode of each beam (about 48 Hz).

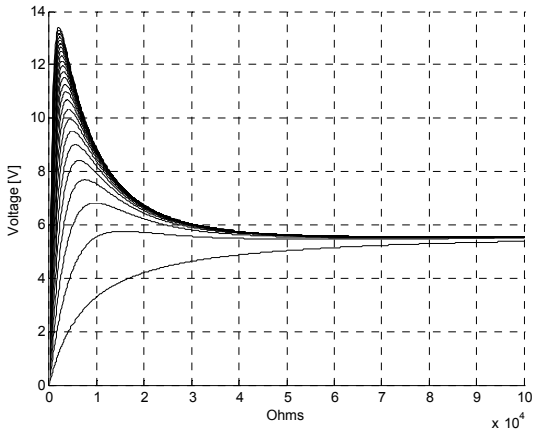


Figure 5: Voltage-Resistance vs number of piezoelectric elements.

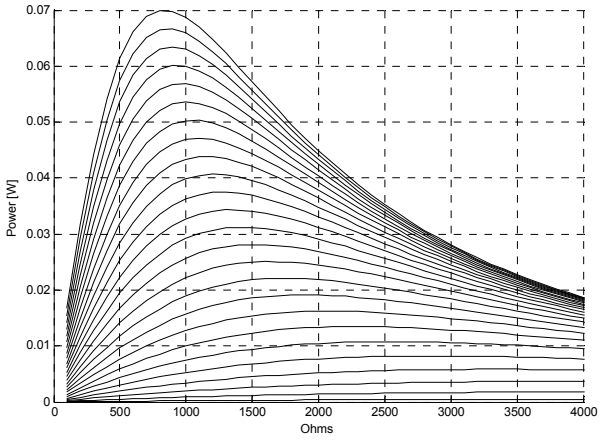


Figure 6: Power-Resistance curves obtained changing the number of piezoelectric elements.

3 Experimental results

An important part of this research activity is focused on the optimization of all those parameters that can improve the efficiency of the piezoelectric power harvesting device. With this goal in mind a series of tests, using a single bimorph, has been carried out.

3.1 Acquisitions on the Freight Train

The measurements made on the freight train during the route from Brindisi to Bari have been performed by using a LMS data recorder, integrated with accelerometers placed in two distinct points of the structure. The monitoring lasted about 5h30' and during this interval of observation the train has stopped, before reaching the destination, 6 times. Five events have been obtained as reported in Table 1.

Tabella 1: Temporal events during the route.

Event	Start time	End time	Description (where available)
1	16:00	16:04	Transfer from the railway station to the station of Brindisi
2	16:42	16:57	Departure from the station of Brindisi and intermediate stop after 15 minutes
3	17:31	18:39	Not available
4	18:54	19:20	Not available
5	19:30	20:40	Arrival at the station of Bari

Three channels were monitored, in particular:

Channel 1: measuring point on the floor (Point_1) in the feed direction (X)

Channel 2: measuring point on the trolley (Point_2) in the lateral direction (Y)

Channel 3: measuring point on the trolley (Point_2) in the vertical direction (Z)

The accelerations are expressed in **g**, which means in fractions of gravitational acceleration. Because of the high amplitudes, just the vertical accelerations of the Point 2 (Figure 7), acquired in the events 3,4 and 5 of **Errore. L'origine riferimento non è stata trovata.**, were considered

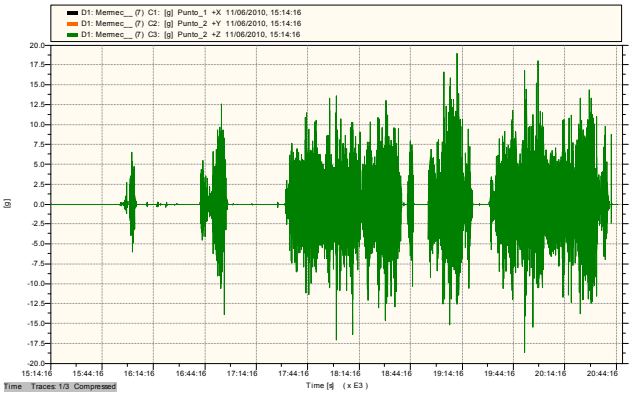


Figure 7: Acceleration acquired on Channel 3.

Figure 8 shows the spectrogram of the acquired accelerations; it is possible to note that, during the entire run of the train, harmonics in the range 40-50 Hz, 70-80 Hz and 95-105 Hz are always present; in particular the highest levels of acceleration are included in the first interval indicated.

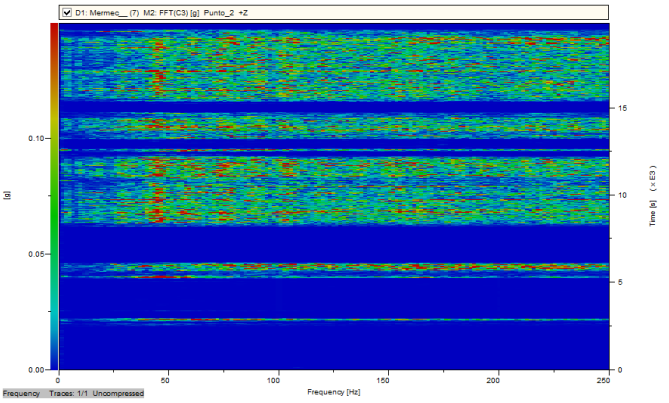


Figure 8: Spectrogram in g of the acquired accelerations.

The experimental validation of the harvesting system has been performed by using a shaker simulating the same acceleration and vibration amplitude recorded during train service. It has been possible to implement a program in LabVIEW and using a random signal generator (DAQPad-6052E, National Instruments). Two time segments have been chosen from the acquisition data and used as input signals for the tests; in this work these signals will be referred as "FIRST SIGNAL" and "SECOND SIGNAL".

3.2 Experimental Validation of the Parallel/Series Configuration

An experimental validation has been carried out by giving in input the first signal. Even if the acceleration given in input is a random signal and not a sinusoidal one, the best way to wire piezoelectric elements is always in parallel. The RMS (Root Mean Square) value of the power generated in both cases, at the same operating conditions (Resistive load=820[Ohm], thickness of the metallic layer=0.6[mm], natural frequency=43.75 [Hz], FIRST SIGNAL) has been calculated and resulted equal to:

$RMS_S = 1.16 \times 10^{-9}\ [mW]$ - Series

$RMS_S = 1.16 \times 10^{-9}\ [mW]$ - Parallel

Other tests have been performed with a higher number of piezoelectric elements (Figure 4.12) and all the results lead to the same conclusion.

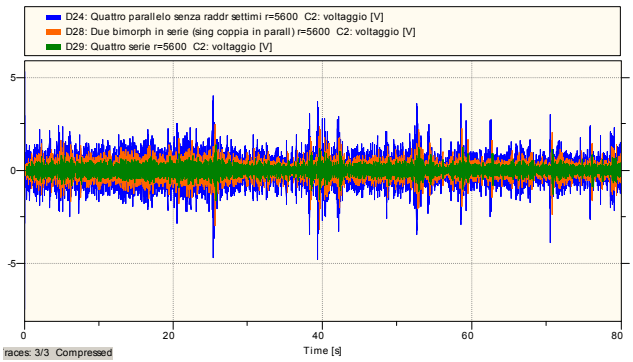


Figure 9: Voltage generated wiring in series and in parallel a higher number of piezoelectric elements under the same test conditions (resistive load R=5600 Ohm).

Reminding that two piezoelectric elements wired in parallel form a "bimorph", in Fig. 9 the blue curve is obtained wiring two bimorph in parallel and, as it is clearly shown, it's the highest voltage generated. The orange and the green curves are obtained by wiring respectively two bimorph in series and four piezoelectric elements in series.

3.3 Optimal configuration with a single bimorph

By varying the end mass value, it is clearly possible to change the natural frequency of the beam. Tests performed with different values of the end mass show that the highest power is generated when the frequency of the driving vibrations is equivalent to the natural frequency of the generator device. In the first column of Table 2 the RMS values of the power generated in the various mass configurations have been reported. The highest voltage and power are generated with the mass configuration "m2". A structural model of the device has been then developed in Patran. A numerical modal analysis has been performed in order to evaluate the natural frequency of the device in correspondence of the "m2" mass configuration, consisting in an end mass of 8.4 [g]. The value of the estimated natural resulted equal to about 43.75 Hz. This result confirms what stated before since the spectrogram of the acquired accelerations (Figure 4.2) shows that the highest level of acceleration is in the interval 40-50 Hz.

Table 2: RMS of the power, using the FIRST SIGNAL.

Mass ID	*P [mW] R = 820 Ohm*	*P [mW] R = 34800 Ohm*
m0	8.86×10^{-2}	
m1	7.73×10^{-2}	
m2	1.89×10^{-1}	3.27
m3	6.36×10^{-2}	
m4	4.58×10^{-2}	
m5	4.91×10^{-2}	
m6	2.50×10^{-2}	
m7	1.93×10^{-2}	

Furthermore, it has been found that the optimal resistive load (the resistive load that maximizes the harvested energy) for a single

bimorph is R=34800 [Ohm], as plotted in Figure 4.11.

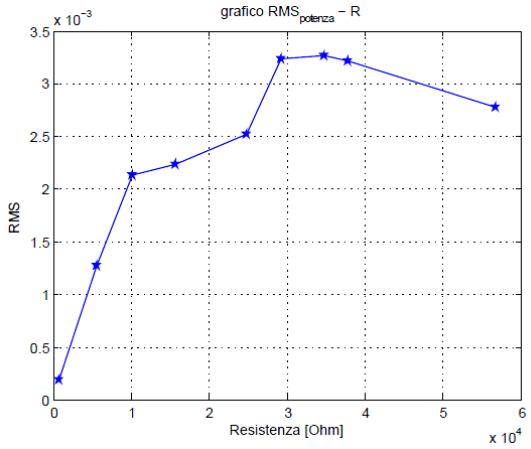


Figure 10: Power (RMS) by a single bimorph wiring different resistive loads.

Several tests have been performed in order to find the most powerful configuration between three different thickness of the metallic layer: 0.6 [mm], 0.8 [mm] and 1.2 [mm]; all the beams have the same natural frequency (43.75 Hz) and work at the same operating conditions (FIRST SIGNAL as driving vibrations and R=820 Ohm). The results show that the metallic layer with a thickness of 0.6 [mm] has to be used to improve the harvested energy.

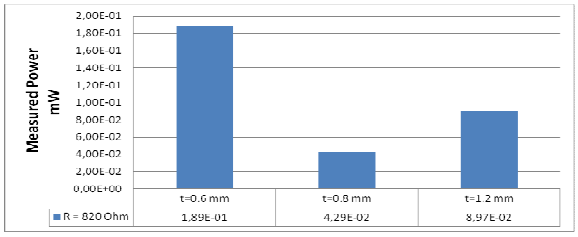


Figure 11: Power generated (rms) with three different thickness of the metallic layer

With reference to the extra length, that represents how the metallic layer is longer than the piezoelectric, the analytical model developed by (Roundy) shows that the optimal electrode length is generally equal to, or very close to, the beam length up to the point where the mass is attached. On the contrary, from the tests performed on the prototype, it has been obtained that the extra length is not a crucial parameter (if it's not too long) because the power generated with an extra length of 35 [mm] is about the same as that generated with no extra length. Summarizing the results obtained until now, it is possible to identify the optimal configuration with a single bimorph:

Table 3: Optimal configuration with a single bimorph.

Optimal Thickness	Natural Frequency	Optimal Resistive Load
0.6 [mm]	43.75 [Hz]	34800 [Ohm]

This configuration has been tested by using the SECOND SIGNAL as driving vibrations. The RMS value of the power calculated in this acquisition is: RMS_{1B}=8.7 [mW].

3.4 The Converse Piezoelectric Effect

The simple analytical model, presented in previous sections, schematizes the piezoelectric element as an ideal generator. However, the presence of the converse piezoelectric effect has to be considered in the design of a real device.

If multiple piezoelectric materials are attached to the same circuit in an attempt to produce more electric energy, the energy loss will be very high. The reason for this is that the energy generated by one piezoelectric transducer causes the converse piezoelectric effect to occur at the other transducers, resulting in consumption of a part or all of the generated energy. Also, further losses occur due to the destructive electric signal interference produced from each piezoelectric device.

According to the previous considerations, a piezoelectric energy harvesting system should include individual diode bridge circuits that may be attached to the outgoing wires from each piezoelectric generator. The outgoing wires from each diode bridge may be connected to a single energy harvesting circuit with minimal or no energy loss. This allows for the use of an unlimited number of piezoelectric generators at the same time on the same, or single, energy harvesting circuit.

Considering the optimal configuration with a single bimorph, a diode bridge has been connected to each piezoelectric element and a test, using the SECOND SIGNAL as driving vibrations, has been performed.

The RMS value of the power calculated in this acquisition is: $RMS_{1B,(bridge)}$=6 [mW], with a loss of approximately 30% compared to the power generated without a diode bridge. It has

been found that the loss due to the use of a diode bridge is higher than the one due to the converse piezoelectric effect, so it has been decided not to use such devices.

3.5 Experimental Results with More Bimorph Wired in Parallel

The optimal configuration with a single bimorph gives, as seen before, a RMS value of the power generated equal to 8.7 [mW]. In order to improve this value, several tests have been performed by wiring in parallel two, four and five bimorph. The optimal configurations and the RMS values of the generated power are shown below.

Table 4: Operating conditions for each test.

ID Test	Driving Vibrations	N. of Bimorph	Optimal Sheet Thickness	Optimal Resistive Load	Natural Frequency
001	SECOND SIGNAL	2	0.6 mm	5600 Ohm	43.75 Hz
002	SECOND SIGNAL	4	0.6 mm	4500 Ohm	43.75 Hz
003	SECOND SIGNAL	5	0.6 mm	2000 Ohm	43.75 Hz

Knowing that the optimal resistive load with one bimorph is 34800 [Ohm] and looking at the values reported in the fifth column of Table 4.5, it is possible to confirm that the optimal resistive load decreases with the number of piezoelectric elements placed in parallel. In Table 5 the RMS values obtained working with one, two, four and five bimorph, operating at the optimal configuration, have been listed.

Table 5:RMS values of the voltage and the power.

N° Bimorph	Power RMS [mW]	Voltage RMS [V]
1	8.70	11.2
2	12.6	6.02
4	20.0	6.91
5	26.6	5.16

As expected, the energy generated does not increase linearly with the number of bimorph used. This happens because of the converse piezoelectric effect and because of the destructive electric signal interference produced by each piezoelectric. The last phenomenon can be reduced by adopting a “comb configuration” that decreases the mechanical phase shift.

3.6 Linear Relationship between Voltage and Acceleration

A final result that has been highlighted during the tests, is the linearity of the voltage developed across a single bimorph with the acceleration provided in input to the shaker.

The measured voltages (giving as driving vibrations the FIRST SIGNAL, whose amplitude is scaled by different factors, and working with: thickness of the sheet t_s=0.6 [mm], resistive load R=820 [Ohm], natural frequency f_n=43.75 [Hz]) are reported (as RMS value) in Table 6 these values are plotted with the scale factor of the accelerations.

Table 6: Voltage generated across a single bimorph with different amplitudes

Voltage RMS Original Amplitude	Voltage RMS 1/2 Amplitude	Voltage RMS 1/5 Amplitude	Voltage RMS 1/6 Amplitude
0.2551 [V]	**0.1313 [V]**	**0.0503 [V]**	**0.0409 [V]**

This linearity is not maintained if more bimorph are wired in parallel. The tests have been performed using the following values: thickness of the sheet t_s=0.6 [mm], resistive load R=4500 [Ohm], natural frequency f_n=43.75 [Hz] and the SECOND SIGNAL, whose amplitude is scaled by different factors, as driving vibrations. Even if it’s hard to see, this new curve is not linear anymore but, increasing the scale factor of the signal, the voltage increases more than linearly. In order to focalize it better, the RMS values of the voltage generated with each scale factor are reported in Table 7.

Table 7: Voltage generated across four bimorph with different amplitudes of the driving vibrations (SECOND SIGNAL).

Voltage RMS Original Amplitude	Voltage RMS 1/2 Amplitude	Voltage RMS 1/3 Amplitude	Voltage RMS 1/4 Amplitude	Voltage RMS 1/7 Amplitude
6.91 [V]	**3.13 [V]**	**2.07 [V]**	**1.51 [V]**	**0.814 [V]**

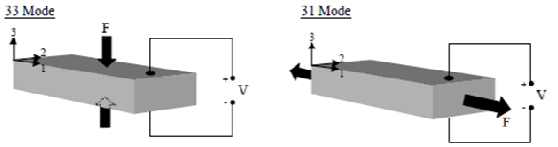
4 Description of the Designed Device

Although piezoceramic (PZT) is widely used, its extremely brittle nature causes limitations in the strain that can be safely absorbed without damages , especially susceptible to fatigue crack growth when subjected to high frequency cyclic loading. In order to eliminate the disadvantages of piezoceramic materials and improve upon

their efficiency, in this application a more flexible, piezoelectric material which exhibits considerable flexibility in the range 10Hz - 1MHz without electrode damage has been used. Furthermore the piezoelectric element chosen has a high coupling coefficient (k) and a high capacitive load (C). The piezoelectric element used for the device is P-876.A11 DuraAct.

Typically, piezoelectric material is used in the 33 mode, meaning that both the voltage and stress act in the 3 direction (Figure 4.11). However, the material can also be operated in the 31 mode, meaning that the voltage acts in the 3 direction (i.e. the material is poled in the 3 direction), and the mechanical stress / strain acts in the 1 direction. Although the electrical/mechanical coupling for 31 mode is lower than for 33 mode, it has been chosen to operate in a 31 mode because there is a key advantage: the system is much more compliant, therefore larger strains can be produced with smaller input forces and the resonant frequency is much lower.

Table 8: Illustration of 33 mode and 31 mode operation of piezoelectric material.



A cantilever beam configuration with a mass placed on the free end has been chosen for two reasons. First, the cantilever mounting results in the lowest stiffness for a given size; second, the cantilever configuration results in the highest average strain for a given force input. Because the converted power is closely related to the average strain in the bender, a cantilever mounting is preferred.

The designed structure is a series of cantilevered beams clamped at their center and here excited along the transversal direction by the vertical acceleration of the train. Each beam consists of two piezoelectric with a metallic layer sandwiched between them. The layer is made of spring steal because of its high yield strength which gives an excellent fatigue behavior. A “comb configuration” has been adopted in order to reduce the mechanical phase shift between the beams and because of its compact shape that optimizes the space requirement. A conceptual design of the device and the prototype used for the tests is shown in Fig. 12

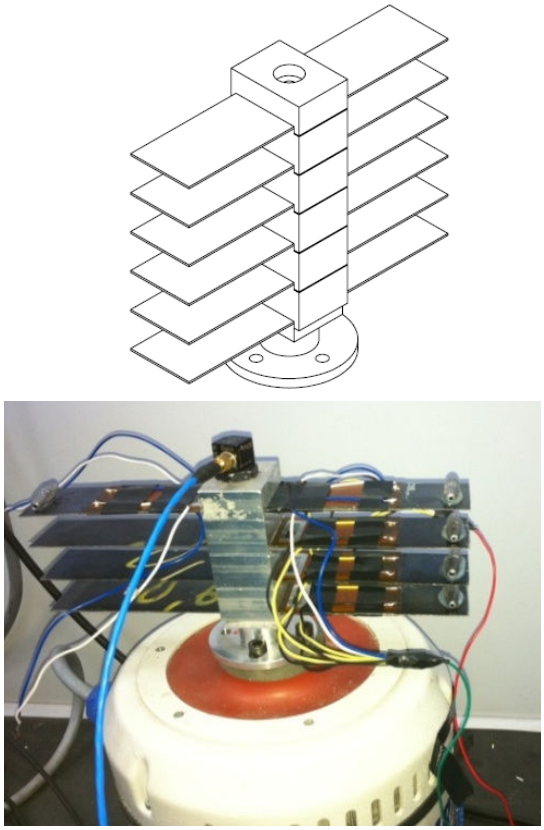


Figure 12: Conceptual design of the devise and the prototype

5 Conclusion

A rational approach based on analytical and numerical model was followed in order to design, to manufacture and to test the energy harvesting system.

In order to study the energy harvesting system based on more interconnected piezoelectric element, a simplified numerical model was developed in Pspice(r) language and validated; it was verified that the parallel connection of piezoelectric elements is more efficient than series connection.

In order to design and to test the device in operative conditions, experimental measurement were performed on freight train. This measurement showed that during the entire run of the train, harmonics in the range 40-50 Hz, 70-80 Hz and 95-105 Hz are always present, but the highest levels of acceleration are within the

first indicated range. It was then designed and manufactured a prototype system with 5 bimorph. In operative condition the realized system produced about 27 mW in rms.

Acknowledgments

The authors wish to acknowledge MerMec SpA technical office for the support to all the activities which made possible this study and the experimental test on freight train.

References

[1] Lee B S, He J J, Wu W J and Shih W P 2006, "MEMS generator of power harvesting by vibrations using piezoelectric cantilever beam with digitize electrode", Proc. Smart Structures and Materials Conf.; Proc. SPIE 6169 61690B

[2] Lee C S, Joo J, Han S, Lee J H and Koh S K 2005, "Poly(vinylidene fluoride) transducers with highly conducting poly(3,4-ethylenedioxythiophene) electrodes", Proc. Int. Conf. on Science and Technology of Synthetic Metals vol 152 pp 49–52

[3] Mohammadi F, Khan A and Cass R B 2003, "Power generation from piezoelectric lead zirconate titanate fiber composites", Proc. Materials Research Symp. p 736

[4] Sodano H A, Lloyd J and Inman D J 2004a, "An experimental comparison between several active composite actuators for power generation", Proc. Smart Structures and Materials Conf.; Proc. SPIE 5390 370–8

[5] Sodano H A, Park G and Inman D J 2004b, "A review of power harvesting using piezoelectric materials", Shock Vib. Digest 36 197–206

[6] Sodano H A, Park G and Inman D J 2004c, "Estimation of electric charge output for piezoelectric energy harvesting Strain", 40 49–58

[7] Baker J, Roundy S and Wright P 2005, "Alternative geometries for increasing power density in vibration energy scavenging for wireless sensor networks", Proc. 3rd Int. Energy Conversion Engineering Conf. (San Francisco, CA, Aug.) pp 959–70

[8] Platt S R, Farritor S and Haider H 2005°, "On low-frequency electric power generation with PZT ceramics", IEEE/ASME Trans. Mechatronics 10 240–52

[9] Ng T H and Liao W H 2005, "Sensitivity analysis and energy harvesting for a self-powered piezoelectric sensor", J. Intell. Mater. Syst. Struct. 16 785–97

[10] Mateu L and Moll F 2005, "Optimum piezoelectric bending beam structures for energy harvesting using shoe inserts", J. Intell. Mater. Syst. Struct. 16 835–45

[11] Roundy S 2005, "On the effectiveness of vibration-based energy harvesting", J. Intell. Mater. Syst. Struct. 16 809–23

[12] D. Benasciutti, E. Brusa, L. Moro, S. Zelenika "Optimised piezoelectric energy scavengers for elder care", Proc. 10th International EUSPEN Conference, Zurich, 41-45, 2008.

[13] Mossi K, Green C, Ounaies Z and Hughes E 2005, "Harvesting energy using a thin unimorph prestressed bender: geometrical effects", J. Intell. Mater. Syst. Struct. 16 249–61

[14] Kymissis J, Kendall D, Paradiso J and Gershenfeld N 1998, "Parasitic power harvesting in shoes 2nd IEEE Int. Conf. on Wearable Computing", (August) pp 132–6

[15] Ericka M, Vasic D, Costa F, Poulin G and Tliba S 2005, "Energy harvesting from vibration using a piezoelectric membrane", J. Physique. Coll. IV 128 187–93

[16] Kim S, Clark W W and Wang Q M 2005a, "Piezoelectric energy harvesting with a clamped circular plate: analysis", J. Intell. Mater. Syst. Struct. 16 847–54

[17] Kim S, Clark W W and Wang Q M 2005b, "Piezoelectric energy harvesting with a clamped circular plate: experimental study", J. Intell. Mater. Syst. Struct. 16 855–63

[18] Kim H W, Batra A, Priya S, Uchino K, Markley D, Newnham R E and Hofmann H F 2004 "Energy harvesting using a piezoelectric 'cymbal' transducer in dynamic environment", Japan. J. Appl. Phys. 43 6178–83

CEAS 2011

Space Systems Design and Analysis

Structural Development of a Detachable Preburner Technical Model

Jaehan Yoo, Ilyoon Moon, Insang Moon, Soo Yong Lee

Future Launch Vehicle Technology Department

KARI (Korea Aerospace Research Institute), Korea

Keywords: *Liquid Rocket Engine, Preburner, Structural, Analysis, Test*

Abstract

A detachable preburner technical model is being developed by KARI. This paper describes the general design, the results of the finite element analyses (FEA) and the experiment results related to structural issues during the development. Before the fabrication, two burst experiments with the specimens simulating the brazing joints were performed and compared with the FEA results, which increases the reliability of the analyses and the design. Also, 2D FEA of the combustion chamber cooling channel were performed for the 1st (before the Lox hole) combustion region and 2nd (after the Lox hole) one, where the hottest spot locates behind the region between the adjacent Lox holes. And various structural modifications were analyzed by 3D FEA and the screen structure was selected to reduce the stress over highly stressed regions. During the development, some proof and leakage tests were performed for the head, the chamber with and without flange and the preburner assembly, of which burst test was performed. Several firing tests of the assembly were performed to show no structural failure.

1 Introduction

Staged combustion liquid rocket engines can minimize the cycle loss and maximize the efficiency. Thus they are broadly used for high performance space launch vehicles. A Preburner is one of the most important components of the staged combustion engine, required high technologies regarding oxygen rich combustion under the high temperature and pressure [1,2]. A detachable preburner technical model is being developed by KARI. The 'detachable' means that its head can be detached from the chamber and switched with another head, which have different injector groups. That makes fire tests for various groups inexpensive.

This paper describes the general design, the results of the finite element analyses (FEA) by ABAQUS v6.9 [3] and the experiment results related to structural issues during the development. The operating pressure and temperature of the model considered is over 200 bar and 2000 K, respectively. And the chamber of the preburner have the regenerative cooling channel by Lox and the lox holes at its middle to emit the Lox and lower the exit or turbine inlet temperature by about 700 K.

2 Materials and Fabrication

Figure 1 depicts the geometry and two materials used for the preburner model, which consists of austenite SS (stainless steel) 316L and chrome-copper alloy [4]. The former material which has relatively high strength is used for reinforcing the model and the latter with high conductivity used for fast cooling the surfaces contacting the hot gas by using liquid oxygen.

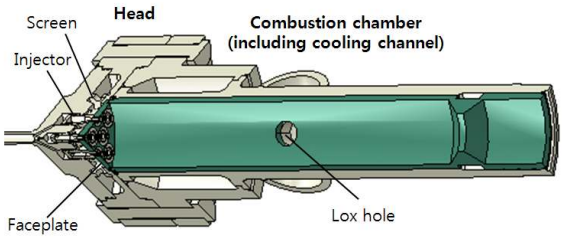


Fig. 1 Configuration of the preburner (white : SS 316L, green : copper alloy)

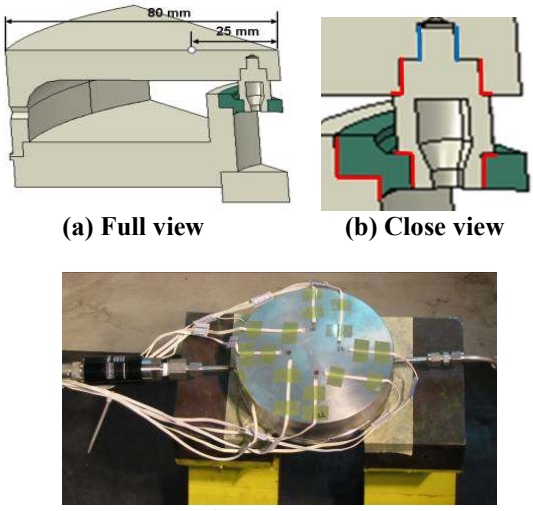
The joints of the 316L injectors and the 316L head and the joints of the injectors and the copper alloy faceplate were joined by the conventional vacuum brazing. On the other hand, the copper alloy cooling channel and the 316L outer wall of the combustion chamber were brazed by vacuum compression one.

The material 316L has somewhat low yield strength for the reinforcing and will be changed to new one that has higher mechanical properties to reduce the weight. For the new stainless steel, yield strength and toughness at room and cryogenic temperature after brazing thermal treatment, joint strength with the copper alloy used and easy procurement may be considered. The selected one, a well known SS, will be used for a single body perburner, which is not detachable, to reduce the weight of the reinforcing parts by about half.

3 Brazing Joint Strength Tests

Before the fabrication, the burst experiments with two specimens simulating the brazing joints were performed and compared with the FEA results, which increases the reliability of the analyses and the design and checks brazing filler settings on the joints and brazing capability of a factory.

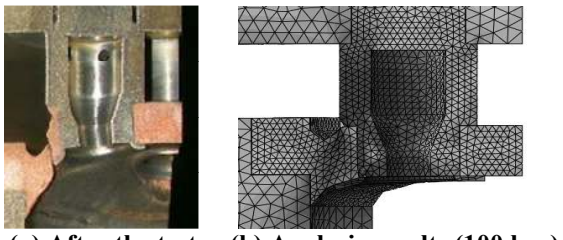
The first specimen depicted in Fig. 2 simulated the brazing joints of the oxygen injectors and the head and one of the injector and the faceplate. In preliminary FEA results, the highly stressed regions were found to be at the latter joints, so the specimen was designed to be under similar stress state as the joints. In the specimen, the four dummy injectors and strain gauges (#1~#4) directed to radial direction are attached to the top cover. The burst test was stopped at about 150 bar where the sudden pressure decrease occurs without external leakage.



(a) Full view **(b) Close view**

(c) Burst test

Fig. 2 Configuration of the head specimen (1/8 model, white circle : strain gauge, red line : brazing, blue line : thread)



(a) After the test **(b) Analysis results (100 bar)**

Fig. 3 Head specimen test results (dense color means high Von-Mises Stress)

Fig. 3 shows the specimen after the test and wire cutting and Von-Mises stress distribution of the static elastic-plastic FEA result (100 bar), where dense color means high stress. All the analyses in this study were performed using the

material stress-strain curve up to the ultimate tensile strength. During the experiment, it was observed that the fractures were occurred not at the brazed joining but in the middle of the face, where the materials are the copper alloy. In addition, the analysis showed that the predicted fracture locations were well matched with the experimental results.

Table 1 shows the strains measured and analyzed for the head specimen. The measured strains from #1 to #4 gauge are represented by percentage differences from the FEA results and have discrepancy each other due to various non axi-symmetries of manufacturing and brazing due to trivial lack of the filler, where the part design was changed for the fabricated preburner. The differences are small below the 100 bar and increases abruptly above 100 bar, which may represent the crack initiation and/or propagation. The corresponding joints of the fabricated preburner at an operating condition are under stress state less than the state of the specimen at 100 bar.

Table 1 Comparisons of head specimen strains between the experiments and analyses

Pressure (bar)	Strain				
	Test (difference, %)				Analysis (μ)
	#1	#2	#3	#4	
20	-11	4	12	4	132
50	-20	-4	10	4	421
70	-13	4	22	15	628
80	-1	17	42	31	737
100	45	70	113	92	972
130	197	325	529	476	1490
150	157	-	598	454	2200

The second specimen as in Fig. 4 simulated the brazing joints of the inner and the outer wall for the combustion chamber cooling channel. The width and the height normalized by the rib thickness and the number of the channel are 1, 3.125, 11, respectively. The specimen has round notch for highly stressed region at the center brazing joints and nine strain gauges are used. Then burst test are performed up to 750 bar, where the copper alloy inner wall fractures at the center as in Fig 5(a) and the external leakage occurs.

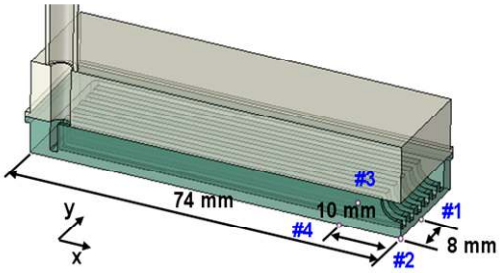
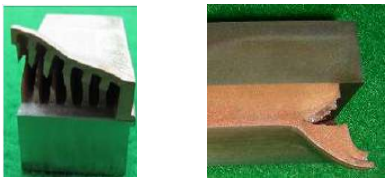


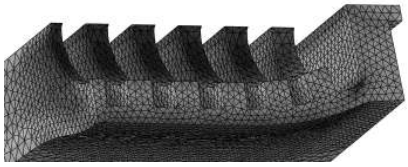
Fig. 4 Configuration of the chamber specimen (1/4 model, white circle : strain gauge)



(a) Full view after test



(b) Section view after test



(c) FEA results (Von-Mises Stress)

Fig. 5 Test and analysis results of cooling channel specimen (dense color means high Von-Mises Stress)

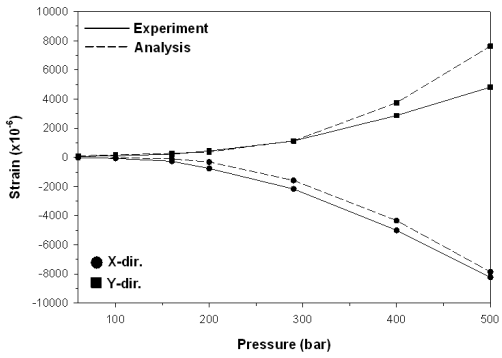


Fig. 6 Comparisons of strains measured and analyzed (gauge #4)

At the test, it was observed that the fractures were occurred not at the brazed joining but in

the middle of the copper alloy cooling wall as the head specimen, which was showed in Fig. 5(b). Also, the predicted fracture locations were well matched with the test as in Fig 5(c). The measured strain at the gauge #4 is compared with FEA results in Fig 6. The results were presented up to 500 bar only, where no crack propagation was thought to be developed. The agreement between the results is well each other.

4 Cooling Channel FE Analyses

Two-dimensional plane strain structural analyses of the combustion chamber cooling channel including thermal and pressure loading were performed for the 1st (before the Lox holes) and 2nd (after the Lox holes) combustion regions. At the Lox holes, a large amount of liquid oxygen is emitted to the chamber.

The analysis procedure consists of two steps. The first is the steady state CFD (computational fluid dynamics) analysis using the commercial program, Fluent [5], to obtain temperature distributions of the channel section. The calculated maximum temperature for the 1st and the 2nd regions is near room temperature and over 500 K, respectively.

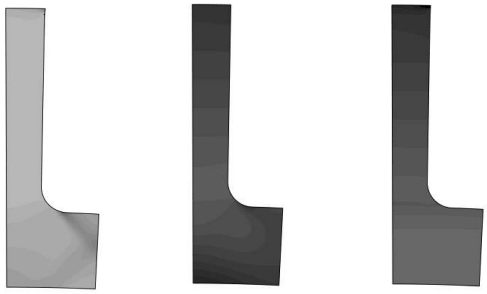
The second is the static FE structural analysis using Abaqus to obtain von-Mises equivalent or effective strain, ε, distributions. The pressure loadings due to the Lox in the channel and the hot gas in the chamber are considered to the analyses, which are elastic-plastic ones without creep and strain rate effects because the mechanical properties related to them are hard to obtain and to be considered in the analyses, also its effects on the ε may be small [6].

During development, the one of important structural targets for the cooling channel is not to fail by low cycle fatigue of several ground firing tests. One of ways to access the fatigue life is using a certain strain quantity like the ε, of which concept is the root of the square sum of three difference terms for the three principal strain components [6]. The ε is normalized by the structural target hereafter. The target is set for the copper alloy part to have over 250 cycles of low cycle fatigue life using the fatigue test datum [7].

For the first region, the half channel model was used and the temperatures of four points were obtained from the first step and the distributions between the points were assumed to decrease linearly to radial direction as in Fig. 7. The pressure difference of the cooling channel from the combustion chamber is 20 bar.



Fig. 7 Temperature distribution of the half channel model for 1st combustion region (dense color means high temperature, datum from CFD analysis in the red circles).



(a) case 1 **(b) case 2** **(c) case 3**

Fig. 8 Effective strains of the copper alloy part for 1st combustion region from FEA (dense color means high value).

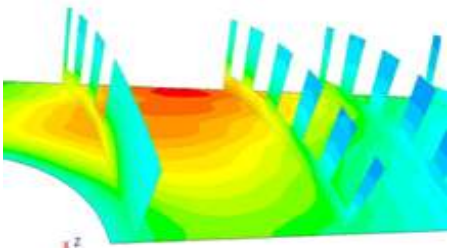
For the special case #1 under pressure loading only, the section is under tensile stress in circumferential direction except the rib and compressed radically. The effective strain distribution is depicted in Fig. 8(a) and the maximum is 0.15.

For the special case #2 under thermal loading only in Fig. 8(b), the channel is compressed for both directions and maximum is 0.28 satisfying the target. The main mechanism is that the outer jacket tends to contract while inner jacket tends to expand, then the constrained force on the inner and the ε occurs.

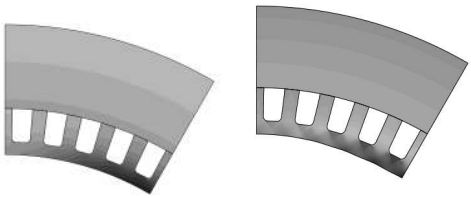
For the case #3 under thermal and pressure loading in Fig. 8(c), the channel is under the similar stress state as the previous. The effect of pressure loading on the ε is relatively small compared to the thermal loading. For the three loading cases above, plastic deformation occurs

and the joints are compressed radically.

The CFD analysis as in Fig. 9(a) show that the hot spot, which have highest temperature points in the whole combustion region, locates behind the region between the adjacent Lox holes in the second combustion region. One section including the spot was selected for the FE analyses. The five-channel model considering Lox hole symmetry was used in the second step because the circumferential temperature variation was significant. Fig. 9(b,c) shows the temperature distribution from the CFD analysis and effective strain one from the FE analysis for the selected section. The ε is relatively high at the channel wall where high temperature regions locate. The max. ε of the section is 1.0 near the target.



(a) 3D temperature from the CFD Analyses

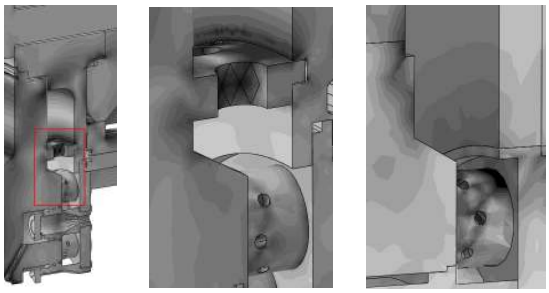


(b) 2D temperature for FEA (c) ε from FEA

Fig. 9 Analysis results of hot spot region (red color means high value in the rainbow contour)

5 Effects of the Head Screen

Since the axial separating load due to the chamber pressure is sustained by the flange bolts and the junctions of the faceplate and the injectors in a certain head section (refer to Fig. 1), moreover the bolt is more far from the head center than the junctions, thus high stress on the junctions is induced. Various structural modifications to reduce the stress were analyzed by 3D FEA and the screen structure in Fig. 10 was selected and optimized for its thickness and size and number of its holes. Additionally, some CFD analyses show it enhances the flow uniformity to the injectors.

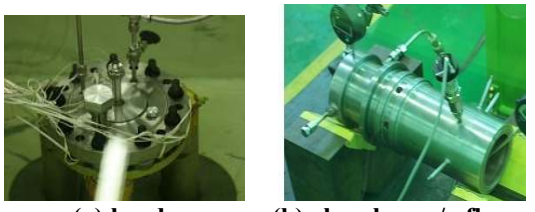


(a) with screen (b) w/o screen

Fig. 10 Screen structure effect at the head from 3D FEA (Von-Mises stress, dense color means high Von-Mises Stress).

6 Proof and Leakage Tests

During the development, some proof and leakage tests were performed for the head, the chamber with and without the flange and the preburner assembly as shown in Fig. 11, where the white lines are for strain gauges. At the water proof test of the chamber without the flange, the pressure exists in the cooling channel only and at the test with the flanges, the pressure in the channel and chamber.



(a) head (b) chamber w/o flange



(c) chamber w/o flange (d) preburner assembly

Fig. 11 Proof tests by water

At all the proof test except the final preburner assembly test, the test pressures were set not to yield except local areas to minimized damages. The test pressure of the assembly was set to 1.2 x off-design condition x operating pressure.

The measured strains during the proof tests were compared with the results computed, where the measures strains at chamber outer surface for preburner assembly are smaller than the predicted, then the assembly is determined to pass the proof test.

The leakage is assured by the bubble test and the pressure transducers. The burst test at a pressure over two times operating condition shows no external leakage and no internal structural failure.

7 Firing Tests

The several firing tests of the final assembly as in Fig. 12 were performed up to 57 % of the operating chamber pressure and any structural failure or large deformation was not detected.



Fig. 12 Firing test of the preburner.

8 Conclusion and Future Works

The structural analyses and tests during the development of the detachable preburner technical model were presented. The structural design including the brazing joints and the Lox cooling channel is thought to be reliable for continuing firing tests in the future through the proof, burst and firing tests and FEA analyses.

In detail, the specimen burst tests for the brazing joints show that the fractures occur not at the joints but at the copper alloy. The screen structure of the head enhances the local area strength effectively. The cooling channel analyses gave some minimum fatigue life of the model. The model was validated at the proof and burst pressure for the room temperature. In several firing tests up to 57 % of the operating pressure, any structural failure was not detected.

The firing tests of the operating condition will be performed only soon, then after fixing of the injector design, the affixed head version of the preburner model will be fabricated with the aforementioned new material.

References

[1] Huzel, D. K., *Modern Engineering for Design of Liquid-Propellant Rocket*, Progress in Astronautics and Aeronautics, Washington DC, 1992, 105.

[2] Sutton, G. P. and Biblarz, O., *Rocket Propulsion Element*, 7th ed., John Wiley & Sons, Inc., New York, USA, 2001, pp. 383-384.

[3] *Abaqus Analysis User's Manual*, v6.9, Dassault Systems, RI, USA, 2009.

[4] Ryu C., Chung Y. H., Choi H. and Lee D., "Structural Analysis of Liquid Rocket Engine Thrust Chamber Regenerative Cooling Channel at Room Temperature", *Journal of the Korean Society of Propulsion Engineers*, Vol. 9, No. 4, 2005, pp. 39-47.

[5] *Ansys Fluent User's Guide*, release 13.0, Ansys, PA, USA, 2010.

[6] Ryu C., "A Study on the Structural Design of Regenerative Cooling Chamber of Liquid Rocket Thrust Chamber," Ph.D. Dissertation, Department of Mechanical Engineering, Chungnam National University, Korea, 2006.

[7] Lee K., Ryu C. and Choi H, "Life Prediction of Copper Alloy of Combustion Chamber", *Journal of the Korean Society of Propulsion Engineers*, Vol. 15, No. 2, 2011, pp. 44-49.

AUDIBLE NOISE CONTROL OF TAS-I MANNED PRESSURIZED MODULES ON INTERNATIONAL SPACE STATION

S.Destefanis(1), N.Riva(1), R.Ullio(1), V. Tumminelli(2)

(1) ThalesAleniaSpace Italy – Strada Antica di Collegno, 253, 10146 Torino, Italy
stefano.destefanis@thalesaleniaspace.com, Tel: 00390117180554, Fax: 00390117180274
nicola.riva@thalesaleniaspace.com, Tel: 00390117180349, Fax: 00390117180274
roberto.ullio@thalesaleniaspace.com, Tel: 00390117180349, Fax: 00390117180274

(2) Sofiter System Engineering S.p.A., C.so Francia, 35 10138 Torino, Italy
valentina.tumminelli@external.thalesaleniaspace.com, Tel: 00390117180554, Fax: 00390117180274

Abstract

On-orbit vibro-acoustics has found exciting applications in modern manned space systems because of audible noise/human vibration requirements aimed at human comfort and demanded to have suitable conditions for scientific investigations.

Due to the severity of these requirements, which implies a system vibro-acoustic control over a wide range of frequencies, the classical design has demanded strong updating in terms of physical representation of the mechanical system and in terms of analysis approach.

Particularly, central aspects are: (1) the on-board disturbance sources experimental characterization, (2) the simulation of the internal acoustic cavity behaviour coupled with the structural behaviour and (3) the use of both deterministic and statistical numerical approaches as complementary tools to assess the requirement covering the whole frequency range of interest.

The integrated test/analysis control approach developed, during these years, by ThalesAleniaSpace-Italy covers both the evaluations of the structural and vibro-acoustic transmission paths and the characterization of the disturbers in terms of force at the equipment interfaces and Sound Pressure Level in the surrounding volume.

1 Introduction

A dedicated methodology on how to derive experimentally the induced disturber's interface forcing functions was also developed in the frame of this method.

In this paper, results of MPLM, NODE 2, ATV and COLUMBUS modules design and verification activities versus the International Space Station (ISS) audible noise requirements are presented.

The principal characteristic of this control approach is the capability to report the most important outcomes during the whole process cycle and the evolution of the project relating to the audible noise requirement. It starts from a preliminary audible noise environment assessment (based mainly on Space-lab experience), identification, design and implementation of passive recovery actions as mufflers, acoustic box and soft-mounted devices for compliance achievement.

Subsequently, the scope is passing through an even more detailed audible noise environment assessments (using the FEM and SEA dynamic response analysis and the disturber's interface forcing functions and SPL derived from unit tests).

Lastly a complete audible noise system test is performed allowing both the characterization of the structural and vibro-acoustic system transmissibility, by the use of artificial exciters, and the evaluation of the audible noise environments by the module disturbance sources activation.

The development of this design methodology began during the late 80's, while its first

industrial application was the MPLM (Multi Purpose Logistics Module) project. The design success achieved showed that this control approach had a great potential. Then the methodology was further refined in the frame of NASA/ASI NODE2 and also for the European project COLUMBUS and it has been also applied for the NODE3, CUPOLA and PMM module.

The NODE2 / NODE 3 and COLUMBUS on-ground qualification by test and analysis against audible noise requirement demonstrated the correctness of ThalesAleniaSpace-Italy developed design, analyses and verification method.

In addition to the on-ground audible noise requirements verification, another key effort, explained in this paper, is focused to the COLUMBUS payload complement on-orbit audible noise compliance verification necessary to demonstrate the compliance of the payload complement with the respective requirements and to define operational guidelines and constraints when element level conflicts are identified.

2 European manned space facilities evolution

Spacelab (figure 1) was designed to guarantee an adequate On-Orbit Vibro-acoustic environment in order to allow Scientific experiments at 10^{-3}g and an internal acoustic environment < NC50 for an acceptable working and speech communication conditions. During the whole Design cycle several recovery actions were implemented to control airborne and structure borne noise. The main designed and implemented noise reduction devices were focused on duct mufflers, acoustic boxes and Anti Vibration Mounts. The performances got during the flights were within:

- $4 \cdot 10^{-5}$ g to $3.6 \cdot 10^{-3}$ g induced by life support system and auxiliary equipment
- Around NC60 (Mission configuration dependent)

Based on this experience it was decided to define dedicated Noise and Micro-g control plans at the beginning phase of the Design. These Control plans defined the Design and the Verification tasks, control model, sub-requirement, and responsibilities on System, sub-system and equipment level. It also organized all the necessary activities from system level down to equipment level to not exceed the specified limits. Moreover the control plan defined all the activities (Analysis and Test) for the implementation and verification of the On-Orbit Vibro-acoustic requirements.

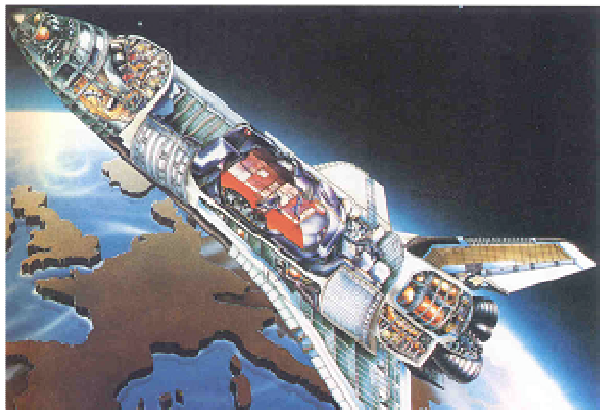


Fig. 1 Spacelab and Shuttle-Past Micro-g space facility

Based on this lessons learned dedicated control plans for the Audible noise, Human Vibration and the Micro-g environment control was implemented at the beginning of the project on all manned spacecraft designed by ThalesAleniaSpace-Italy and concurred by EADS, ASI and ESA. The manned modules designed by this design strategy are: MPLM, Columbus, Cupola, Node#2 and 3 and partially ATV.

Figure 2 shows the ISS configuration.

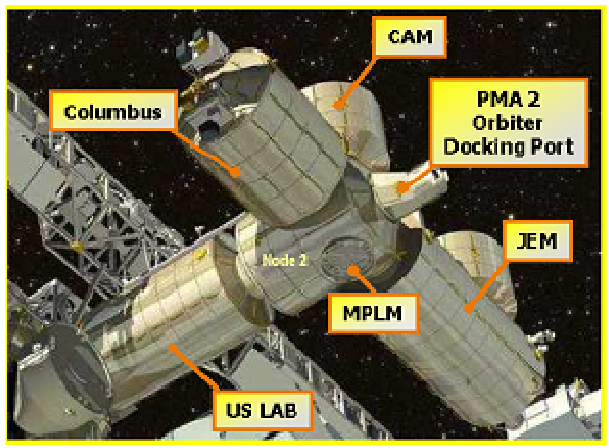


Fig. 2 ISS–International Space Station.

3 European on-orbit vibro-acoustic environment control approach

ThalesAleniaSpace-Italy has defined with EADS and ESA a design verification strategy to assess by analysis the vibro-acoustic environments (i.e. audible noise aspect as crew comfort) induced by the on-board disturbance sources (mechanism, valve, fan, water pump, ducting, etc.). The innovative

budget allocation concept has been introduced and applied to the design cycles. Based on manufacturer's data and spacecraft structural and vibro-acoustic transmissibility, tailored contribution factors have been allocated, starting from the global audible noise environment, down to each identified disturbance source (as shown in fig.3).

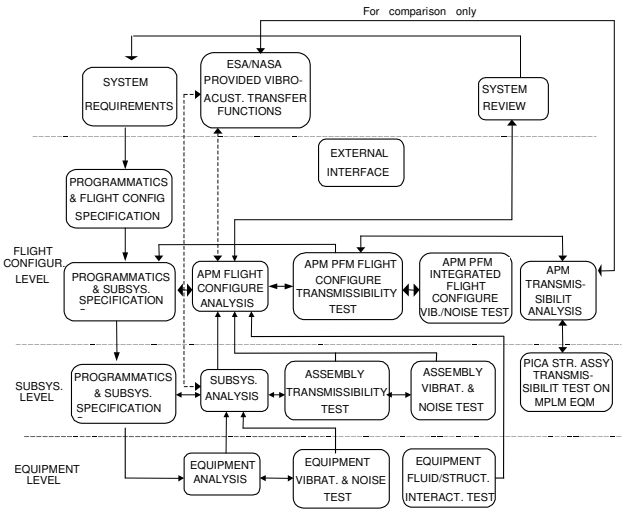


Fig. 3 Typical On-Orbit Vibro-Acoustic Control Approach - Design, Specification and Verification Responsibilities

Following each analysis/test cycle, the budget is updated to reflect the current disturbance source contribution. This budget apportionment allows effective and real-time design verification at each stage of the project and helps to assess margins and critical areas as well as suitable recovery strategies, if needed.

3.1 Design strategy to Audible noise environment control

For the *Audible Noise* environment control, the system sound pressure level requirement cannot be used as a function of disturber emitted acoustics requirement. It defines a sound pressure with regard to the acoustical properties of the receiving volume (the module cabin) which is not available for verification noise measurements at lower level. Instead, sound power levels and budgets shall be considered as sub-requirements. The adopted noise control approach (Figure 4) shall maintain a sound power balance between all acoustical power sources after transmission into the module cabin and the acoustical power absorbed by the cabin interior such that the resulting sound pressure level criterion is not exceeded. Sound power is that quantity radiated by the source which is not dependent on environmental factors.

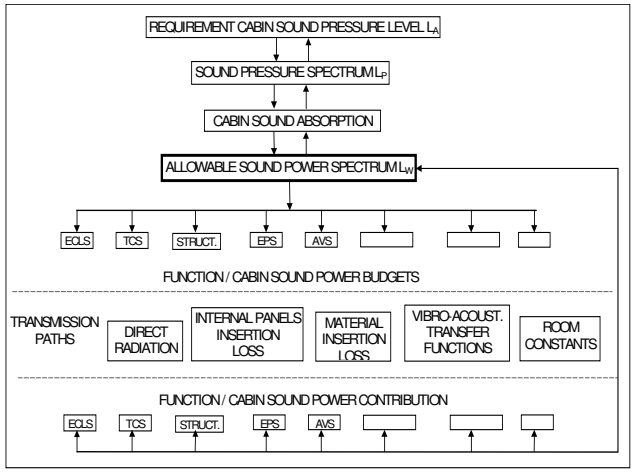


Fig. 4 Typical Sound Power Noise Control Model

4 First industrial application: MPLM

The MPLM project is the first pressurized module that applies the design approach discussed in this paper for the vibro-acoustic control.

Dedicated control plan (for the audible noise environment control) was implemented as a guideline for the project development during its design cycle, controlling all levels, from the equipment up to the system, considering the contributions of all the disturbance sources in order to converge to the system compliance. Initially an assessment was done to derive a preliminary estimate of the potential induced internal acoustics. Also, to define the budget allocation for each subsystem and equipment in terms of induced sound power considering the TAS-I database in terms of forcing functions (I/F structural forces), acoustic power and vibro-acoustic transmissibility). Then, the assessment was refined on the basis of the disturbance measurement and of the system mathematical models availability. By this cross-checking activity, it was possible to recover passive control means mainly for ECLS and TCS equipment. Even if the ECLS and the TCS are the major noise contributors, design effort was spent also on the elimination of the stick/slip effect (which is due to thermal gradients in the MDPS structural joint areas) by implementation of isostatic boundary condition.

The total cabin sound power was derived by superimposing the noise induced by the ECLS, TCS, AVS and Structure subsystems.

This assessed performance was compared to the allowable cabin sound power and then converted into sound pressure level, considering the acoustic properties of the cabin (i.e. averaged room absorptivity).

During the MPLM design cycle, the following design improvements have been implemented:

- Cabin fan acoustic box
- Cabin fan inlet and outlet mufflers
- Cabin fan AVM
- Teflon ring at the TCS ducting I/F's
- Soft ring at the ECLS ducting I/F's

The MPLM flight system configuration employed the above noted design interventions and the final assessment has been performed, showing that the compliance with respect to the NC50 requirement is fully met over the whole frequency range (see Table 1).

Frequency	31.5	63	125	250	500	1000	2000	4000	8000
ECLSS LW		60.4	49.9	57.6	54.0	52.6	49.1	46.0	44.9
ATCS LW	64.2	56.2	52.3	47.7	43.2	39.8	36.4	35.7	35.1
STR LW	63.0	55.0	51.1	46.5	42.0	38.6	35.2	34.5	33.9
AVS LW	No emission at steady-state condition								
SYSTEM LW	66.7	62.6	56.0	58.3	54.6	53.0	49.5	46.7	45.6
Cabin PWL Specified (NC 50)	77.0	71.0	67.1	61.5	56.0	53.1	50.2	49.5	48.9

Table 1 MPLM predicted performance vs. NC50 requirement

The QR audible noise test was conducted at TAS-I premises on the flight model, simulating the on-orbit open hatch condition by means of a dummy hatch expressly designed. The QR test activity confirmed the predicted compliance as shown in Figure 5.

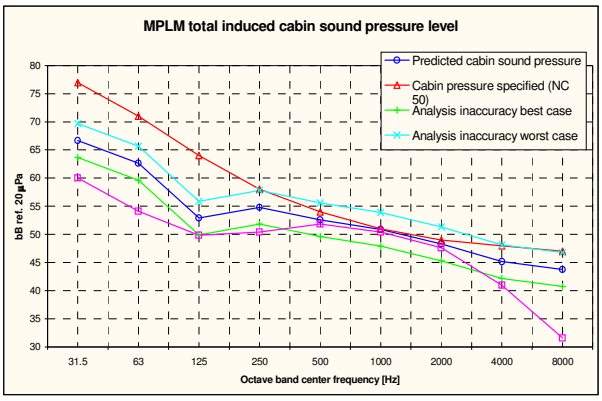


Fig. 5 MPLM predicted and measured performance vs. NC50 requirement

Moreover, it should be noted that the on-ground test configuration is conservative, due to the lack of payload racks and partition walls. By these considerations, it can be assumed that the cabin noise could be further reduced.

This industrial application demonstrates that the design strategy based on the integrated test/analysis approach fulfils the NASA NC50 requirement for the audible noise comfort of the crew.

5 Consolidation of the industrial experience: Node#2 and Columbus project

Node 2 (Fig. 6) and Columbus (Fig. 7) are two pressurized modules of the International Space Stations. They allow, through ports, the attachment of the Space Station elements. As of today, they are both docked to the ISS.

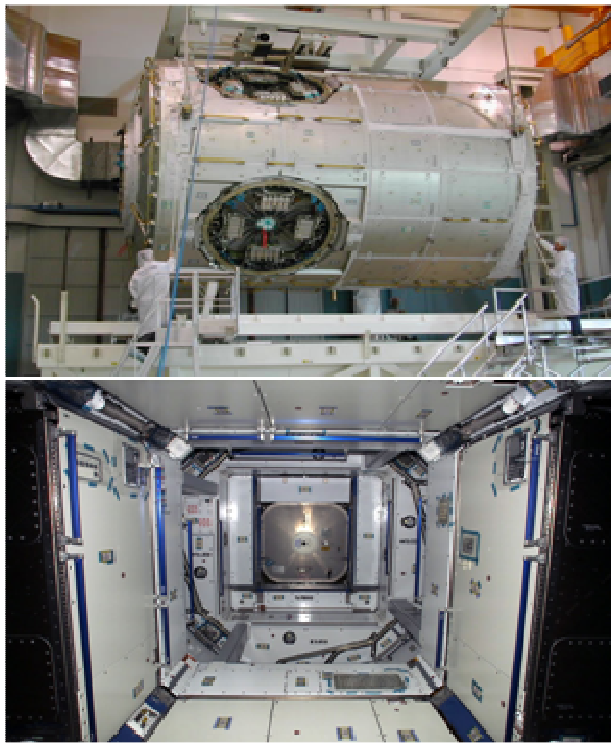


Fig. 6 Node#2 external and internal architecture at ThaleAleniaSpace-Italy integration facility

From structural point of view the two modules are very similar. They show an overall length of about 7 meter a diameter of 4.6 meter and a mass of about 14 tons. Node#2 and Columbus include an electrical power distribution system, an active and passive thermal control system, an environment control system (air revitalization and distribution, internal temperature control, pressure control and fire detection and suppression) and an audio and video distribution system. All these subsystems during on-orbit phase generate noise that needs to be

contained as much as possible for two main reasons: to reduce crew annoyance and more important to avoid permanent damage to hearing of astronauts. NASA puts particular attention on this subject. For this reason the audible noise requirements imposed to Space Station Modules are very severe.

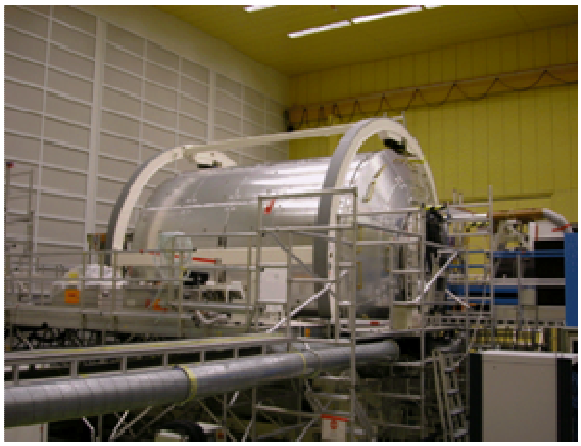


Fig. 7 Columbus external architecture at EADS integration facility

Performed Node#2 and Columbus qualifications by tests and analysis against audible noise environmental comfort requirements (i.e. internal vibro-acoustics) demonstrated (on industrial high complex projects) the correctness of the European developed design method.

The uniqueness of such method comes mainly from controlling mechanical induced disturbances, generated by individual sources like mechanisms, aerodynamic fluxes, fluid dynamics and thermal gradients. Such gradients give origin to intermittent "clicks" at structural junction level, totally random in terms of time history, which could make the crew grow nervous and, in the end, become less efficient. Disturbers control approach (quite uncommon) is based on mechanical and acoustic power generated by the disturber itself. Mechanism designers, aerodynamic circuits' designers, fluid dynamics designers and of course structural designers, using this approach are allowed to verify the correctness of their design "stand-alone", without the need of having a whole-integrated module available. Another key factor is evaluating the structural-acoustic interaction between the module shell dynamic behavior and the internal cabin acoustic environment. Thanks to this new methodology becomes possible, for each single disturbance source, to evaluate the impact generated both at cabin acoustic environment level and at scientific payload micro-dynamic environment level.

5.1 NODE 2 audible noise predicted performance

The following figure 8 shows the TAS-I Node#2 module measured audible noise performances vs. the specified NASA requirements.

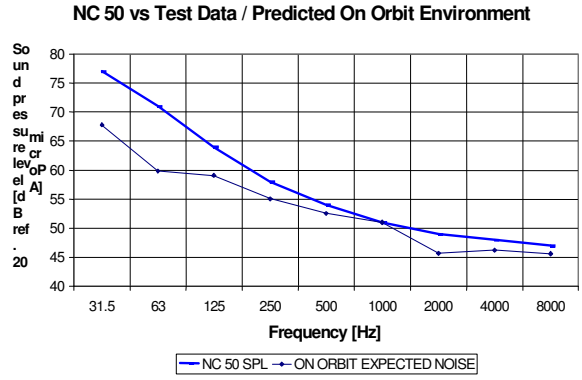


Fig. 8 Node#2 on orbit acoustic environment (prevision based on ground test) vs NC50 requirement (SPL dB ref. 20 µPa) - dB(A) = 56.1

NASA, delighted by Node#2 design success, awarded the whole ThalesAleniaSpace-Italy team with the following rationale: "*for the resounding success of Node 2 Audible Noise, Human Vibration and Microgravity test, the culmination of several years of design and integration efforts*".

5.2 COLUMBUS audible noise predicted performance

The following figure 9 and figure 10 shows the TAS-I Columbus module measured audible noise performances vs. the specified NASA requirements.

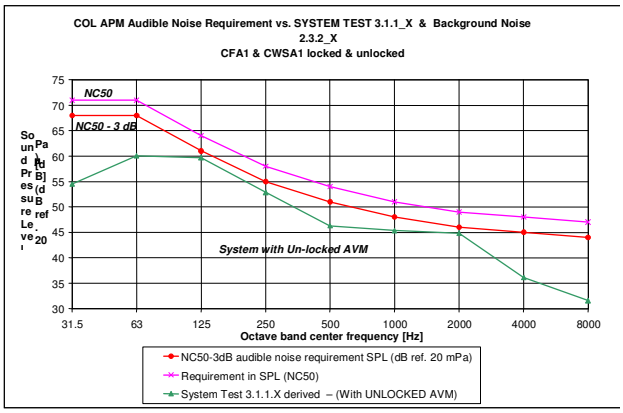


Fig. 9 Columbus Audible noise Steady State quality vs. the NC50 requirement

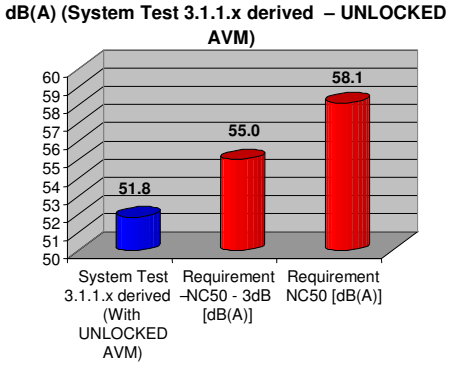


Fig. 10 Columbus Audible Noise Overall performance

Lastly, the measured T60 = 0.4s at 500 Hz inside the cabin vs. the preferred Reverberation Time of 0.3s (for the cabin volume of 25 m^3 see fig.11) shows that the acoustic quality of the Columbus Cabin Room is closer to the Conversational speech than the General-Purpose curve. Therefore, the material used for the design of the cabin guarantees a good intelligibility of the human speech inside cabin.

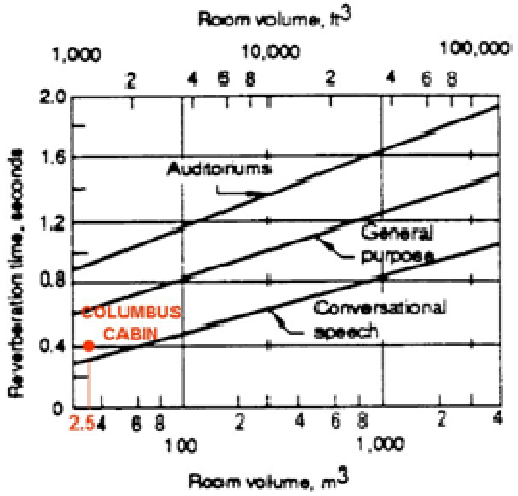
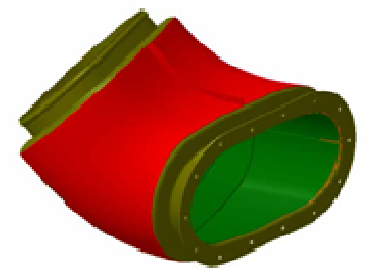


Fig. 11 Columbus Reverberation Time quality vs. T(60) (s) requirement



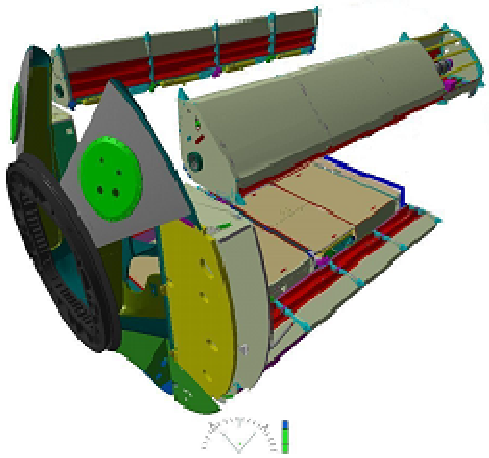


Fig. 12 Typical passive recovery action – Airborne control - Muffler solution and Partition walls

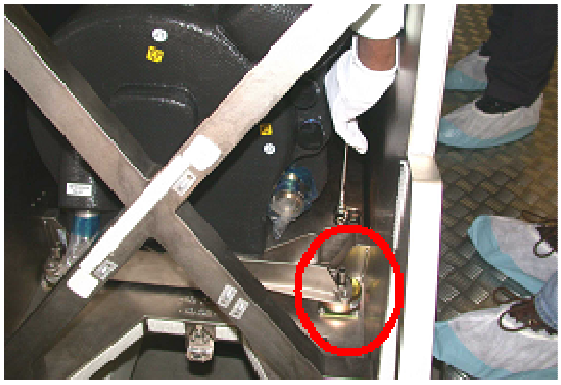


Fig. 13 Typical passive recovery action - structure borne control – AVM

5.3 Design recovery actions on Audible Noise Performances

MPLM, Node#2 and Columbus are (as of today) the only manned modules of the International Space Station to gain compliance versus NASA NC50 environmental requirement (see Figure 8 and 9). Such performances were achieved by means of vibro-acoustic environment control systems, like AVMs (Anti Vibration Mounts) to control structural disturbances and mufflers / silencers / partition walls to control airborne noise (Figure 12 and 13).

Supplementary investigation was also conducted, for example, on Columbus WPA2. In fact an extra noise anomaly was detected during the Columbus QR testing and confirmed by subsequent diagnostic test done at assy level.

Sound intensity testing (figure 14) showed clearly that the cause of this acoustic emission was due to the hammering effect of the accumulator internal expansion vessel at those frequencies. The recovery action selected to recover this extra noise was to wrap the accumulator by a multilayer cover (Nomex case, Poron and Durette multilayer barrier).

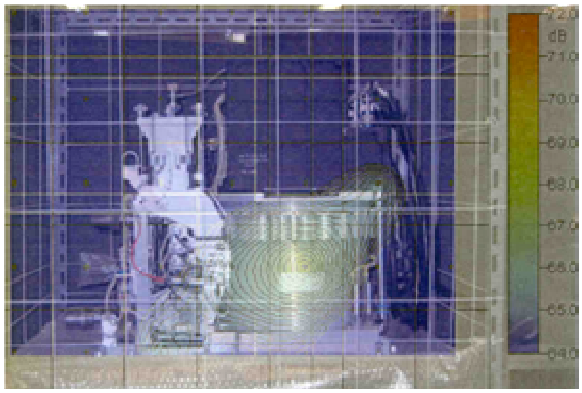


Fig. 14 Columbus WPA Sound Intensity Measurement – Equipment Sound disturber evaluation

5.4 Columbus/Payload integrated system: expected on orbit audible noise performance

The Columbus laboratory is nowadays the ESA's biggest single contribution to the International Space Station.

The Columbus cylindrical module is equipped with flexible research facilities that offer extensive science capabilities, as, for example:

- **BIOLAB**. It supports experiments on micro-organisms, cells and tissue cultures, and even small plants and small insects.
- The European Physiology Modules Facility (**EPM**). It is used to investigate the effects of long-duration spaceflight on the human body. Experiments also contribute to an increased understanding of age-related bone loss, balance disorders and other ailments back on Earth.
- The Fluid Science Laboratory (**FSL**). It accommodates experiments in the strange behaviour of weightless liquids.
- The European Drawer Rack (**EDR**). It is a modular and flexible experiment carrier system for a large variety of scientific disciplines, providing basic accommodation and resources for experiment modules housed within standardised drawers and lockers.
- The European Transport Carrier (**ETC**). It accommodates items for transport and stowage. In orbit ETC serves as a workbench and stowage facility.
- The Microgravity Science Glovebox (**MSG**). The device allows astronauts on board the ISS to perform a wide range of experiments in a fully sealed and controlled environment, completely isolated from the rest of the Station. It shares nevertheless the weightlessness of orbit.

These payloads can be activated or deactivated and sometimes provided of particular experiments to exploit the Columbus capabilities as an orbital Laboratory to its maximum.

The Columbus audible noise on orbit environment must be evaluated w.r.t. its active payloads (how many of them are active, according to detailed timelines). Payloads can operate both as "***steady state***" disturbers and as "***intermittent***" disturbers: each operating mode is regulated by dedicated requirements.

A designated period of on-orbit ISS configuration after each flight which adds capability to the ISS is defined as "*Stage*", while the time period (typically 2 or 4 months of ISS operations) from the undocking of a Soyuz spacecraft to the undocking of the next Soyuz is defined as "*Increment*".

A certain combination of active payloads is identified in each "*stage/increment*" that must be separately analyzed for requirement compliance.

Up to now 13 stages and 5 increments different configurations have been analyzed. Hereafter is reported, as an example, the results of the **stage 1J audible noise analysis**.

Operational constraints and ESA MedOps Approval when necessary have been introduced.

Steady State operative modes

The audible noise assessment was based on the combination of the entire operative conditions of COLUMBUS integrated with the European Payloads EDR, FSL, EPM, BIOLAB, ETC and the NASA Payloads MSG, ER-3, HRF-1 and HRF-2.

The following power scenario constraints have been introduced:

- Power scenario 1: BIOLAB+EPM+ER3+MSG
- Power scenario 2: ER3+FSL+MSG
- Power scenario 3 & 5: BIOLAB+EPM+FSL+MSG
- Power scenario 4: BIOLAB+EPM+HRF2+MSG
- Power scenario 6: BIOLAB+EPM+HRF1+HRF2
- Power scenario 7: EPM + ER3 + HRF-1 + HRF-2

Each aforementioned power scenario operates with each Columbus steady state continuous operative mode *for a total of 1477 operative cases*.

The Columbus stage 1J audible noise assessment shows that Columbus integrated with European payloads fulfils the NC50 Audible noise requirement; except for few slight exceedances (<1.66dB) at 250Hz and 1KHz octave bands. The audible noise disturbers that mainly contribute to these slight exceedances have been detected.

Anyway the SPL of 58.1dbA (which is the NC50 overall SPL) is not exceeded, then the payload racks can be operative without any voice contamination, annoyance and hearing conservation reasonable impact. So no operational constraints or guidelines have been identified.

All scenarios analyzed for steady state operations indicate no need to switch off payloads to comply with the applicable audible noise requirements.

Intermittent operative modes

For each Columbus stage/increment configurations, the constraints for intermittent operations scenarios are mainly driven by the maximum allowed operations of WOOVs and WFSVs. However taking into account the number of valves present in Columbus and the limit (set at 30 minutes cumulative at 69 dBA), it comes to a maximum of 18 activations per valve per day, which is operationally not realistic and consequently it is not considered as a constraint.

When Payload racks intermittent operative modes are not compliant with the intermittent noise level requirements, agreement on how to proceed is concurred with ESA MedOps.

5.5 Columbus On-Orbit Audible Noise test

For each ISS stage/increment configurations NASA provides the ISS Sound Level Meter Survey and usually the audible noise inside the Columbus module is around the NC-45 requirement. However in the Increment 26 measurement Columbus was so quiet as to be almost unbelievable NC-41.1

On-Orbit measurements were taken in the Columbus module at seven locations on the module centerline shown in Figure 15 and the noise levels are shown in Table 2 and Figure 16.

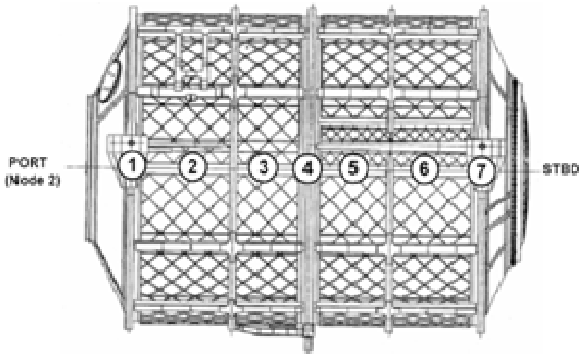


Fig. 15 Measurement (on orbit) locations on Columbus centerline

Columbus measurements			12/29/2010	day 363						
Control #	1	2	3	4	5	6	7			
List #		List #60		List #61						
	Port Hatch 1ft	Rack Bay 1	Rack Bay 2	Center	Rack Bay 3	Rack Bay 4	Stbd endcone	average	NC50	NC47, Columbus Rqmt w/o Payloads
	49.0dBA	47.5dBA	45.7dBA	44.7dBA	43.9dBA	42.2dBA	40.2dBA			
	NC45.2	NC42.6	NC41.0	NC39.5	NC38.1	NC36.2	NC34.4			
Measure-ment #		M56		M58						
Corrected time tag	15:05:59	15:06:24	15:06:45	15:07:11	15:07:37	15:08:03	15:08:43			
63	50.4	46.2	44.1	44.6	43.9	40.6	40.4	45.7	71	68.6
125	47.2	46.4	44.5	42.7	42.3	41.1	43.1	44.7	64	61.6
250	45.1	45.7	42.5	42.7	41.3	41.6	38.4	43.1	58	56.0
500	43.7	42.8	40.8	39.8	39.5	37.1	35.4	40.7	54	51.0
1000	48.2	43.6	42.0	40.5	39.1	37.2	35.4	42.1	51	48.0
2000	38.1	38.8	37.7	36.2	36.7	34.3	32.8	36.9	49	46.0
4000	38.9	36.9	33.8	35.2	33.2	32.3	29.1	35.1	48	45.0
8000	34.8	35.4	33.0	32.2	31.6	29.8	26.2	32.7	47	44.0
dBA	49.0dBA	47.5dBA	45.7dBA	44.7dBA	43.9dBA	42.2dBA	40.2dBA	46.7dBA	58.1dBA	55.4dBA
NC	NC-45.2	NC-42.6	NC-41.0	NC-39.5	NC-38.1	NC-36.2	NC-34.4	NC-41.1	NC-50.0	NC-47.0
SIL(4)	41.7dB	40.5dB	38.5dB	37.9dB	37.1dB	35.2dB	33.2dB	38.7dB	50.5dB	47.5dB
Inter-ference Noted by Crew:										

Table 2 Columbus on orbit measured noise levels (29 December 2010)

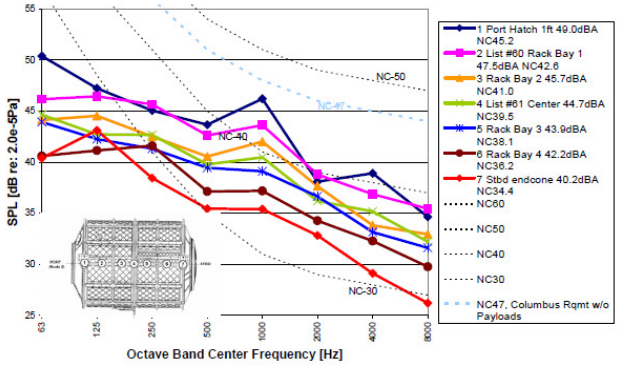


Fig. 16 Columbus on orbit measured noise levels (29 December 2010)

6 ATV On-Ground Diagnostic Test

In the frame of ATV project (Figure 17), the aforementioned design control approach was not fully applied due to the mission of the spacecraft considered as a "cargo" for the ISS refurbishment.

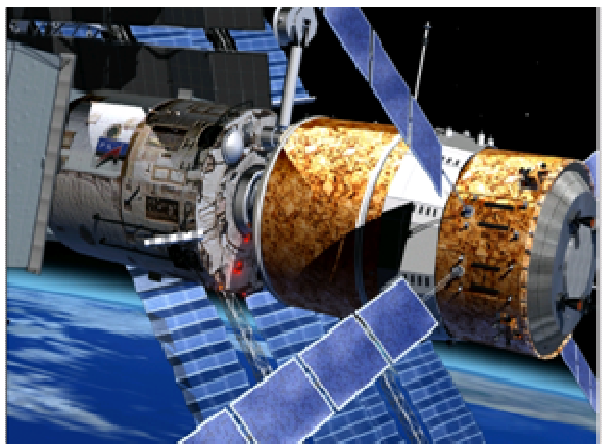


Fig. 17 ATV and Node#2 on orbit, docked to the International Space Station (artist concept)

During the qualification process a not acceptable internal acoustics was detected. Since no local design recovery action was possible to implement on disturber design, a diagnostic strategy was set-up to see the major contributor to the cabin noise. This diagnostics was based on an audible noise test using a set of microphones positioned from the main disturbers and in the centre of cabin as shown in figure 17.

The outcomes of the testing indicate that there are three most important sound source disturbers:

- Intake system: Cabin Return Grid
- Cabin Fan: Cabin Fan Assy
- Exhaust system: Diffusers and End Restrictors

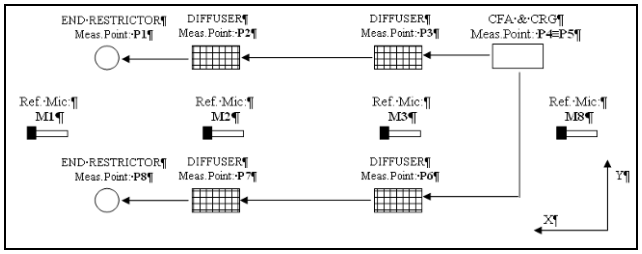


Fig. 17 ATV Diagnostic Microphones locations

A dedicated assessment has been done to derive the expected SPL for each of the three identified disturber families and to see their impact on the overall SPL.

By this way it was possible to evaluate the effect of recovery actions on the overall ATV SPL simulating the new SPL expected by Reference Microphones reducing in a certain percentage one or more sound sources.

After this iterative design verification process the following passive recovery actions have been implemented on ATV Jules Vernes for internal acoustics control (Current baseline configuration):

- ➔ **CRG acoustic box**: double wall concept, to control the noise directly injected into the cabin by the Intake system
- ➔ **Conveyer:** implementation of a conveyer on CRG acoustic box to guide the inlet noise to a phono-absorbent panel implemented in front of the conveyer
- ➔ **Multilayer partition wall:** a phono-absorbent panel is implemented on the shell in front of the conveyer to reduce the reflected acoustic field by the shell.
- ➔ **CFA Inlet and Outlet mufflers**
- ➔ **CFA wrapping:** for CFA airborne noise reduction.
- ➔ **CFA soft-mounted interface:** selected rings and bushing concept design for 4 CFA feet.

In the frame of the On-Ground diagnostic test one accelerometer was also placed on the panel close to the fan interface to evaluate if the panel structural vibrations can influence the measured internal SPL. A hammer test on the CFA interface panel was performed to evaluate if the avionic panel has a resonance in the octave band of 500Hz that could be excited by the CFA forcing function inducing a synergetic effect on emitted noise. The response of this accelerometer is shown in figure 18. By this check a clear resonance frequency around 550 Hz was found.

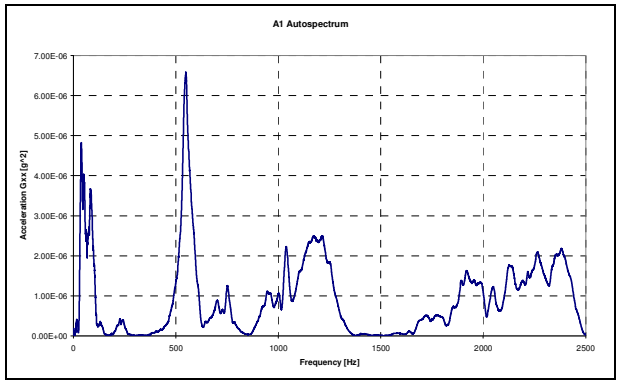


Fig. 18 hammer test➔ acceleration (auto-spectrum) acquired on avionic panel.

In order to evaluate if the avionic panel is really excited at this resonance (producing noise as a loudspeaker), the fan was structurally decoupled from the avionic panel (Figure 19). In this figure is clearly evident that the vibrations of Cabin Fan panel influences the SPL measured by Internal Microphones. Consequently a better effective anti-vibration mount at CFA-to-panel mechanical interface is needed.

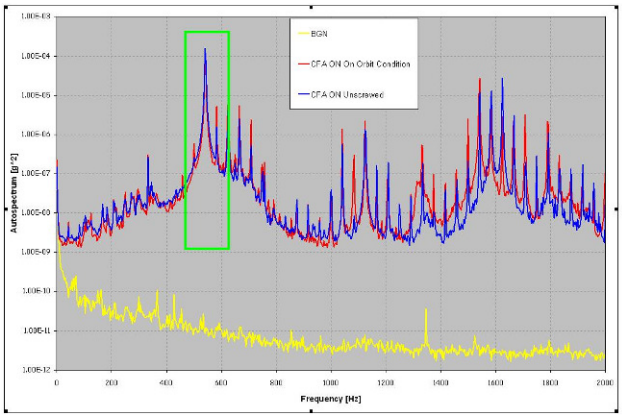


Fig. 19 CFA "on"➔ acceleration (auto-spectrum) acquired on avionic panel

The following acoustic passive means have been implemented on ATV Jules Vernes and tested:

- ➔ CRG acoustic box
- ➔ Conveyer
- ➔ Multilayer partition wall
- ➔ CFA acoustic box
- ➔ CFA wrapping
- ➔ CFA Inlet & Outlet mufflers
- ➔ CFA soft-mounted by Armaflex washer by reducing on orbit by 1/3 of the nominal screw torque

Even if the expected ATV internal SPL is non compliant vs. NC50 in some octave bands (up to +9dB at 2000Hz), the expected overall value (63.05dBA) is lower than the ESA SpaceLab one (67.51dBA) as shown in table 3 and in figures 20, 21 and 22.

Based on this consideration and by the fact that ATV is a cargo with a specified maximum continuously acoustic exposure of 8 hours the non compliance vs. NC50 can be accepted without any voice communication degradation and hearing damage risk.

	October 2005 acoustic test	January 2006 acoustic test			March 2007 acoustic test
	Red	Blue	Green	Purple	Yellow
CRG acoustic box		YES	YES	YES	YES
Conveyer		YES	YES	YES	YES
Multilayer partition wall		YES	YES	YES	YES
CFA acoustic box		YES	YES	YES	
CFA wrapping					YES
CFA Inlet & Outlet mufflers					YES
CFA hard-mounted	YES	YES			
CFA soft-mounted with Armaflex washers			YES		YES
CFA decoupled				YES	
Bags in racks slots	YES	YES	YES	YES	YES
Hatch closed	YES	YES	YES	YES	YES

Table 3 Diagnostic tests performed vs. implemented recovery actions

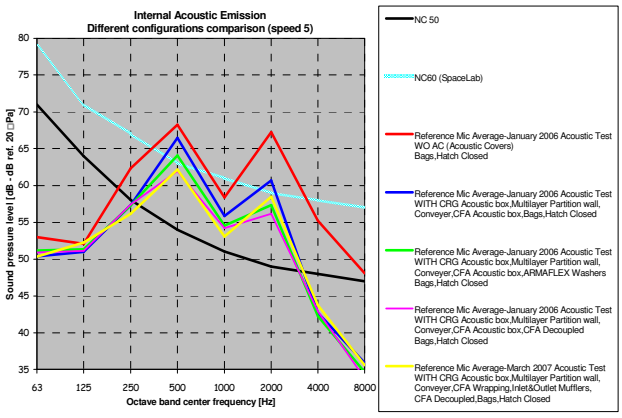


Fig. 20 Effect of the different recovery actions vs NC50 and vs Spacelab performance

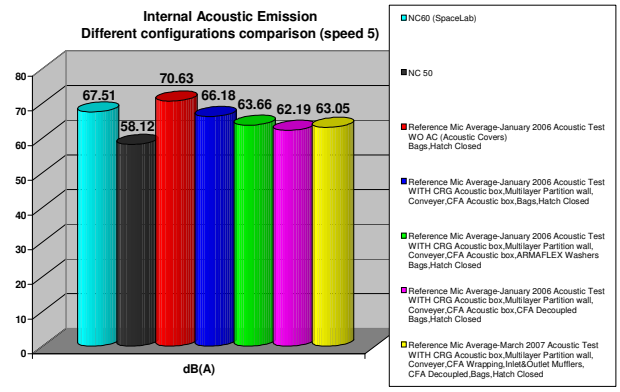


Fig. 21 Comparison between dBA due to the Effect of the different recovery actions vs NC50 dBA and Spacelab ones

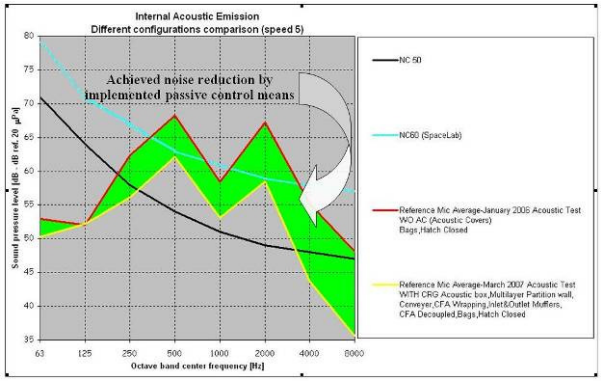


Fig. 22 Achieved noise reduction by implementing the selected acoustic recovery actions. vs NC50 dBA and Spacelab noise criteria curve.

7 Conclusion

These successful industrial experiences shows the suitability of the TAS-I Design Verification process based on the Budget allocation via Mechanical Power and integrated Analysis/Test approach developed to verify and control both the Micro-g and the Audible Noise/Human Vibration Environments requirements.

All the manned modules as MPLM, Columbus and Node 2 & 3, designed by this control strategy are fully compliant to both the Comfort requirement and to the Micro-g ones.

It can therefore be stated that there is a sound chance to fulfil NC50 requirement also on other future modules / projects as (for instance) the Moon and Mars Sites.

An additional advantage of this integrated Design Strategy is that by limiting the mechanical power also the structure-borne and the airborne noise contributions are mutually controlled with obvious benefits on the on-Orbit Vibro-acoustic Environment

(i.e. Audible Noise, Human vibration & Micro-g requirements).

This design methodology, besides the technical effectiveness, has an additional benefit, the financial cost reduction if compared to "classic" design approaches (based on sound pressure level). Experience shows that without any Design Control process from equipment up to system level there is the potential for strong impact on schedule and cost to retrieve the eventual non-compliance.

References

1. Audible noise control in Spacelab as Part of manned spacecraft environmental control – ESA SP-139, November 1978.
2. Low-Gravity Environment in Spacelab Acta Astronautica Vol.9, No.4, pp 187-198, 1982 Pergamon Press Ltd
3. Ullio R., Marucchi-Chierro P.C., "Microgravity (Microdynamic) test facility for disturbance source characterisation through indirect method technique", *Conf.: "Spacecraft structures and mechanical testing"*, Noordwijk, The Netherlands, 27-29 March 1996.
4. Francesconi D., Statistical energy analysis application on pressurised modules: test and analysis of a typical structure", *Conf.: "Spacecraft structures and mechanical testing"*, Noordwijk, The Netherlands, 27-29 March 1996.
5. Marucchi-Chierro P.C., Trittoni L., Marta F., Boggiatto D., "Microgravity and comfort design of pressurized modules under on-orbit vibro-acoustic environment", Conf: IAF 98, Melbourne, Australia, 28 September - 2 October 1998.
6. Marucchi-Chierro P.C., M. Martini "Columbus Verification Control approach versus micro-gravity requirement", *Conf.: "5th ESA/ESTEC "*, Noordwijk, The Netherlands, 15-17 June 2004

ACRONYMS

ASI	Agenzia Spaziale Italiana
ATV	Automated Transfer Vehicle
AVM	Anti Vibration Mount
AVS	Avionics
CFA	Cabin Fan Assy
CRG	Cabin Return Grid
EADS	European Aeronautic Defence and Space Company
ECLS	Environmental Control Life System
ER3	Express Rack Number 3
ESA	European Space Agency
FEM	Finite Element Method
HRF-1/-2	Human Research Facility – Rack 1 / 2
I/F	Interface
ISS	International space Station
MedOps	Medical Operations
MDPS	Meteoroids & Debris Protection System
MPLM	Multi-Purpose Logistics Module
MSG	Microgravity Science Glove Box
NASA	National Aeronautics & Space Administration
QR	Qualification Review
SEA	Statistical Energy Analysis
SPL	Sound Pressure Level
TAS-I	ThalesAleniaSpace-Italy
TCS	Thermal Control System
WFSV	Water Flow Separation Valve
WOOV	Water On Off valve
WPA	Water Pump Assembly

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Design and Analysis of an Inflatable Airlock for Space Applications

E. Carrera, S. Brischetto and T. Salis
Department of Aeronautics and Space Engineering, Politecnico di Torino, Italy

M. Nebiolo and A. Rapisarda
Thales Alenia Space Italia, Torino, Italy

Keywords: *inflatable structures, inflatable airlock, space vehicles, multi-purpose applications.*

Abstract

Inflatable structures are specific applications of membrane or any ultra-low-mass hybrid structures with an extensive use of membranes. They have many applications such as solar arrays, communication systems, human habitats, planetary surface exploration, radar and reflect arrays. The Inflatable Airlock (IAL), here given, has been developed in the framework of the Piemonte regional project STEPS (Sistemi e Tecnologie per l'EsPlorazione Spaziale), and its main application is the extra vehicular activity. Its inflatable structure characteristics could also permit the multi-purpose applications since it can be lower in mass and packaged in smaller volumes. The main IAL functional elements are the deployment mechanism and the inflatable wall. The final configuration uses telescopic beams and floor panels which are the simpler elements from a kinematic point of view. A Finite Element analysis has been performed after a preliminary sizing and the CAD modeling.

1 Introduction

Inflatable manned modules offer the possibility to deploy on-orbit large structures and they have high habitable volumes in spite of a reduced volume at launch due to their high packaging capability. The inflatable wall of these modules is a multifunctional system composed of several layers, each layer has to accomplish a specific function and to satisfy a set of different requirements [1], [2].

The idea of inflatable structures has been considered for many years because their applications could be numerous. Inflatable structures range from hot air balloons and airships to structural air beams and space suits with interesting design challenges. However, when these design challenges are met, inflatables often perform a function that would be difficult, if not impossible, to achieve with stiff metallic or composite structures. Whether the end is to have a module with a working volume larger than the volume in the launch vehicle or a space suit that must be flexible when it is under pressure, inflatable fabrics provide a design which creates structures that satisfy these needs. One of the latest design challenges comes from the need to work, live, and travel in space for long periods. The large

volume needed to live in space is not only used to store the equipment and to sustain life, but also to provide important psychological benefits for the crew [3].

A structure capable of sustaining life in the harsh space environment is a big design challenge, when full integrated structures need to be larger than the volumetric payload capacity of conventional launch vehicles, fabric inflatable modules become an attractive alternative to metallic structures. There are many design challenges involved in the developing of large fabric structures for space applications. The first one is to create the structure itself, which must be able to withstand the extremely high loads caused by both the size and internal pressure. The second challenge is the protection of the structure from the harsh space environment (see for example micro meteorites, extreme thermal changes, radiations and the atomic oxygen). In real operative missions, the first step is the substitution of a rigid classic module with an inflatable one, but the design of this structure could be appropriate only for specific uses. This is the easier way to test the inflatable technology for manned modules in space field keeping the maximum safety. Surface exploration missions in Lunar environment needs a joint airlock to enabling Extra Vehicular Activity (EVA) missions, in this case the main problem is the lunar dust. An airlock is a device which permits the passage of people and objects between a pressure vessel and its surroundings minimizing the change of pressure in the vessel and the loss of air from it. The lock consists of a small chamber with two airtight doors which cannot simultaneously be opened. The idea is an Inflatable Airlock which will be deployed only by gas inflation [1]-[3].

The Inflatable Airlock (IAL) developed in the framework of the Piemonte regional project STEPS (Sistemi e Tecnologie per l'EsPlorazione Spaziale) has the extra vehicular activity as main application. However, it could be extended to many other applications with slight adjustments in order to obtain a multi-purpose structure.

A preliminary design of the IAL has been given in order to have a reference configuration which could be improved. The IAL is attached to a manned module and it must allow the crew egress to the external, moreover it is subjected to pressurization and de-pressurization cycles. The airlock has an inflatable segment which permits the crew egress from the Rover and it must a priori be pressurized. The IAL, here proposed, has a diameter of 1900 mm and a length of 2400 mm. The deployment of the IAL must be completely passive. The main requirements for the whole structure can be classified as: functional and system requirements for the IAL wall, environmental requirements, physical requirements, operational requirements, structural wall material configuration requirements, structural restraint requirements, bladder requirements, internal barrier requirements. These requirements will be discussed in details at the conference and not in this paper for the sake of brevity.

The present work is structured in five main parts. The main possible configurations are discussed in Section 2, the best one is a combination of the second and third configuration and its deployment mechanism will be analyzed in Section 3 via a preliminary design. The inflatable structural restraint and global analyses are performed in Sections 4 and 5, respectively. Finally, Section 6 gives the main conclusions.

2 Evaluated Configurations

The first step of this work is the definition of a preliminary configuration for the inflatable airlock. A reference design is highlighted in order to develop the whole project. The main task of the IAL is to allow the link between a rigid habitation module (for example the Rover in this case) and the external environment. Connections to the exit and re-entry must be guaranteed by the module. The structure must be able to support loads generated by crew passage as well as by several pressurization and de-pressurization cycles. The main purpose is to allow crew operations in an optimal way.

The general configuration of an IAL can mainly be divided in the following two functional elements:

- inflatable wall;
- deployment mechanism.

The inflatable wall is a multilayered structure made of several layers, each one has a specific function. The deployment mechanism is thought as a passive element, it follows the inflatable structure without motors, springs or others third part assistances. The deployment movement is actuated only by the inflation and, therefore, by the differential pressure between the internal volume and the external environment. This structure must accomplish two main tasks: the first one is to bind up the inflatable wall in a right deployment sequence, the second task is to block off it in two particular positions (retracted and completely deployed). The mechanism must provide the appropriate structural integrity for the whole IAL and it allows the operative load mission support. An example for a general configuration of the IAL is given in Fig. 1.

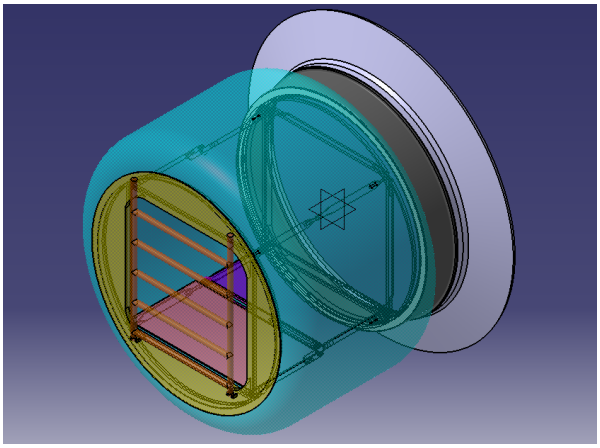


Fig. 1 IAL general configuration

Three main configurations have preliminary been analyzed, they are described in details in the following subsections. The best one appears a combination between the second and the third configuration.

2.1 Configuration 1: deployment mechanism with compass opening arms

This deployment mechanism is made of several parts, each one with different features. As first approximation the structure can be divided into:

Fixed frame, this element is an integral part of the Rover. Therefore, it is fixed while inflatable shell can be deployed and retracted. The main functions are:

- to provide a connection between the Rover and the IAL;
- to provide interfaces for the hatch between the Rover and the IAL.

Mobile frame, this element is similar to the fixed frame, but it moves (along the longitudinal axe of the IAL) from an initial position (retracted configuration) to a final position (deployed configuration). The main functions are:

- to provide interfaces for the hatch between the IAL and the external environment;
- to provide interfaces for the supporting framework (in the deployed configuration the structure is like a cantilever beam, therefore the IAL must be anchored to the Rover in order to avoid an excessive free hand and deflection);

Arms, in the deployment mechanism there are four arms. They are positioned in the IAL cross section and they coincide with the four vertices of a square inscribed in the circle drawn by the frame (fixed or mobile). Each element is made of two segments with a compass opening. In the retracted configuration the first segment is opposite to the second one, during the deployment the angle between the two segments increases from zero up to one hundred eighty degrees in the final position (completely deployed IAL).
The main functions are:

- to allow the mobile frame translation;
- to limit the movement of the mechanism between the two outermost positions.

Crosspiece, this element is positioned in the lower part of the airlock. It is a beam with two slots which allows the integral movement of the crosspiece together the two lower arms. The crosspiece must accomplish the following main tasks:

- to provide the necessary structural integrity of the IAL;
- to provide the interfaces for the panel floor.

The mechanism configuration must have two main characteristics: the small overall dimension in the retracted configuration, and the small frictional resistance force during the deployment. From this point of view the configuration with compass opening arms is an excellent solution. This mechanism provides both the characteristics and its arms provide a reduced deployment dimension and limited friction areas. These are good characteristics for a passive mechanism. Thanks to this opening mechanism the overall dimension is smaller than other configurations. For example, a telescopic arm configuration provides a bigger overall dimension and a bigger resistance, during the movement, than this selected configuration. Moreover, this configuration can use the volume between the two frames to allocate the deployment mechanism.

This first configuration (see Fig. 2) is a good solution to have a retracted configuration with very small overall dimension and, at the same time, a low frictional resistance force against the deployment. However, at this point of the work, the attention is focused on the reliability of the mechanism. For this reason a simpler configuration must be defined in this work. From the structural reliability point of view the weak point of this configuration is the use of the compass opening arms. Each element works in a good way to achieve the previously discussed ends, but they do not ensure the best performance in the phase of deployment of the structure. At this stage, although the structure is subjected to its only weight, the particular configuration of the arms is difficult to be modeled and exhaustively analyzed. For this reason, a simpler and more reliable configuration is studied in this work.

The chosen configurations are two but they have very similar mechanisms:

- Deployment Mechanism with Telescopic Beams (acronym used is DMTB);
- Deployment Mechanism with Telescopic Floor (acronym used is DMTF).

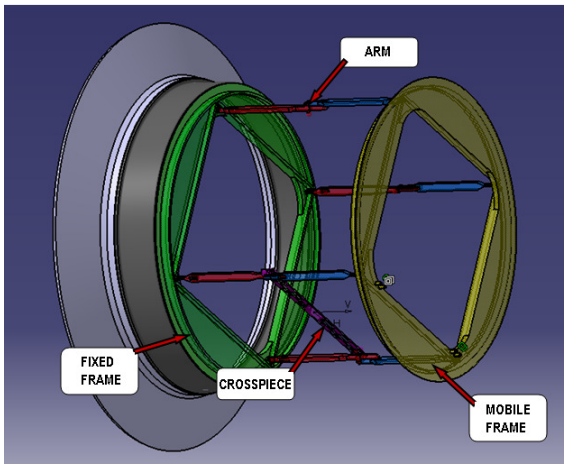


Fig. 2 Configuration 1 of the IAL

2.2 Configuration 2: deployment mechanism with telescopic beams (DMTB)

The deployment mechanism of this configuration (see Fig. 3) is a passive element as in the previous case. The deployment movement is actuated only by the inflation and therefore, by the differential pressure between the internal volume and the external environment. The DMTB is constituted by several parts, each one has different features. As first approximation the DMTB can be divided into:

Fixed frame, this element is very similar to the previous configuration. The main difference is the connection part. However, the main functions are the same:

- to provide connection between the Rover and the IAL;
- to provide interfaces for the hatch between the Rover and the IAL.

Mobile frame, the mobile frame is quite similar to the fixed one, and it moves (along the longitudinal axe of the IAL) from the initial position (retracted configuration) to the final position (deployed configuration). The main functions are:

- to provide interfaces for the hatch between the IAL and the external environment;
- to provide interfaces for the supporting framework (in the deployed configuration the structure is like a

cantilever beam, therefore the IAL must be anchored to the Rover in order to avoid an excessive free end deflection);
- to provide interfaces for secondary elements, such as the contact surface ladder.

Telescopic beam, in the deployment mechanism there are four telescopic beams. This is the core of the mechanism which characterizes this configuration. Each element is made of three cylinders, which slide into each other. The main functions of the telescopic beam are:
- to permit the mobile frame translation;
- to limit the movement of the mechanism between the two outermost positions.

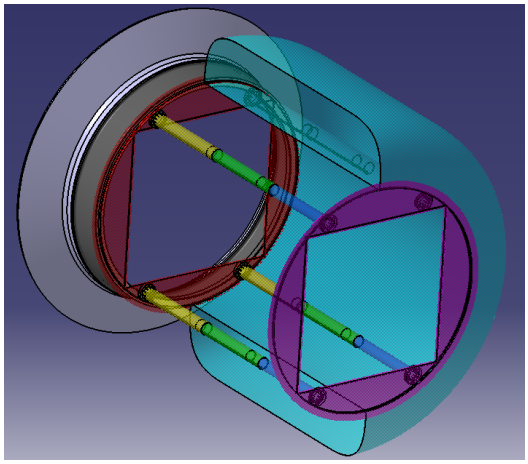


Fig. 3 Configuration 2 of the IAL

2.3 Configuration 3: deployment mechanism with telescopic floor (DMTF)

The telescopic floor employed in this configuration is a passive element too. The two lower telescopic beams are replaced by only one telescopic element which is a panel floor. This is the main difference between this configuration (see Fig. 4) and the second one with four telescopic beams. The mechanism is the same if we consider all the other features. The DMTF does not need any other element (for example the panels of the compass opening arms in the configuration described in the Section 2.1) to provide the passage of crew up to the external hatch. The described deployment mechanism, except for the telescopic floor, is constituted by the same parts of the previous configuration, and they also have the same functions. The working sequence is very similar.

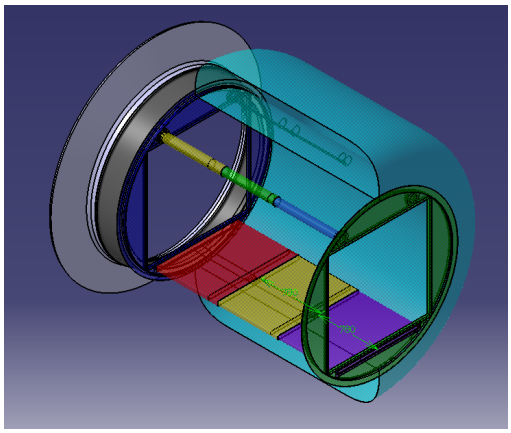


Fig. 4 Configuration 3 of the IAL

Summarizing these aspects:

Fixed frame, this element is very similar to the fixed frame of the previous configuration. The main difference is the lower connection zone. The main functions are exactly the same:
- to provide connection between the Rover and the IAL;
- to provide interfaces for the hatch between the Rover and the IAL.

Mobile frame, it is very similar to the fixed equivalent frame. The main functions are:
- to provide interfaces for the hatch between the IAL and the external environment;
- to provide interfaces for the supporting framework (in the deployed configuration the structure is like a cantilever beam, therefore the IAL must be anchored to the Rover in order to avoid an excessive free end deflection);
- to provide interfaces for secondary elements, such as the contact surface ladder.

Telescopic beam, in this deployment mechanism there are only two telescopic beams which are located on the upper section of the structure. The main functions of the telescopic beams are:
- to permit the mobile frame translation;
- to limit the movement of the mechanism between the two outermost positions.

Telescopic floor, the two lower telescopic beams of the second configuration are replaced by a telescopic floor element. The telescopic floor is made of three box-type elements, with rectangular section, which slide into each other. The telescopic floor functions are exactly the same of the equivalent beams.

The final configuration analyzed in this work is a combination of the DMTB and DMTF.

3 Preliminary Design of the Deployment Mechanism

The structure analyzed in this section is a combination of the DMTB and DMTF configurations, it has four telescopic beams and the telescopic floor positioned over the two lower telescopic beams. A preliminary design of the IAL deployment mechanism is given via a simple analysis. A FE (Finite Element) model of the structure is created after a preliminary evaluation of each element.

The geometrical data are summarized in Tab. 1, the loading data in Tabs. 2 and 3, and the material data in Tab. 4.

	[mm]
Diameter	1900
Length	2400

Tab.1 Overall dimensions

	Mass[Kg]
Crew Member	75
Space Suit	100

Tab.2 Crew and space suit masses

	Gravity acceleration [m/s^2]
On Earth surface g_0	9.81
On Lunar surface g_{Moon}	1.622

Tab.3 Gravity acceleration on the Earth and Moon surface

Material	Al7075 Aluminum Alloy
Elastic Modulus [MPa]	71044
Elastic Modulus Compression [MPa]	73114
Tensile Yielding [MPa]	317
Tensile Ultimate [MPa]	393
Compression Yielding [MPa]	310
Shear Ultimate [MPa]	255
Bearing Yielding [MPa]	524
Bearing Ultimate [MPa]	661
Density [kg/m^3]	2795
Poisson Modulus	0.33

Tab.4 Al7075 Aluminum Alloy properties

The preliminary design is based on an oversimplified structural configuration. From this point of view the IAL structure could be considered as a beam fixed at one edge. This preliminary analysis considers the maximum load applied on the beam free edge as shown in Fig. 5. This case allows a simple preliminary bending design of the structure to be accomplished.

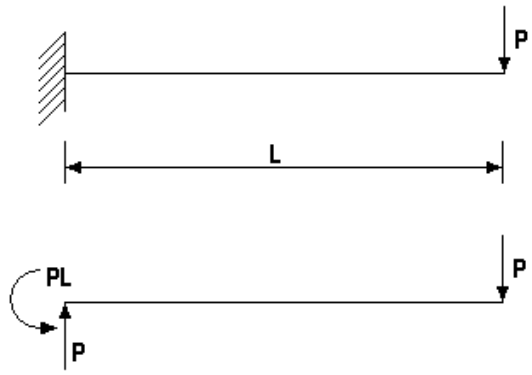


Fig. 5 Oversimplified IAL structure

3.1 Finite element analysis

A Nastran FE Model [4]-[6] is here discussed for the only deployment mechanism. It is based on the configuration given in Fig. 6 where the telescopic floor is added to the two lower telescopic beams. The internal structural core design is based on a cage like structure, made of four telescopic longerons connected to fixed and mobile frames. Each frame is like a bulkhead ring which provides interfaces for all the inflatable shell functional layers [4].

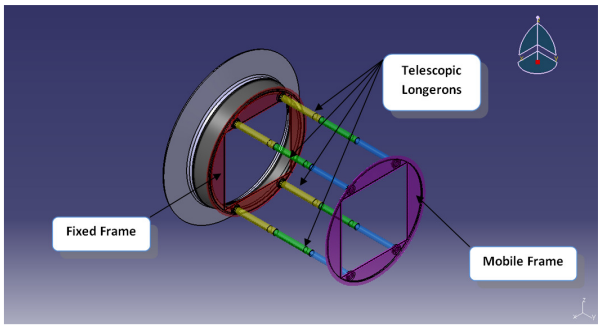


Fig. 6 Chosen deployment mechanism

The transverse displacement due to the force applied on the mobile frame is shown in Fig. 7. The maximum displacement is 0.79 mm. In order to limit the transverse displacement, the structure is stiffened by two cables between each upper and lower telescopic beam. The maximum value of displacement is 3.44 mm in the mobile frame and 2.88 mm in the fixed frame. These values are not the real displacements because of three reasons. The first reason is that the mobile frame constraint interests only few nodes. The second one is that the pressure is applied only inside the frames, in the real case there is also a pressure in the exterior face of the frame and for this reason the gradient is smaller. The third reason is that the employed pressure is four times bigger than the operative pressure of the inflatable airlock.

The Von Mises stresses caused by forces are shown in Fig. 8. The maximum value is located in the beams and it is the same of the cable stress. This stress is much smaller than the tensile yielding value.

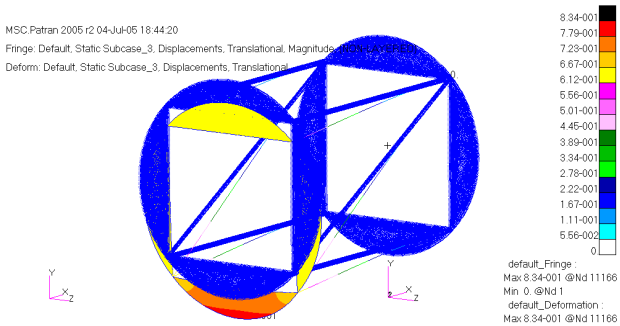


Fig. 7 Deployment mechanism transverse displacements

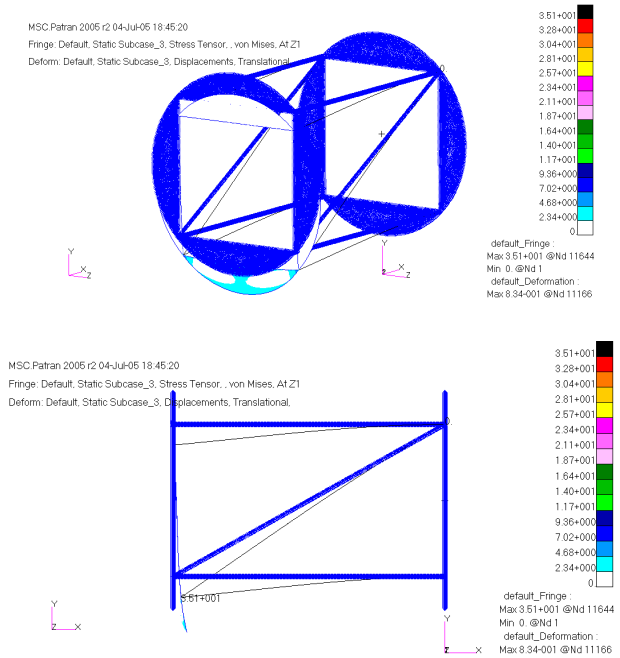


Fig. 8 Deployment mechanism stresses

4 Inflatable Structural Restraint

The structural restraint configuration (see Fig. 9) is described in this section. It is made of a cylindrical barrel combined with two toroidal ending parts. Then, a simplified FE model of the inflatable shell [4], using Nastran solver [5], [6], is given and the load fluxes are evaluated when 1 bar internal pressure is applied. This evaluation is related to the ribbon number required to sustain the applied qualification pressure level in the different areas of the inflatable shell. The minimum detected allowable flux for ribbons is obtained.

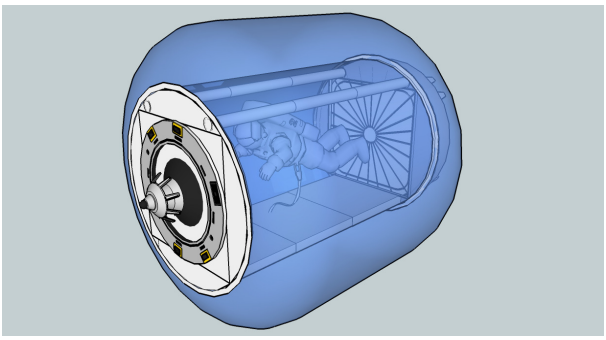


Fig. 9 Inflatable shell shape

The Nastran FE Model (see Fig. 10) is based on a configuration with a uniform shell thickness of 1 mm in order to perform a first check of the load fluxes in the restraint layer. The target is to validate the results obtained with the preliminary simplified analytical solutions. A structural restrain model has been obtained via MSC.Patran pre/post-processing software [5], [6].

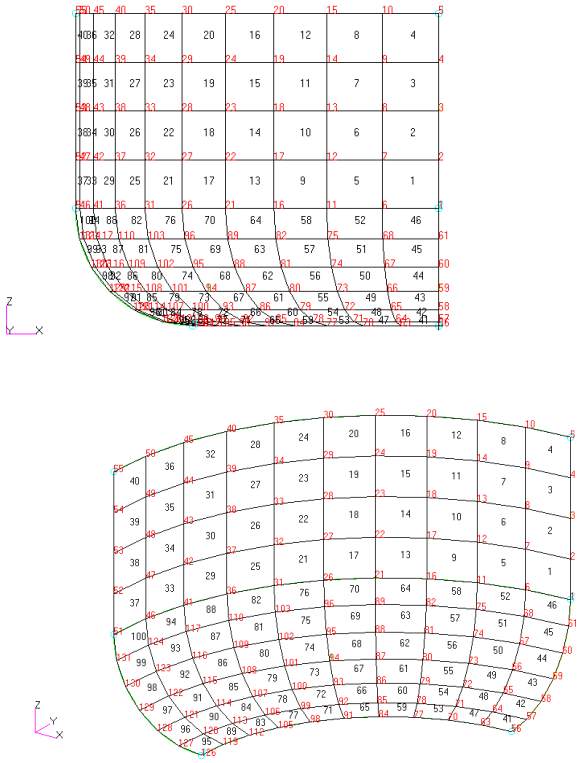


Fig. 10 Lateral and isometric views of the mesh.

The displacement and stress evaluations for the shell are given in Figs. 11 and 12, respectively.

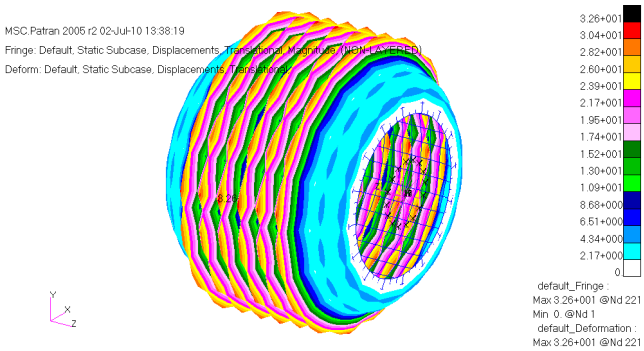


Fig. 11 Overall shell displacements

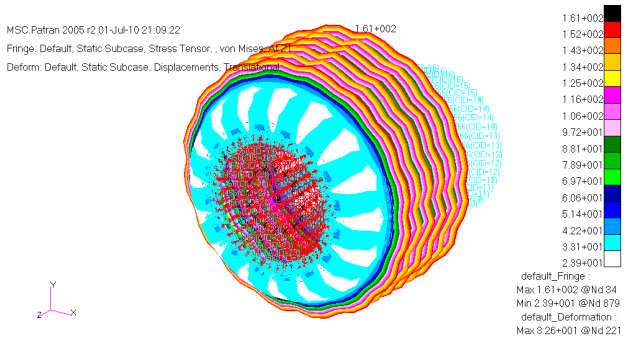


Fig. 12 Overall shell stresses

5 Inflatable Airlock Global Analysis

The FE models previously created and analyzed for both the deployment mechanism and the inflatable structural restraint are here assembled to create a mathematical model for the entire inflatable airlock. The FE model created is analyzed to verify the whole structure [4]-[6].

The mathematical model is built through the assembly of the two previous FE models created. The base model is the deployment mechanism, the restraint FE model is built around the rigid mechanism.

Figure 13 shows the overall structural mathematical model, Figure 14 gives the applied loads and the constraints applied to the structure:

- 1 Force (yellow markers);
- 1 Pressure Load (red markers);
- 28 Nodes with 6 degrees of freedom constrained (cyan markers).

The main mathematical model characteristics are:

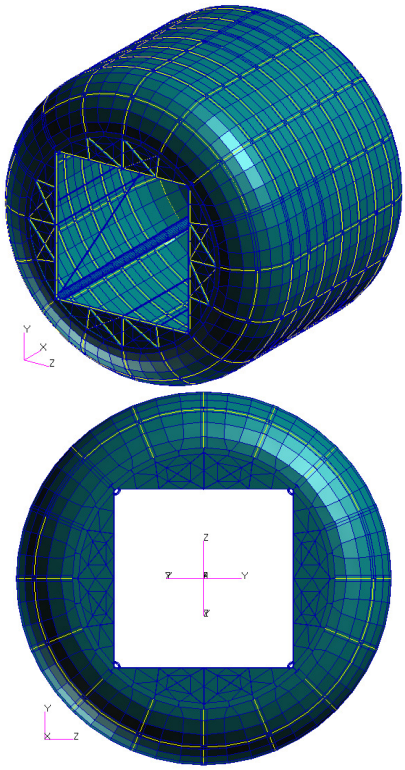


Fig. 13 Inflatable airlock model, isometric and frontal view

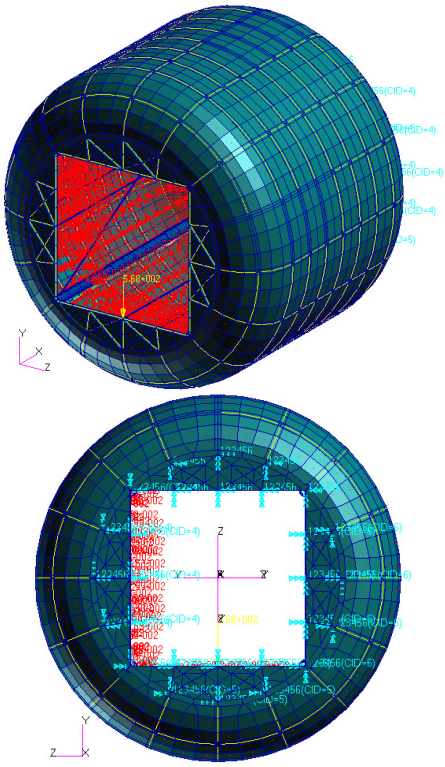


Fig. 14 Inflatable airlock model, applied loads and constraints

- Node Number: 1524;
- Element Number: 2898;
- Element Type: bar2 (stiffened ribs, beams, cages and ribbons), tria3 (frame panels, restraint 2D fabric) and quad4 (restraint 2D basic fabric).

Materials: Al7075 and PBO-Zylon.

- Constraints: 28 constrained nodes for 6 degrees of freedom;
- Load: force and constant pressure applied inside the shell.

The global inflatable airlock displacements and stresses are shown in Figs. 15 and 16, respectively.

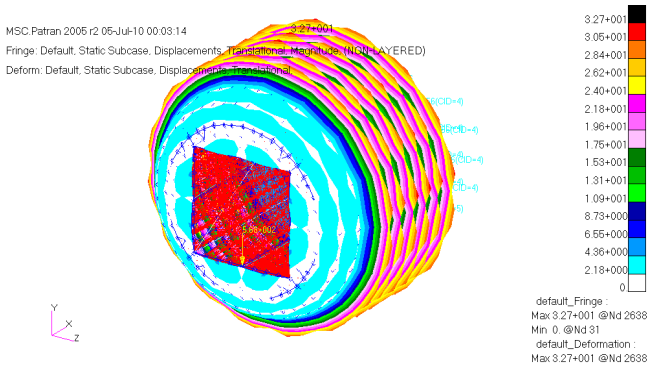


Fig. 15 Global inflatable airlock displacements

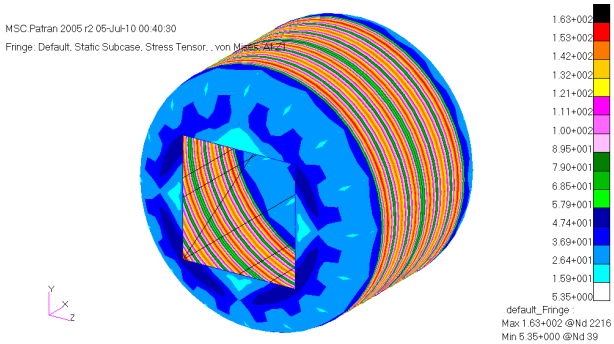


Fig. 16 Global inflatable airlock stresses

6 Conclusions

The final configuration of the deployment mechanism has been obtained after a first

preliminary analysis. This preliminary structure has been analyzed with the MSC.Nastran FE code using qualification pressures and interface loads coming from structural restraint analyses. Results have shown stress levels which give positive margins of safety for all the deployment mechanism metallic parts. The proposed concepts for the structural core are considered as achievable. The analysis performed has provided a detailed evaluation of the material configurations based on a ribbon grid fixed to a below fabric. The results have defined the required number of ribbons in the cylindrical and toroidal sections of the structural wall, based on the applied load fluxes at the qualification pressure level. A detailed FE model analysis made with MSC.Nastran code has shown stress levels which give positive margins of safety for all the structural restraint parts. Finally, a complete Inflatable Airlock FE model has been analyzed, the single structural element behaviors have confirmed the good quality of the project. The Inflatable Airlock is suitable for space applications and for multipurpose ends after opportune modifications.

References

[1] Cadogan D., Stein J. and Grahne M., "Inflatable composite habitat structures for lunar and mars exploration", *Acta Astronautica*, Vol. 44, No. 7-12, 1999, pp. 399-406.

[2] Jenkins C.H.M., *Gossamer spacecraft: membrane and inflatable structures technology for space applications*, Published by American Institute of Aeronautics and Astronautics Inc., Virgina, 2001.

[3] de Jong M. and Lennon A., "Pressure Restraint Design for Inflatable Space Habitats," *10th European Conference on Spacecraft Structures, Materials and Mechanical Testing and 1st CEAS European Air and Space Conference*, Berlin, 2007.

[4] Kraus H., *Thin elastic shells*, John Wiley and Sons Inc., New York, 1967.

[5] MSC Nastran, "Quick reference guide," MSC.Software, 2010.

[6] MSC Patran, "Reference manual," MSC.Software, 2001.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Computational simulation of far-field pyroshocks including experiments and optimization

A. Lacher, N. Jüngel and Utz von Wagner
Technische Universität Berlin, Germany

A. Bäger
Deutsches Zentrum für Luft- und Raumfahrt (DLR), Germany

Keywords: *pyroshock, SRS, wave propagation, impact, plate*

Abstract

Pyroshocks are caused by explosions of pyrotechnic devices attached to spacecraft structures and can lead to failures in electronic subsystems. In the so-called far-field, acceleration levels up to 1.000g within a frequency range of 10kHz are expected. According to existing standards pyroshocks have to be simulated for qualification of space modules which, for the far-field case, can be done by in-plane or out-of-plane hammer excitation of a plate carrying the module. By switching system parameters the acceleration signal measured at the specimen can e.g. be systematically influenced in order to fulfill strict test requirements given in terms of the Shock Response Spectrum (SRS). The paper investigates analytical mechanical models of an existing test setup and allows for numerical prediction of expected acceleration signals. The fast analytical procedure is verified with experiments and further on used for optimization with the help of an evolution strategy as well as numerical trial and error.

1 Introduction

The failures in space structures caused by pyrotechnic devices have been listed in [1] showing that most accidents in aerospace are based on excitation within high frequency ranges. In several standards by governmental organisations, such as [2], or administrations such as [3] and [4] pyroshock tests are defined including guidance for performing pyroshock tests. *Procedure III* from *Method 517* in the Military Standard [2] of the US Department of Defense includes tests that are based on transient mechanical excitation other than by electro- dynamic shakers. Mechanical impacts caused by hammer pendulums are able to excite structural frequencies of up to 10kHz at very high acceleration levels. For facilitating test procedures and synthesizing prescribed spectra several setups have been developed in order to simulate far-field pyroshocks. Examples are [5], [6], [7] and [8] where mechanical excitation by in-plane and out-of-plane impacts are achieved. In various shock handbooks such as [9], [10], [11] and [12] other test facilities for pyroshock simulation are presented including far-field and mid-field tests. In some cases, Finite Element models have been programmed in order to allow for prediction of experiments, see e.g. [13], [14]

and [15] while in [16], [17] and [18] pure analytical procedures with detailed material and wave propagation modeling have been developed. However, there is a lack of fast analytical procedures with high modeling depth and practical applicability allowing for accurate prediction of pyroshock tests. Also, optimization strategies are rarely discussed in this concern.

The present paper includes the mechanical and mathematical modeling of an existing pyroshock test device including verification by experiments. The calculations are mainly analytical and result in very short processing times which allows for fast computational pyroshock test simulation.

From the mechanical point of view, an in-plane and out-of-plane impact of a sphere on a plate is described in the present work with the help of HERTZian contact law. The solution of the mechanical field equations is obtained with the help of the RAYLEIGH-RITZ procedure. Additional masses are attached to the carrying plate which models the influence of the specimen on the wave propagation on the one hand and accounts for a possible control of the structural resonances on the other hand. For the mechanical solution of the problems textbooks such as [19], [20] and [21] are mentioned. At this point, the authors do not go into a detailed literature study about existing theoretical approaches of solving impact problems of solids with subsequent wave propagation. However, one can mention that this field has been investigated within the last two centuries for one dimensional problems such as the impact of a rigid sphere on an elastic rod or beam. Also, the out of plane impact on a plate has been studied even for the case of additional masses attached to it, as can be found in the textbooks. In contrast, the analytical solution of two dimensional in-plane impact of a sphere hitting the edge of an unsupported plate has not been found in the literature, to the best of the authors' knowledge. The mentioned problem is the origin for the modeling of pyroshock test setups from [5] and [6] as shown in Fig.s 1 and 2. In both cases, the impact is directed in-plane and achieved by the hit of a hammer pendulum corresponding to a rigid sphere in the mechanical model. The plate guiding and reflecting the shock wave which is caused by the transient excitation carries the test specimen. The inertia of latter reacts on the plate and influences its structural resonances. Latter effect

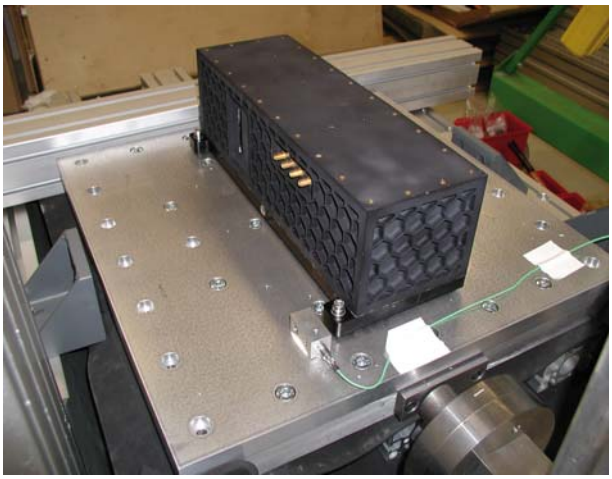


Fig. 1 Pyroshock test setup from [6]

can be specifically utilized in order to influence the expected spectral properties of the measured signal.

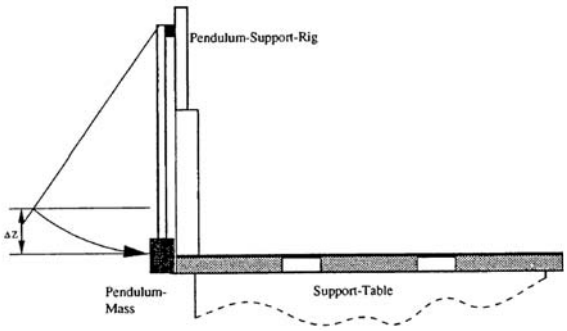


Fig. 2 Pyroshock test setup from [5]

According to the handbooks [9], [10], [11] and [12] the common procedure of evaluating a pyroshock in terms of its frequency and acceleration content is the *Shock Response Spectrum* (SRS). An acceleration signal base point excites an array of N single degree of freedom oscillators with different undamped natural frequencies f_i. The maximum acceleration $\mathbf{max}|\ddot{x}_i(t)|$ of each oscillator resulting from the input acceleration $\ddot{y}(t)$ is defined as one point in the double logarithmic SRS graph. Figure 4 shows the mechanical

model for calculating SRS. The numerical calculation of the SRS data points is done by the widely-spread and fast SMALLWOOD algorithm from [22]. In order to fulfill a pyroshock test, usually reference SRS are given which have to be achieved at certain points on or near the specimen to be tested. An example for a typical far-field SRS including the reference SRS and tolerance lines is given in Fig. 4.

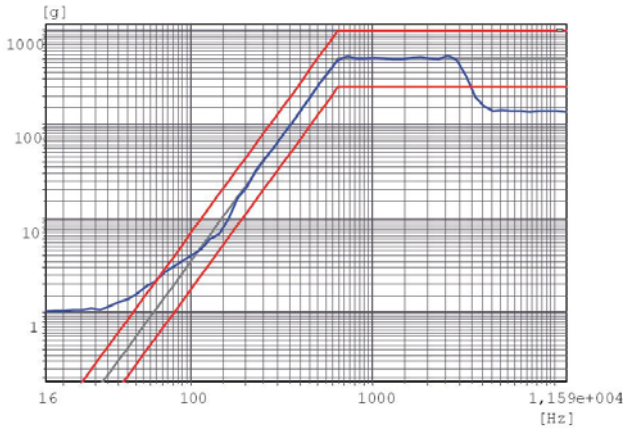


Fig. 3 SRS (blue) with tolerance lines (red)

Examples for further SRS graphs will be shown in Sec. 4 and following. The procedure of finding the appropriate set of system parameters for the pyroshock test in order to meet the requirements is one of the main challenges of the pyroshock engineers. Hence, experimental *trial and error* is often chosen as a strategy which is also combined with long testing durations and high material costs. The following sections include a theoretical mechanical modeling of the mentioned test setups including a numerical solution and verification with experiments.

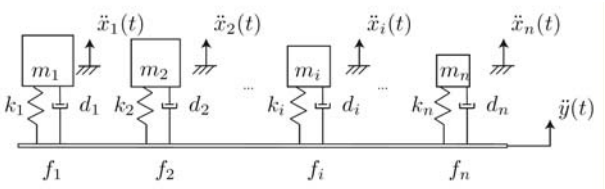


Fig. 4 scheme of the SRS calculation

In addition, the analytical results are evaluated based on the variation of system parameters concerning the striking sphere as well as the position of the test specimen on the plate and of additional masses. The results are visualized and further optimization is made with the help of an *evolution strategy* as well as systematic *trial and error*. With the help of the developed material, a method of finding appropriate system parameters for successful completion of pyroshock testing of aerospace modules (sometimes also called *qualification*) is suggested.

2 Mechanical modeling

This section deals with a brief description of the mechanical models used for calculating acceleration signals at selected points of the carrying plate. Two cases are investigated, which are, on the one hand, the out-of-plane and, on the other, the in-plane impact of a sphere on a free plate. In both cases, the plate is initially at rest and a rigid sphere is hitting a material point. Figures 5 and 6 show the sketches of the corresponding models, where for brevity, no point masses are included. E, E_1 and ν, ν_1 denote for the elastic constants YOUNG's modulus and POISSON's ratio of the plate and the sphere respectively whereas v_0 is the impact velocity of the rigid sphere having the radius of curvature R_1 and mass m_1. The dimensions height, width and thickness are given by a, b and h respectively.

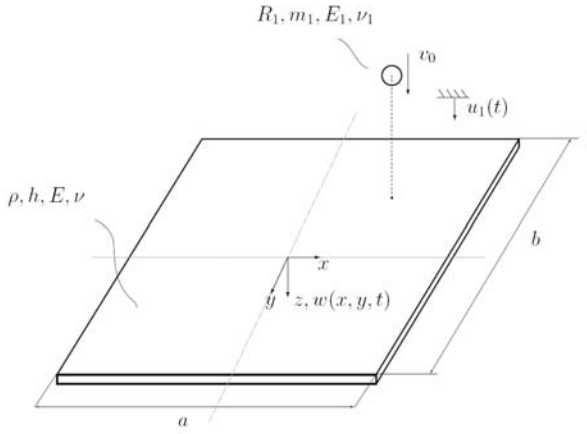


Fig. 5 mechanical model for out-of-plane impact

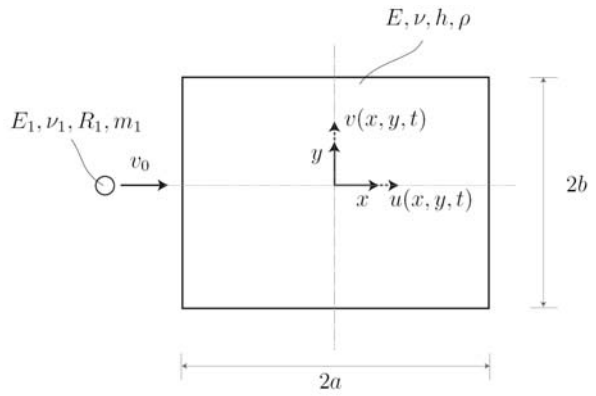


Fig. 6 mechanical model for in-plane impact

The resulting contact problem can be formulated with the help of the RAYLEIGH-RITZ procedure (see e.g. [21]) by using the 2nd order LAGRANGE equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\rho}_i} - \frac{\partial L}{\partial \rho_i} = Q_i \qquad (1)$$

where q_i and Q_i are the generalized coordinates and forces and L is the LAGRANGE functional defined in terms of the potential energy U and the kinetic energy T by

$$L = T - U. \qquad (2)$$

For both investigated cases, the energies of the plate Ω write as

$$T = \frac{1}{2}\int_{\Omega} \dot{u}^2 \rho dV \quad \text{and} \quad U = \int_{\Omega} \sigma \cdot\cdot \varepsilon \, dV \qquad (3)$$

with mass density ρ, CAUCHY stress tensor σ and work conjugate GREEN deformation tensor ε combined by the Rayleigh product second order. The velocity $\dot{u}$ corresponds to either the out-of-plane displacement w or in-plane displacement u and v respectively. The energy T_s of the striking sphere with initial velocity v_0 can be written as

$$T_s = \frac{1}{2}\dot{u}_1^2 m_1 \qquad (4)$$

Both, striking sphere and plate are coupled by HERTZian contact force F(t) during impact, see e.g. [20], which is accounted for in the generalized forces Q_i and writes

$$F(t) = \begin{Bmatrix} k\, d^{3/2} & \text{if d positive} \\ 0 & \text{if d negative} \end{Bmatrix} \qquad (5)$$

in terms of the indentation $d(t) = u_1(t) - u(x_0, t)$ between the colliding bodies. k is a proportionality factor and gives

$$k = \frac{4}{3}\left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu^2}{E}\right)^{-1} \sqrt{R_1} \qquad (6)$$

in the specific case. For solving the problem a RITZ approach is formulated for the unknown displacements $w(x,y,t)$ (out of-plane) and $u(x,y,t)$, $v(x,y,t)$ (in-plane) in the sum form

$$w(x,y,t) = \sum_{i=1}^{N} W_i(x,y)\, \rho_i(t)$$
$$\begin{pmatrix} u(x,y,t) \\ v(x,y,t) \end{pmatrix} = \sum_{i=1}^{N} \begin{pmatrix} U_i(x,y) \\ V_i(x,y) \end{pmatrix} \rho_i(t). \qquad (7)$$

The generalized coordinates of the problem are hence the N unknown time functions ρ_i and the single unknown displacement u_1 of the sphere. By substituting Eq. (2) – (7) into (1) and assuming HOOKE's linear elastic law, see e.g. [19], one obtains an initial value problem for the unknowns ρ_i and u_1 with homogeneous initial conditions except $\dot{u}_1 = v_0$. For $W_i(x,y)$ and $U_i(x,y), V_i(x,y)$ form functions have to be used that fulfill at least the geometrical boundary conditions. For the given case of unsupported structures with missing geometrical boundary values in-plane and out-of-plane eigenforms of free quadratic plates are used, see e.g. [23] and [24]. This procedure shows good convergency of the obtained solutions. Latter is calculated by fast numerical integration with MATHEMATICA.

If one intends to add M point masses $\bar{m}_k$ onto the plate at positions x_k and y_k of the plate the term

$$T_m = \frac{1}{2}\sum_{k=1}^{M} \bar{m}_k \dot{u}^2(x_k, y_k, t) \qquad (8)$$

has to be added to the kinetic energy in Eq. (3). In section 4 analytical results are compared with experiments conducted.

3 Experimental setup

For verifying the analytical results experiments have been performed concerning impact on plates. Therefore a steel ball mounted on cables hits a structural point of the plate and excites in-plane or out-of-plane waves respectively.

In the in-plane case, the plate is supported on steel cylinders allowing for translation after impact (free boundary conditions). In the out-of-plane case the plate is freely suspended on ropes ensuring to the best traction free edges. Fig. 7 shows a list including the properties of the involved structural elements.

The sphere's as well as the plate's velocity are both measured by laser vibrometers at a sampling frequency of 102.4kHz. The resulting acceleration in order to calculate the SRS (up to a bandwidth of 10kHz specific for near-field pyroshocks) is obtained by digitally deriving the velocity signal with respect to time. For the measurements no filters have been used.

Object	Material	ρ	E	ν	Dimensions
Ball	Steel	$7640\frac{kg}{m^3}$	210.00GPa	0.30	$R_1 = 0.05$m
Rod	Aluminum	$2710\frac{kg}{m^3}$	70.60GPa	0.33	$l = 1$m, $A = 3.14 \cdot 10^{-4}$m^2
Plate	Aluminum	$2650\frac{kg}{m^3}$	71.10GPa	0.33	$a = b = 1$m, $h = 0.03$m

Fig. 7 list of material and geometrical properties of involved structures

In Fig. 8 and 9 the experimental setups are shown in a schematic.

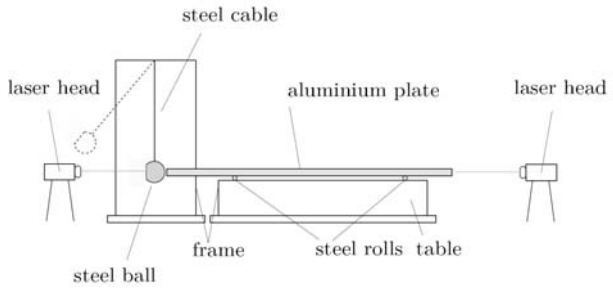


Fig. 8 schematic of in-plane test

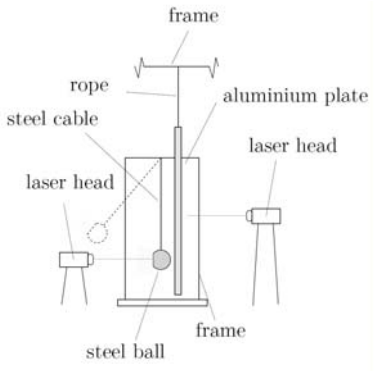


Fig. 9 schematic of out-of-plane test

4 Results

Based on the analytical and experimental investigations a comparison between themis presented in the present section. Therefore, measured (red dashed line) and calculated (blue line) acceleration signals due to impact are shown in Fig. 10 and 12. The corresponding SRS calculated with the SMALLWOOD algorithm are visualized in Fig.s 11 and 13.

In the out-of-plane case the sphere hit the plate at position x = 0.0m and y = -0.4m and at impact velocity of 0.292 m/s whereas the resulting velocity has been measured at the center of the plate.

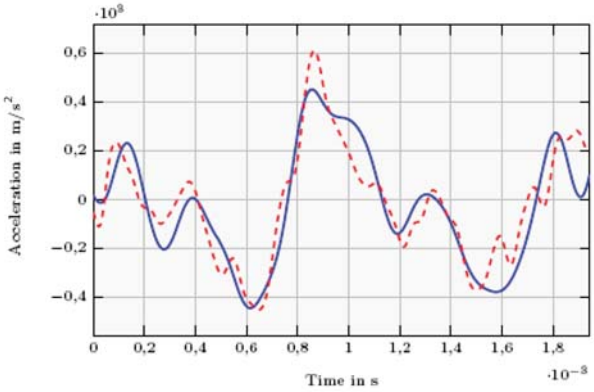


Fig. 10 comparison of out-of-plane acceleration

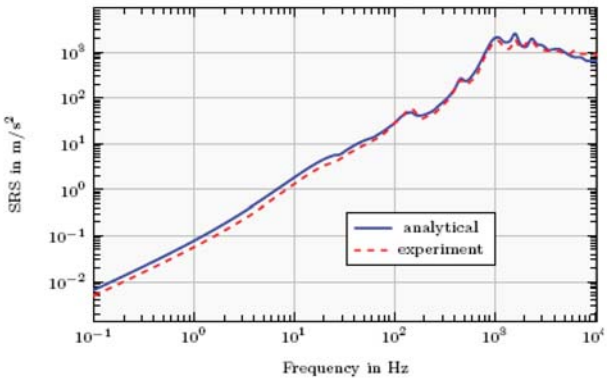


Fig. 11 comparison of out-of-plane SRS

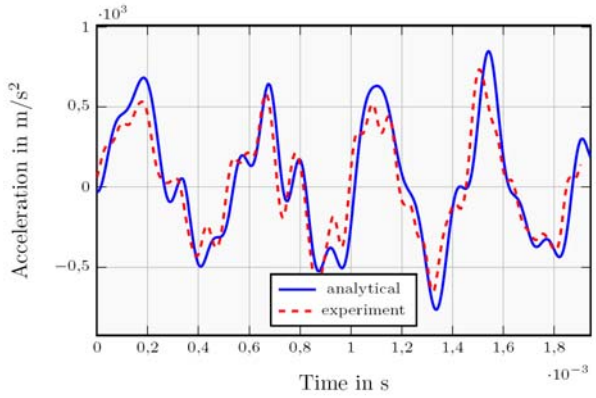


Fig. 12 comparison of in-plane acceleration

In the in-plane case the plate was hit by the sphere in the center of the left edge as seen in Fig. 6. The impact velocity of 0.774 m/s has been measured at the back side of the sphere, where the plate velocity was measured at the opposite side of the impact. The given plots do not include measurements with additional masses. These have been successfully conducted and presented in [25] delivering results of similar quality.

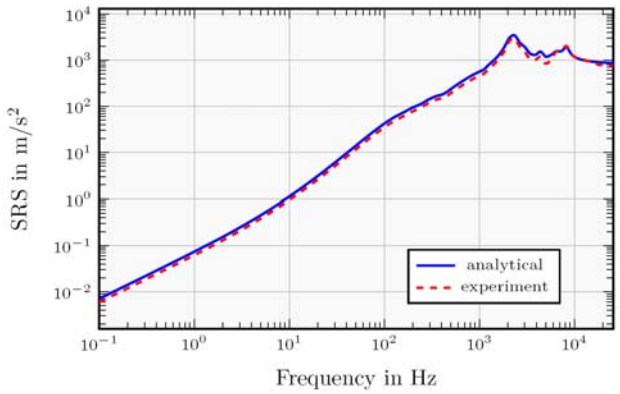


Fig. 13 comparison of in-plane SRS

Based on the presented data one can conclude that the analytical models represent the test setups well and can be used for further analytical investigations.

5 Parameter studies

In order to make predictions of pyroshock tests of the presented test setups possible the authors present results of parameter studies obtained with the help of the analytical models from section 2. Therefore selected system parameters such as impact velocity, radius of curvature, material and mass of the sphere are varied individually while all other parameters are kept constant. The default configuration of the system parameters will be further on called *basic configuration* (and is different from Fig. 10 to 13). The same procedure is conducted with the specimen's mounting position as well as location and position of additional masses. The procedure is similar to a partial derivative and reveals qualitative tendencies of SRS characteristics due to the change of system parameters. For brevity the authors present selected results which are shown in Fig. 14 to 17.

In the out-of-plane case the basic configuration contains the geometrical values of Fig. 7 with a sphere's impact velocity of 0.5m/s, the impact location at x = 0.0m, y = -0.4m and specimen's mounting position at x = 0.0m, y = 0.0m. In the examples shown no additional mass is attached. The basic configuration for the in-plane case is also represented by the values of Fig. 7 as well as an impact velocity of 1.0m/s. The impact takes place at x = -0.5m and y = 0.0m whereas the specimens mounting point is assumed to be at x = 0.5m and y = 0.0m, the same location at which an additional mass is attached in the model. In Fig. 14 to 17 the black solid lines represent the basic configuration which allows for qualitative conclusions concerning the influence of the parameter assigned in the corresponding caption.

Fig. 14 shows, for the out-of-plane case, the influence of the SRS in terms of increasing impact velocity resulting in an upward shift of

the curve whereas an increasing impact mass leads to increasing SRS values in the lower frequency range, see Fig. 15. For the in-plane case Fig. 16 visualizes the effect of an additional mass attached to the carrying plate. The frequency knee shifts to the left with increasing mass which is, from the physical point of view, plausible. Finally, in Fig. 17 the influence on an SRS curve by a change of the specimen's mounting point can be seen.

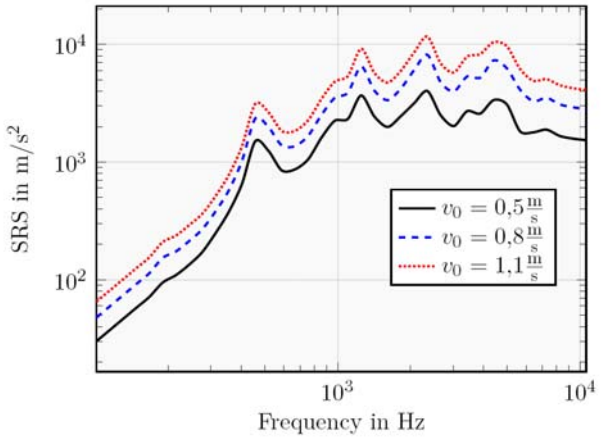


Fig. 14 influence of SRS by variation of impact velocity, out-of-plane case

Similar to the presented plots, all mentioned variations are predictable within the chosen analytical model which finally leads to the authors' idea to implement the model within an optimization strategy or, at least, to systematically influence the SRS using the qualitative results from this section.

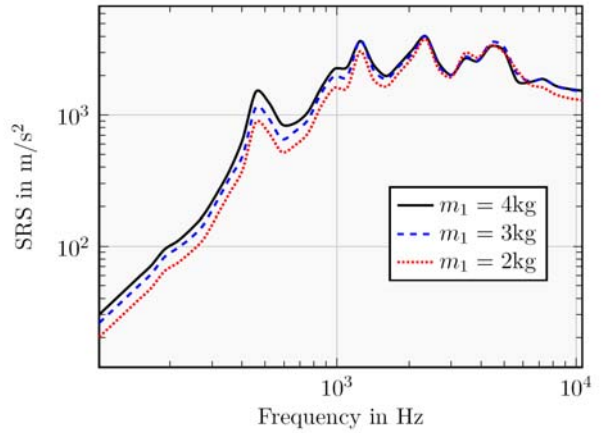


Fig. 15 influence of SRS by variation of sphere's mass, out-of-plane case

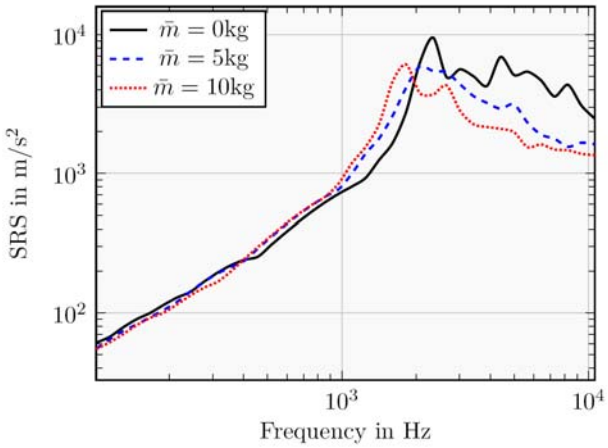


Fig. 16 influence of SRS by variation of additional mass, in-plane case

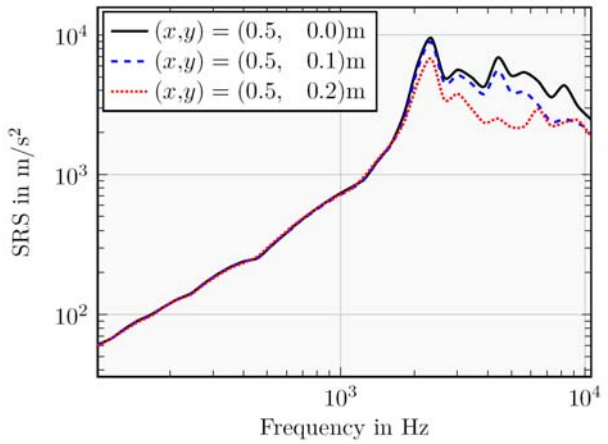


Fig. 17 influence of SRS by variation of specimens mounting point, in-plane case

6 Optimization

This section presents results from an evolution strategy which allows for finding the optimum system configuration leading to SRS that meet specific test requirements. Finally, the authors show an example of achieving the same goal by analytical trial and error.

6.1 Evolution Strategy

The evolution strategy according to RECHENBERG [26] is based on the formulation of a target function Q(r) which has to be optimized in terms of the set of N variables r. In the present case the authors define

$$Q(r) = \sum_{i=1}^{M} \frac{1}{M} \left(\frac{SRS_i(r) - SRS_{ref,i}}{SRS_{ref,i}} \right)^2 \qquad (9)$$

where $SRS_i(r)$ denotes for the ith of M data points of the SRS corresponding to the set r of system parameters. The value $SRS_{ref,i}$ is the "desired" reference data point required by the test specifications. This SRS will be further on denoted by *Reference SRS*. In order to find a minimum error, the following algorithm is used

$$\begin{aligned} r_A^i &= r_P^i + \Delta z^i \\ r_P^{i+1} &= \begin{Bmatrix} r_A^i & \text{if} \quad Q(r_A^i) \le Q(r_P^i) \\ r_P^i & \text{else} \end{Bmatrix} \end{aligned} \qquad (10)$$

where, based on a arbitrary starting configuration, an *descendant* configuration r_A^i is found with the help of the random number z^i within a radius Δ around the *parent* configuration r_P^i. The descendant configuration becomes the parent configuration of the next generation i+1 if it leads to a better quality $Q(r_A^i)$. Based on an arbitrarily chosen starting configuration (start SRS) for the in-plane pyroshock test the authors successfully generated an optimum system configuration that meets the requirements of the test specification. Here, Fig. 18 shows the reference SRS (red solid line), the tolerance lines (red dashed line) as well as the start SRS (blue solid line) which obviously does not meet the test conditions. The quality for the start configuration is Q = 0.406. 857 mutation steps later using the presented strategy the result shown in Fig. 19 has been obtained giving Q = 0.053 and leading to promising results. Basically, the evolution strategy can also be pursued experimentally. However, by analytically finding optimum configurations material costs are minimized.

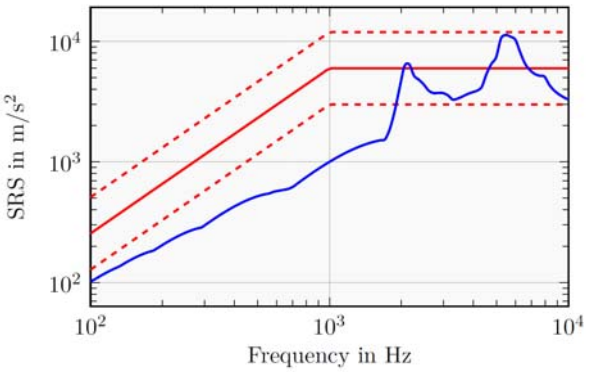


Fig. 18 Example for evolution strategy: *start SRS* (blue), reference SRS (red solid) and tolerance lines (red dashed)

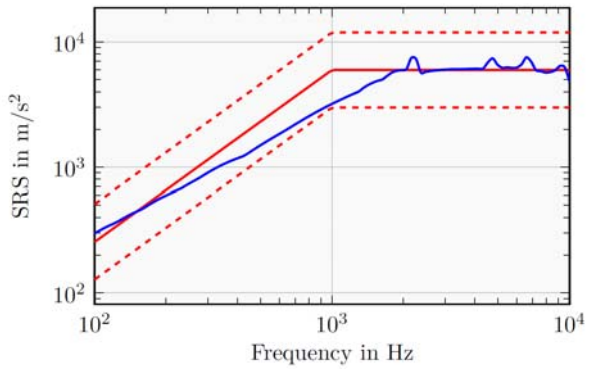


Fig. 19 optimum SRS found with evolution strategy

6.2 Trial and Error

As an alternative utilizing the qualitative results from the parameter studies in Sec. 5 the authors present the results from analytical trial and error procedures representing, in a purely computational way, the SRS synthesizing technique applied by many pyroshock test engineers. In order to allow for a comparison with the results from the evolution strategy the start SRS is chosen to be identical with Fig. 18 from the previous subsection. By successively changing only one system parameter in small increments the authors achieved the SRS from Fig. 20 within 30 *mutations* (compared to 857 in Sec. 6.1). The specific modifications included *[1.] increasing v_0, [2.] decreasing R_1, [3.] increasing m_1, [4.] changing the mounting point and [5.] decreasing v_0*. Hence, on can argue that, at least for the specific case, the qualitative knowledge about how to influence the SRS by changing system parameters can be helpful for

systematic analytical trial and error. Compared to experimental trial and error, the presented computational procedures promise a high potential in order to increase the efficiency of pyroshock tests performed by the initially given test setups.

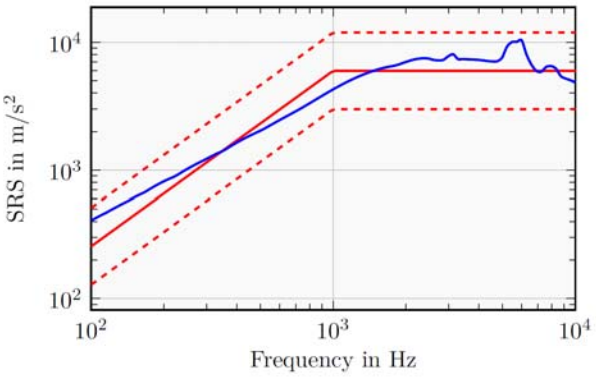


Fig. 20 optimum SRS found with analytical trial and error

7 Conclusion

After giving an overview on existing works dealing with pyroshock simulation and pointing out the lack of analytical procedures for reliable prediction of test results an analytical procedure has been presented including a high depth in mechanical modeling with the help of the HERTZian contact theory and the RAYLEIGH-RITZ procedure. The comparably simple setups, which are the origin for the analytical considerations in this work, allow for a fast computation and can be implemented into common optimization strategies which has been exemplary done with an evolution strategy as well as with analytical trial and error. The analytical results have been compared to those obtained from experiments conducted. The results are promising and allow for further investigation concerning tests with real modules as well as including effects of structural damping and vibrations caused by the internal structural elements of the test specimen.

References

[1] Moening C. J., "Pyrotechnic Shock Flight Failures," *ATM, Institute of Environmental Sciences*, 1985.

[2] N.N., *MIL STD 810 - Test Method Standard for Environmental Engineering Considerations and Laboratory Tests*, US Department of Defense, 2000.

[3] N.N., *NASA STD 7003 Pyroshock Test Criteria-NASA Technical Standard*, National Aeronautics and Space Administration, USA, 1999.

[4] N.N., *IEST-RP-DTE032.2 Pyroshock Testing Techniques*, Institute of Environmental Sciences and Technology, 2009.

[5] Schweickert, G., "The Dornier Shocktable – A new Facility for Shocktesting of Components." *Proceedings of the Third International Symposium on Environmental Testing for Space Programmes, ESTEC,* 1997.

[6] Bäger A., "Pyroshock Entwicklung eines Teststandes", Gesellschaft für Umweltsimulation (GUS), 2007.

[7] Dwyer T. J. and Moul D.S., "Pyro Shock Simulation: Experience with the MIPS Simulator", *Proceedings of the 15th Space Simulation Conference,* 1988.

[8] Smith, J.L., "Effects of Variables Upon Pyrotechnically Induced Shock Spectra, Technical Paper 2603", NASA, 1986.

[9] N.N., *Dynamic Environmental Criteria - Technical Handbook*, NASA, 2001.

[10] Lalanne C., *Vibration and Shock Handbook*, Taylor & Francis Group, 2005.

[11] Davie N. and Bateman V., *Shock and Vibration Handbook*, McGraw-Hill, 1995.

[12] Bernaudin J.B. et al., *Shock Handbook*, ESA Workshop, 2008.

[13] Sutra M., Mesnier D., Berlioz A. and Combes B., "Development of a simulation process of the behaviour of space equipment subjected to pyrotechnic shocks, based on characterization and environmental tests", *Proceedings of the European Conference on Spacecraft Structures, Materials and Mechanical Testing*, 2005.

[14] Wattiaux D., Verlinden O. and Conti C., "Prediction of the Vibration Levels Generated by Pyrotechnic Shocks Using an Approach by Equivalent Mechnanical Shock", *Journal of Vibration and Acoustics*, Vol. 130, No. 4, 2008, pp. 041012-1 - 041012-11.

[15] Bodin E., Brévart B., Wagstaff P. and G. Borello, "Pyrotechnic shock response predictions combining statistical energy analysis and local random phase reconstruction", *Journal of the Acoustical Society of America*, Vol. 112, No. 1, 2002, pp. 156-163.

[16] Ruotolo R., "A spectral element for laminated composite beams: theory and application to

pyroshock analysis", *Journal of Sound and Vibration.* Vol. 270, No. 1-2, 2004, pp. 149-169.
[17] Gherlone M., Lomario D., Mattone M. and Ruotolo R., "Application of wave propagation to pyroshock analysis", *Shock and Vibration*, Vol. 11, No. 3-4, 2004, pp. 145-156.
[18] Hampton R.D., Nygren K.P. and Li T.H., "Analytical shock response of a transversely point-loaded linear rectangular plate", *Proceedings of ESDA 2006 8th Biennial Conference on Engineering Systems and Design Analysis*, 2006.
[19] Goldsmith W., *Impact - The Theory and Physical Behaviour of Colliding Solids*, Dover Publications, 1960.
[20] Johnson K.L., *Contact Mechanics*, Cambridge University Press, 1985.
[21] Wauer J., *Kontinuumsschwingungen*, Vieweg + Teubner, 2008.
[22] Smallwood D.O., "An Improved Recursive Formula for Calculating Shock Response Spectra", *Shock and Vibration Bulletin*, Vol. 51, No. 2, 1981, pp. 211-217.
[23] Iguchi S., "Die Eigenschwingungen und Klangfiguren der vierseitig freien rechteckigen Platte", *Archive of Applied Mechanics*, Vol. 21, No. 5, 1953, pp. 303-322.
[24] Wauer J., "In-plane vibrations of rectangular plates under inhomogeneous edge load", *Eur. J. Mech., A/Solids*, Vol. 11, No. 1992, pp. 91-106.
[25] Renning M. „Messtechnische Untersuchung des Einflusses von Sperrmassen auf die Stoßwellenausbreitung in einer Scheibe als Anwendung für die Pyroschocksimulation", Diploma Thesis, TU Berlin, 2011.
[26] Rechenberg I., *Evolutionsstrategie `94*, Frommann Holzboogs, 1994.

CEAS 2011

Spacecraft Formation Flying and Attitude Control

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Fusion of inter-satellite Ranging and Carrier-Phase Differential GPS for autonomous relative navigation in LEO

Alfredo Renga and Michele Grassi
Department of Aerospace Engineering, University of Naples "Federico II", Italy

Urbano Tancredi
Department for Technologies, University of Naples "Parthenope", Italy

Keywords: *Formation Flying, Relative Navigation, Differential GPS, Inter-satellite Ranging, Augmented Filter*

Abstract

This paper focuses on the autonomous relative navigation of LEO satellite formations. Specifically, a CDGPS technique is addressed using a cascade-combination of dynamic and kinematic filters which processes double differenced code and carrier measurements on two frequencies to get centimeter-level baseline estimates. In order to improve navigation filter robustness against poor GPS geometry and noisy measurements, the CDGPS filter is augmented by inter-satellite local ranging measurements with decimeter-level range accuracy, as the ones provided by ranging transponders or GNSS-like systems. Performance of the augmented filter is tested by numerical simulations of orbital configurations characterized by short, varying baselines, relevant to next generation Synthetic Aperture Radar missions. Numerical results show that the augmented filter is able to preserve the centimeter-level accuracy of the relative navigation also in the cases in which the CDGPS-only filter exhibits a degraded accuracy due to a poor GDOP or a limited number of GPS satellites in view.

1 Introduction

Many present and future space missions for Earth remote sensing, observation of the universe, and Earth gravity field mapping rely on multiple platforms flying in formation. In some cases, the platforms act coordinately to synthesize a unique payload; in other cases, the platforms complement to each other so to magnify mission scientific goals. Formation flying offers many performance and operational advantages with respect to traditional space systems, but at the same time the maintenance of prefixed relative orbital paths among the various platforms is crucial to achieve specific mission goals. This requires implementing a robust communication interlink between the various platforms, and performing autonomous relative navigation and control.

Many applications require the determination of the satellite separations with centimeter-level accuracy to control the satellite relative positions, especially when they come closer as a result of the relative orbital path, or at the sub-centimeter level to achieve scientific goals [1]. For determining the relative position, advanced technologies based on GPS, vision and laser systems have been proposed, and a number of

missions are being performed for in flight demonstration of these technologies [2].

With reference to LEO formations, Carrier-phase Differential GPS (CDGPS) is a promising technology for relative navigation. Previous studies show that processing Single (SD) or Double Difference (DD) carrier-phase measurements within dynamic filters allows determining the satellite relative positions with high accuracy. Sub-centimeter level accuracy has been demonstrated for satellite formations characterized by separations from a few kilometers to hundreds of kilometers [3]-[6], although for large separations only in post-processing.

A major technical problem in implementing CDGPS is solving for the unknown integer number of cycles (integer ambiguity). Moreover, accuracy and robustness of the integer solution has to be preserved against the varying number of common-in-view satellites, and ephemeris and ionospheric differential errors, especially when separations are highly variable, with the satellites continuously coming closer and departing from each other because of the relative orbital path. This is the case of remote sensing applications in which several orbits (parallel, pendulum, helix) are typically proposed to reflect next generation monostatic/bistatic spaceborne Synthetic Aperture Radar (SAR) mission needs in LEO [1],[7].

In addition, navigation filter performance and robustness can be degraded in case of poor GPS geometry, and when a reduced number of GPS measurements is available or anomalous data are gathered. Finally, it is well known that GPS-only measurements commonly attain poor radial accuracy.

To overcome the limitations of a GPS-only navigation system, augmented CDGPS technologies can be implemented [8]-[11]. Specifically, local inter-satellite ranging signals can be exploited to improve accuracy and robustness of CDGPS-based relative navigation. In particular, in [9] it is demonstrated that for baselines of a few kilometers local ranging measurements can be properly combined with GPS measurements to provide a sufficient number of measurements and adequate geometry to solve for the relative position and velocity states with high accuracy. In [11] the augmentation of CDGPS with direct ranging measurements is suggested as a way to enable more rapid and stable converge of the navigation filter solution, especially when the separation among the satellites increases as a result of the relative orbit.

The local ranging measurements can be provided by an Inter-Satellite RF (Radio Frequency) or optical link [8]. In particular, the RF inter-link can be used also for large separations and satisfies the requirement of data distribution among the satellites of the formation. Indeed, GPS data have to be exchanged between the various satellites to form the differential GPS observables needed for CDGPS relative navigation. The RF link can be implemented by means of ranging transponders [8],[11]-[13] or GNSS-like signal based systems [8]-[9], as the one experimented in the PRISMA mission [10].

The availability of local ranging measurements not only improves the accuracy and robustness of a GPS-only relative navigation system, but also improves the accuracy in the radial component determination. Indeed, as shown in [14], an inter-satellite ranging system measuring the along-track baseline variations over the time can be in principle used to evaluate the differences in semi-major axis between the satellites. Any difference in semi-major axis is related to the radial separation between the satellites [15].

In a previous work [16] the authors presented a CDGPS-based filtering approach for precise relative navigation of LEO formations characterized by a computational load compatible with real-time on-board operation requirements. The approach relies on using an Extended Kalman Filter (EKF) combined in cascade with a kinematic filter. The dynamic filter outputs accurate estimates of the floating values of the DD integer ambiguities, whose integer values are extracted by the LAMBDA method [17]. They are then processed in the kinematic filter to obtain precise relative position estimates. Centimeter level accuracy

has been demonstrated in [16] even though a degradation of the accuracy has been observed under worst-case conditions (poor GPS geometry, reduced number of satellites in common view, large separations), especially in the radial direction.

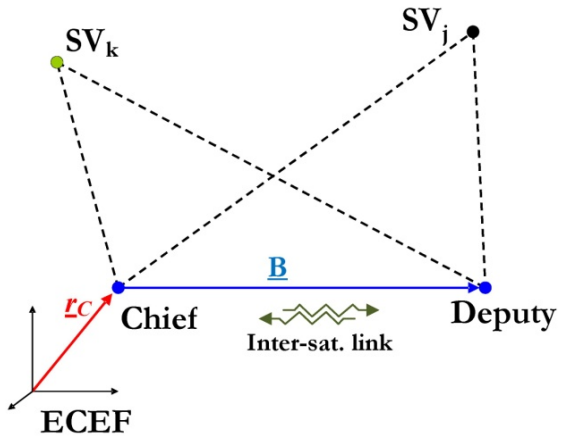


Fig.1Illustration of relative observation geometry with GPS satellites (SV) and ISL.

To overcome these limitations, in the present paper, the proposed filtering approach is generalized by introducing Inter-Satellite Link (ISL) ranging measurements in combination with the GPS ones (Fig. 1). The benefits of fusing ISL and GPS measurements are assessed by comparing the filter performance when processing both observables to the standard GPS-only filter proposed in [16]. The filter performances are ascertained by means of numerical simulations of a formation of two satellites, a chief and a deputy, flying according to different orbital configurations characterized by varying inter-satellite distances. Worst case conditions in terms of GPS satellite observation geometry and ionosphere activity are accounted for.

2 Inter-satellite ranging

Different technologies can be used for inter-satellite ranging. Those technologies can be classified in RF, optical or laser and laser interferometry. In the considered scenario, however, optical or laser metrology cannot achieve satisfactory performance being typically characterized by limited range and/or limited field of view. RF metrology based on ISL relies on local transmission/reception of ranging signals without using the Global Navigation Satellite System (GNSS) constellations. In this context different technologies can be potentially exploited:

- GNSS-like signals
- Transponders
- Ultra-Wide Band (UWB) signals
- Wireless based

The technology based on GNSS-like signal is undoubtedly the most mature, taking advantages of well-assessed experiences on GNSS hardware and software operational issues. These systems rely, in principle, on one-way ranging signals whose structure can be the same as conventional GPS C/A transmissions [8]. This technology is therefore influenced by classic GPS error sources and uncertainties: the ranging measurement is actually a pseudo range, including also transmitter and receiver clock biases and signal propagation delays. Carrier-phase observables can be also derived and used for ranging but only after ambiguity resolution and removal.

A ranging system, based on GNSS-like signals, was embarked, tested and validated on PRISMA mission [9]. The system, whose operational range is a few tens of km, is actually characterized by a dual-frequency S-band configuration, which allows performing carrier ambiguity resolution by wide-lane techniques. In addition, each transmitter is also able to receive GNSS-like signals generated by the other satellites of the formation. In this way relative clock drifts between the platforms can be monitored. Finally, a multi-antenna configuration allows precise line-of-sight (LOS) determination by properly processing multiple carrier phase observables. The overall system mass is about 10 kg with 30 W power consumption. The system demonstrated 1 m ranging accuracy by code acquisition, and 1 cm accuracy by processing carrier phase observables after ambiguity resolution. Similar approaches can be found also in [19]-[20]. Specifically, a GNSS-like signals based system was proposed for RF metrology in Ka-band.

Again, for such a system clock biases must be calculated and carrier-phase measurements have to be used for accurate LOS determination by multiple antenna systems.

As far as transponder-based technology is concerned, a two satellite scenario can be taken as a reference. The inter-satellite distance is determined by the two-way path delay of a signal transmitted by the first satellite, received by the second and then transmitted back by the second satellite, usually using a different carrier, after a default, short, time span. This two-way round-trip ranging technique eliminates the need for high accurate clock synchronization. The technique however is still influenced by propagation delays depending on the selected carrier frequency and the considered inter-satellite distance. The performance achievable by available transponders can greatly vary depending on the transponder technology and adopted transmission protocol. An example of such a system is the Tracking and Data Relay System (TDRSS) developed by NASAs and including a S-band ranging transponder [21]. In [8] the same concept is applied to LEO formations and a system design is presented able to guarantee 30 cm ranging accuracy up to 1000 km baselines with reduced mass (4 kg) and power consumption (about 10 W). If an orbital propagator is used also line-of-sight (LOS) estimations can be performed. A different approach was developed in [11] where a transponder is presented based on long code sequences (P-code like) modulated on Ku-band signals able to achieve 1 cm ranging accuracy, thus avoiding ambiguity resolution issues. The operational range of such a system is limited to a few km. Likewise, a transponder-based system for ranging measurement is proposed in [22] able to guarantee 5 cm ranging accuracy by code only measurements. Specifically, short code sequences are used for signal tracking and locking whereas long-code sequences guarantee high accuracy ranging measurements. For completeness it is important to point out that even though range measurements derived by transponders are not affected by clock biases, the synchronization of those data to a common mission time reference (e.g. GPS time) could also involve managing clock biases.

UWB ranging techniques involve the transmission of short electromagnetic pulses via a current radiating antenna. These short pulse width signals (less than 1 ns pulse width) have very large signal bandwidths (1-2 GHz). In contrast to most continuous wave ranging techniques, ultra-wideband signals have no carrier and are inherently spread spectrum, making them more immune to multipath errors due to interference from a reflected carrier wave signal. In [23], an UWB ranging system is developed and tested in anechoic chamber. The experimentation showed the possibility to guarantee sub-centimeter ranging accuracy also eliminating the measurement bias ambiguity problem. However, important limitations still remain which prevent the implementation of such a technology in satellite formation flying applications, including the design of UWB antennas and spectrum allocation and pollution.

Among wireless protocols the most promising is the 802.16e, able to guarantee long range communications (up to 10 km) [8]. However this standard is not space qualified yet.

From the reported analysis some considerations can be drawn. To satisfy the requirement of autonomous relative navigation and control of LEO formations the ISL system has to be used for augmenting the CDGPS in real-time. Therefore, the selected ISL system must be simple, sufficiently accurate and robust. This means that only range measurements are considered for CDGPS augmentation . This actually excludes carrier-based techniques which are very accurate but are not simple to implement, and in general not robust due to ambiguity resolution issues. In this context, both one-way (GNSS-like) and two-way (transponders) code-based techniques represents the best choice to provide the navigation algorithm with robust decimeter accuracy ranging measurements in a simple way. Different technologies, such those using UWB signals or wireless-based, are very promising but, at the present time, neither mature nor robust enough to be considered for the relative navigation in satellite formation flying missions.

3 Filter Model

The filter model processing both CDGPS and ISL measurements represents an extension of the filter presented in [16] and using only GPS carrier phase and pseudorange observables. Therefore, for the sake of clarity, the main characteristics of the CDGPS only filter will be briefly recalled in the following sub-section before putting into evidence how ISL ranging measurements can be integrated in such a filtering scheme.

3.1 CDGPS only

The filter model is developed with reference to the Earth-Centered-Earth-Fixed (ECEF) reference frame and a formation of two satellites, named chief and deputy. The relative navigation problem is well described by the following nonlinear discrete stochastic model

$$\begin{cases} x_{n+1} = g_n(x_n, v_n) \\ y_n = h_n(x_n, w_n) \end{cases} \quad (1)$$

where x is the system state vector, y is the measurement vector, g is the non-linear state propagation function, h is the non-linear observation function, w is the process noise vector, v is measurement noise vector, and the subscript n is used to denote the variable value at the time t_n. Both noises are assumed to be additive, white, Gaussian with zero mean, uncorrelated in time, mutually uncorrelated and uncorrelated with the state vector at the same time sample. The following covariance matrixes for the variables defined in (1) are introduced: P_{xx} for x, P_{yy} for y, Q for v and R for w. The measurement model relies on dual-frequency, double difference observables. The state and the measurement vectors are thus given by

$$\begin{aligned} x &= \left(\underline{B} \quad \underline{\dot{B}} \quad VTEC_1 \quad VTEC_2 \quad {}_{L1}AM \quad {}_{L2}AM\right)^T \\ y &= \left({}_{L1}PR \quad {}_{L2}PR \quad {}_{L1}CP \quad {}_{L2}CP\right)^T \end{aligned} \quad (2)$$

where

$$\begin{aligned} x &\in \mathbb{R}^{\left[8+2(N_{sat}-1)\right]\times 1} \\ y &\in \mathbb{R}^{\left[4(N_{sat}-1)\right]\times 1} \end{aligned} \quad (3)$$

and, N_{sat} is the number of commonly viewed GPS satellites, $\underline{B}$ is the baseline vector, $\underline{\dot{B}}$ is the relative velocity, $VTEC_1$ and $VTEC_2$ are the vertical total electron content from, respectively, the chief and deputy receivers to the upper bound of the ionosphere, ${}_{L1}AM$ and ${}_{L2}AM$ are the DD integer ambiguity vectors on the L1 and L2 carriers, respectively, ${}_{L1}PR$ and ${}_{L2}PR$ are the pseudorange measurement vectors, and ${}_{L1}CP$ and ${}_{L1}CP$ are the carrier phase measurement vectors.

A double difference can be formed by subtracting two single difference equations of the same type and frequency, taken by the same two GPS receivers at the same time from two different GPS satellites, one of which, named pivot, is taken as a reference. The pivot satellite is selected as the one with the highest elevation with respect to the chief.

Fig. 2 shows a conceptual representation of the filtering scheme. At the generic time step n the filter receives in input the estimates of the state vector and covariance matrix, P_{xx}, at the previous step, along with the GPS observables from chief and deputy satellites, respectively. A coarse evaluation of the chief satellite ECEF position and velocity has to be provided as well. Measurements from commonly available GPS satellites are extracted in order to calculate DD observables. Since both the number of commonly viewed GPS satellite and the pivot satellite (i.e. the reference satellite to calculate DD observables) can change between subsequent observations a rearranging/re-initialization step is needed to correctly relate the bias term of the state vector to the current measurements.

The rearranged state vector and covariance matrix are used as inputs to an Extended Kalman Filter (EKF). The purpose of using the EKF is to perform a robust floating estimate of the DD ambiguities, whose integer values are then extracted by the LAMBDA method. This allows correcting the individual DD carrier phase observations, which can be then used

within a kinematic approach based on a Weighted Least Square (WLSQ) algorithm to produce a highly accurate baseline estimation. Residual errors in the carrier phase observables depend mainly on DD ionosphere path delays estimation errors. A common procedure to eliminate this effect is evaluating ionosphere free combinations of corrected carrier phase observables.

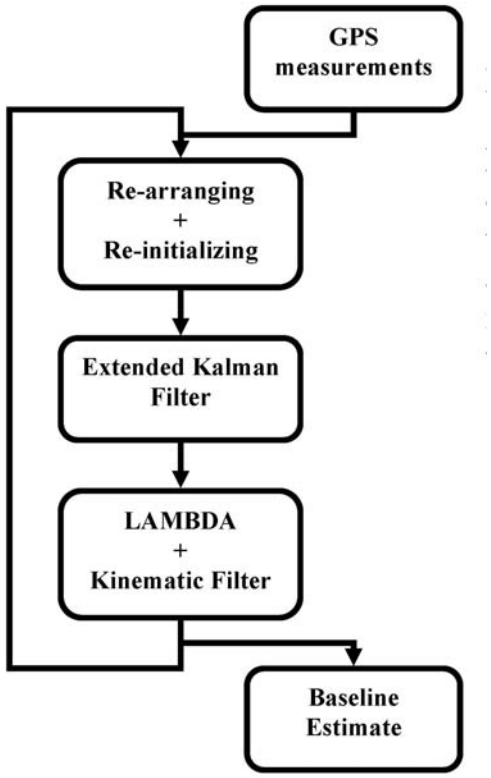


Fig. 1 Block Diagram of the CDGPS-only filtering scheme [16].

Non-linear Keplerian equations have been selected for modeling the satellites' relative motion in ECEF. The selected model has to be interpreted as a trade-off between using a nonlinear model to improve accuracy and having a computational load adequate to real-time implementation. In this way it is possible to get good performance by modeling the small differential perturbations (e.g. differential J2 and drag) as process noise while keeping low the computational effort. Indeed, formation flying satellites usually lie on orbits whose parameters slightly differ and have similar ballistic coefficients so to minimize control efforts to maintain the formation. VTEC terms and carrier phase ambiguities are treated as random walk plus random constant processes. It is worth noting that, once fixed, the carrier phase ambiguities do not change with time, in the absence of cycle slips. Thus, the random constant model should be considered as exact for these bias terms. However, introducing a small amount of process noise on carrier phase ambiguities allows keeping the filter receptive to new measurements, being it able to update these bias terms over the time, even if the convergence of filter covariance matrix has been already achieved.

3.2 CDGPS plus ISL

The addition of a ranging measurement to the filtering scheme reported above involves the modification of the observation model

$$y_n = h_n(x_n, w_n) \qquad (4)$$

because the measurement vector now includes the range measurement D

$$y = \left({}_{L1}PR \quad {}_{L2}PR \quad {}_{L1}CP \quad {}_{L2}CP \quad D \right)^T \qquad (5)$$

Moreover, the expression of Jacobian matrix of this non-linear model that is necessary for EKF implementation

$$H = \frac{\partial y}{\partial x} \qquad (6)$$

must be modified too, together with the measurement noise matrix R.

For the sake of deriving the non-linear observation function and the relevant Jacobian matrix, an observation model for the ISL ranging measurement must be introduced. In general, for both transponders and GNSS-like systems operating in LEO orbits, the ranging measurements can be affected by Time-Of-Flight (TOF) errors and ionospheric delays. The time of flight correction accounts for the motion of the platforms during signal travel time. Actually, for any given instant of reception of the ranging signal, the measured range does not represent the true baseline at that time instant, but the distance between the transmitting platform at the instant of transmission and the receiving platform at the instant of reception.

The difference between the desired baseline and the measured distance depends on the signal time of flight and on platforms velocity. To a first order it can be estimated as the ratio between the baseline and platform velocity; this means that if only short separations are considered (1-2 km at most) the expected TOF error is on sub-centimeter scale. The ionospheric delay, *IO*, depends on the electron density ρ_e at the considered LEO altitude [24]

$$IO = \frac{40.3}{f^2}\int_0^B \rho_e dB \qquad (7)$$

where f is the carrier frequency in Hz and lengths must be expressed in meters. A typical electron density value for 450-650 km altitude is $0.5 \cdot 10^{11}$ particle/m^3 thus leading to a ionospheric delay of less than 0.1 cm for S-band operation at 2 km.

According to those considerations, for the selected orbital scenario (see section 4.1) with limited inter-satellite separations the following observation model can be assumed for the ISL ranging measurements

$$D = \|\underline{B}\| + \sigma_D \qquad (8)$$

where σ_D represents the measurement noise. Therefore the relevant line in the H matrix is

$$\left(\frac{B_X}{\|\underline{B}\|} \quad \frac{B_Y}{\|\underline{B}\|} \quad \frac{B_Z}{\|\underline{B}\|} \quad 0_{1\times 5} \quad 0_{1\times 2(N_{sat}-1)} \right) \qquad (9)$$

In addition, a purely diagonal term accounting for the variance of the ranging measurement must be added to the R matrix. Finally, it is essentially to point out that the additional ranging measurement is used both in the EKF and in the kinematic filter processing only carrier-phase measurements corrected for integer ambiguities.

4 Numerical Results

4.1 Simulation Environment

Filter performance has been numerically evaluated by using the Satellite Navigation Toolbox 3.0 for Matlab® developed by GPSoft, Inc. [25]. The toolbox is capable of simulating the GPS satellite constellation, the propagation environment and the receiver measurements, where the nominal GPS satellite constellation [25], based on 24 satellites distributed on 6 orbital planes, has been assumed.

Since the Satellite Navigation Toolbox was developed mainly for Earth-based GPS receivers, a major modification to the ionosphere delay model has been introduced to adequately model first and second order ionosphere delay terms for LEO satellites [16]. In addition, the Toolbox implicitly assumes a constant value for the VTEC. In order to overcome this limitation the toolbox has been complemented with a IONEX (IONosphere map Exchange file format) reader. Indeed, VTEC maps are provided by the International GPS service as IONEX files [26]. Unfortunately, these maps are relevant only to ground-based users. Nevertheless, for the sake of using IONEX products to estimate the VTEC behaviour along the orbit, a coefficient can be defined to scale the IONEX VTEC values to a LEO receiver as in [27].

Filter performance is evaluated by comparison with a simulated truth obtained by running an orbit propagator including a seven-order model of the gravity field and the aerodynamic drag, and with reference to orbit scenarios relevant to important Earth remote sensing applications. Specifically, an helix-like [28] configuration is considered relevant to interferometric and bistatic SAR applications, such as the generation of digital elevation model (DEM) by SAR interferometry.

Table I Parameters of the selected helix formation

	Chief	Deputy
Semi-major axis (km)	6997.94	
Eccentricity	0.00118	0.001308
Anomaly of perigee (°)	90	
Inclination (°)	97.87	
Ascending node right ascension (°)	0	0.0049
Initial mean anomaly (°)	0	

As shown in Tab. I, the considered orbits share inclination, semi-major axis and anomaly of perigee. Fig. 3 reports the variation of the baseline along one orbit for the considered helix formation with respect to the chief satellite Orbiting Reference Frame (ORF), whose origin is in the chief satellite center of mass, the x-axis is directed along chief velocity, and the z-axis is directed towards the Earth's center. It is worth noting that the axes of the ORF are computed considering the velocity of the chief satellite in ECEF, for this reason the familiar helix shape cannot be recognized from Fig.3.

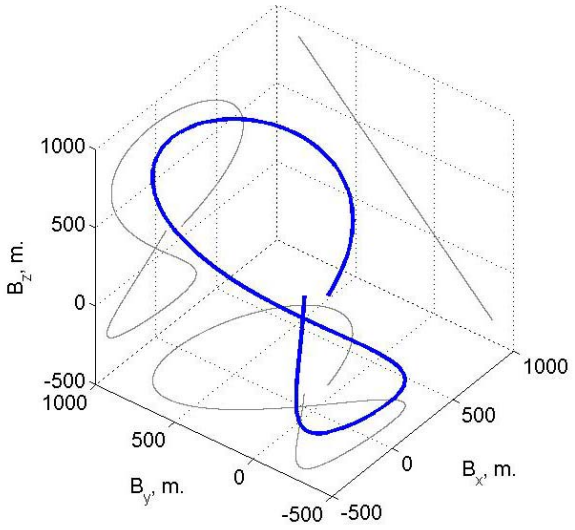


Fig. 3 Variation of the baseline in ORF during a single orbit for the selected helix formation

In all the simulation the Italian COSMOSkyMed mission is taken as a reference for the master satellite and present generation GPS receiver nominal performance is used, specifically 0.2 m and 0.25 m pseudorange measurement noise on L1 and L2 channels, respectively, and 1 cm carrier phase noise both on L1 and L2. In addition, 1-Hz GPS data frequency update and 10° mask angle are considered. Finally, ISL ranging measurement are simulated considering 10 cm measurement noise. No meaningful propagation effects are expected for the ISL system due to the selected helix configuration characterized by limited inter-satellite separations. The chief satellite is supposed to receive deputy GPS observables by means of the ISL. The filter is supposed to run on the processing unit of the chief satellite.

All the results are derived with reference to a single orbit scenario. The considered GPS satellite geometry and ionosphere status are the same selected in [16], representing an example of a worst case condition characterized by very unfavorable GPS geometries, with Geometric Dilution of Precision (GDOP) up to 18, and anomalous ionosphere conditions, with VTEC up to 12 TECU (Total Electron Content Unit, corresponding to 10^{16} electrons per square meter) as reported in Fig 4 and Fig 5. More in detail, it is important to point out that due to the inter-satellite separation (less than 1.5 km at most) VTEC values for master and slave satellites are very close, the maximum difference being less than 0.1 TECU.

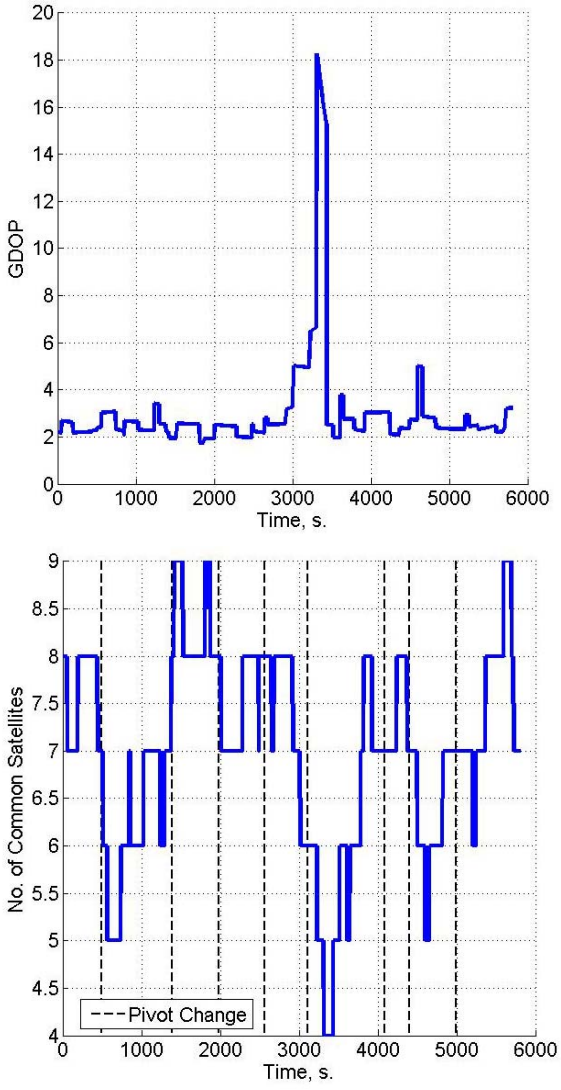


Fig. 4 GDOP (top) and number of common GPS satellites (bottom) during the selected single orbit scenario

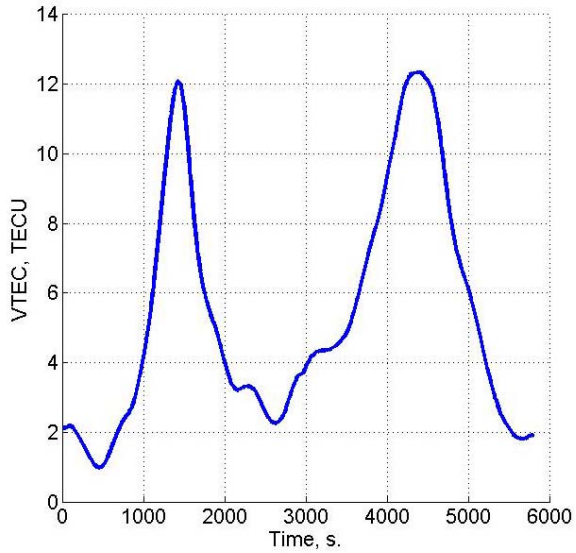


Fig. 5 Simulated VTEC time profiles for the chief satellite

4.2 Performance Verification

The performance of the CDGPS augmented filter presented in this paper is compared to that obtained by the GPS-only filter in [16]. The two filters share the same structure but the latter does not use ISL ranging measurements. Before presenting the results it is essential to point out that when GPS satellite geometry is good in terms of both GDOP and number of common satellites, and when the filters are able to correctly estimate integer ambiguities, the results must be necessarily very similar, since the accuracy of the baseline calculated with the kinematic algorithm is essentially determined by carrier-phase measurements, that are one order of magnitude more accurate than ISL ones.

Two cases have been then selected as representative of the improvement achievable by CDGPS augmentation.

In the first case, both filters are able to correctly fix integer ambiguities, however due to the fact that there are only 4 common GPS satellites (see Fig. 4), the CDGPS-only filter cannot fully exploits the kinematic algorithm, with an evident growth of the baseline error. Specifically, the baseline norm exhibits a maximum error larger than 0.85 m (Fig. 6). This error is determined by the errors in the radial and along-track components of the baseline, with the radial error contributing for the most part (Fig. 7). As expected, the possibility of using an additional ranging measurement can be properly exploited by the present filter structure, thus reducing notably both the vertical and along-track estimation errors and allowing filter solution to be less sensitive to GPS satellite geometry. Table II summarizes filters' performance. The utilization of ranging measurements reduces both the maximum baseline norm error and the maximum radial component error by a factor of five. In addition, the maximum along-track error is reduced by a factor close to four and also all the rms errors improve. Finally, it is worth noting that when only 4 common GPS satellites are available the CDGPS-only filter cannot derive an estimate of the filter error covariance by the kinematic algorithm, whereas the additional ranging measurement (Fig. 7) guarantees proper covariance estimations.

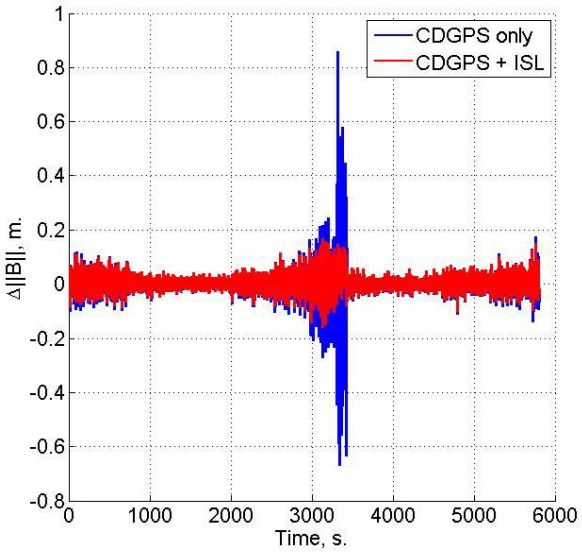


Fig. 6 Baseline norm error (first case).

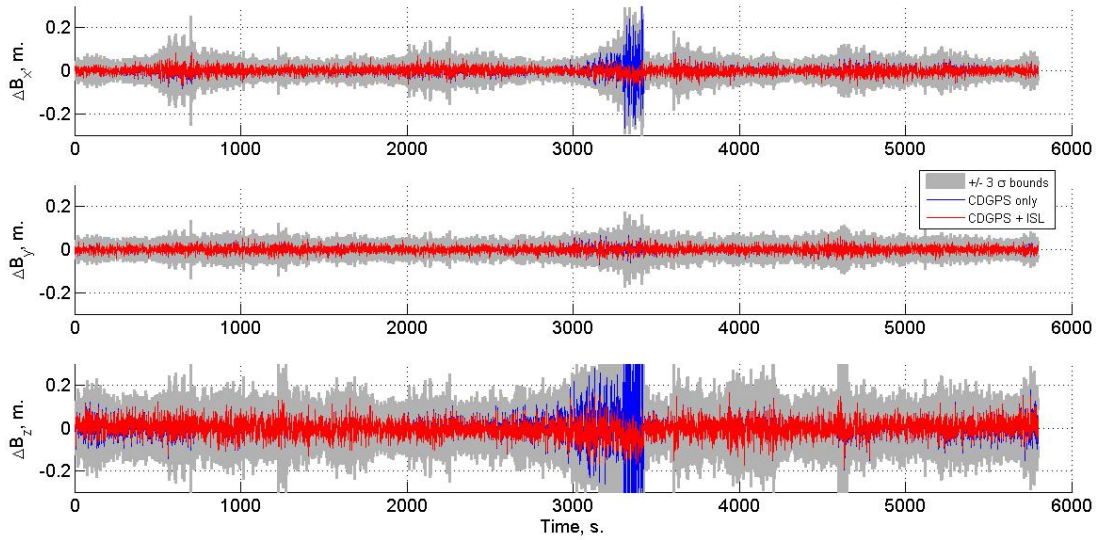


Fig. 7 ORF baseline components error (first case); +/- 3 s bounds are relevant to CDGPS + ISL.

Table II Performance Statistics (first case)

	$\|\underline{B}\|$		B_x		B_y		B_z	
	Max	RMS	Max	RMS	Max	RMS	Max	RMS
CDGPS-only	86 cm	5 cm	30 cm	2 cm	-7 cm	2 cm	-97 cm	6 cm
CDGPS + ILS	17 cm	3 cm	8 cm	2 cm	7 cm	1 cm	17 cm	4 cm

In the second case, characterized by worse measurement noise conditions, the filter with no range measurements is not able to correctly fix integer ambiguities in presence of a poor GPS satellite geometry . Specifically, a wrong integer ambiguity value is derived around t = 3300 s. In the following time instants, even if the GDOP improves the filter is not able to recover good ambiguity estimates, and a diverging behavior can be thus observed in Fig. 8-9. In the same condition, in the CDGPS-augmented filter the ranging measurement effectively aids the integer ambiguity estimation process, thus preventing the diverging behavior observed in the CDGPS-only filter. Performance statistics for this second case are not reported since those of the CDGPS-augmented filter are very similar to the previous case, while the performance statistics of CDGPS-only filter are strongly influenced by wrong integer ambiguities and filter divergence.

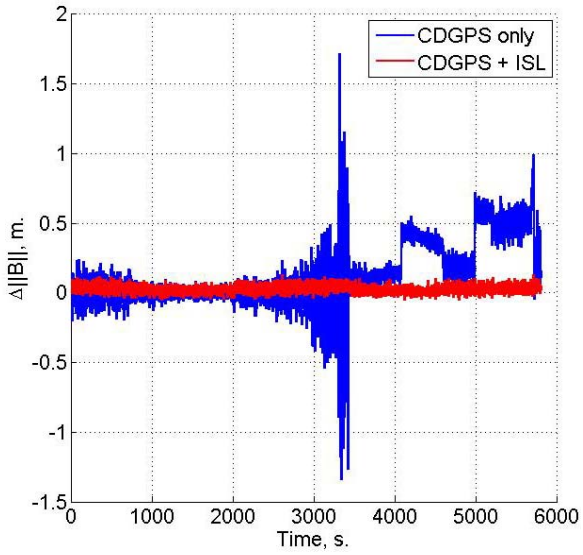


Fig. 8 Baseline norm error (second case).

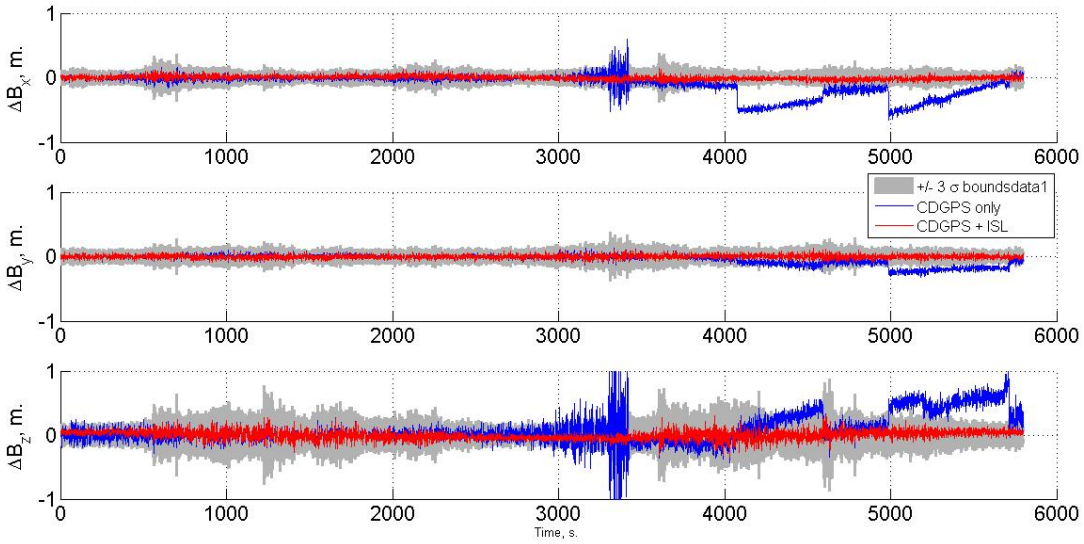


Fig. 9 ORF baseline component errors (second case); +/- 3 s bounds are relevant to CDGPS + ISL.

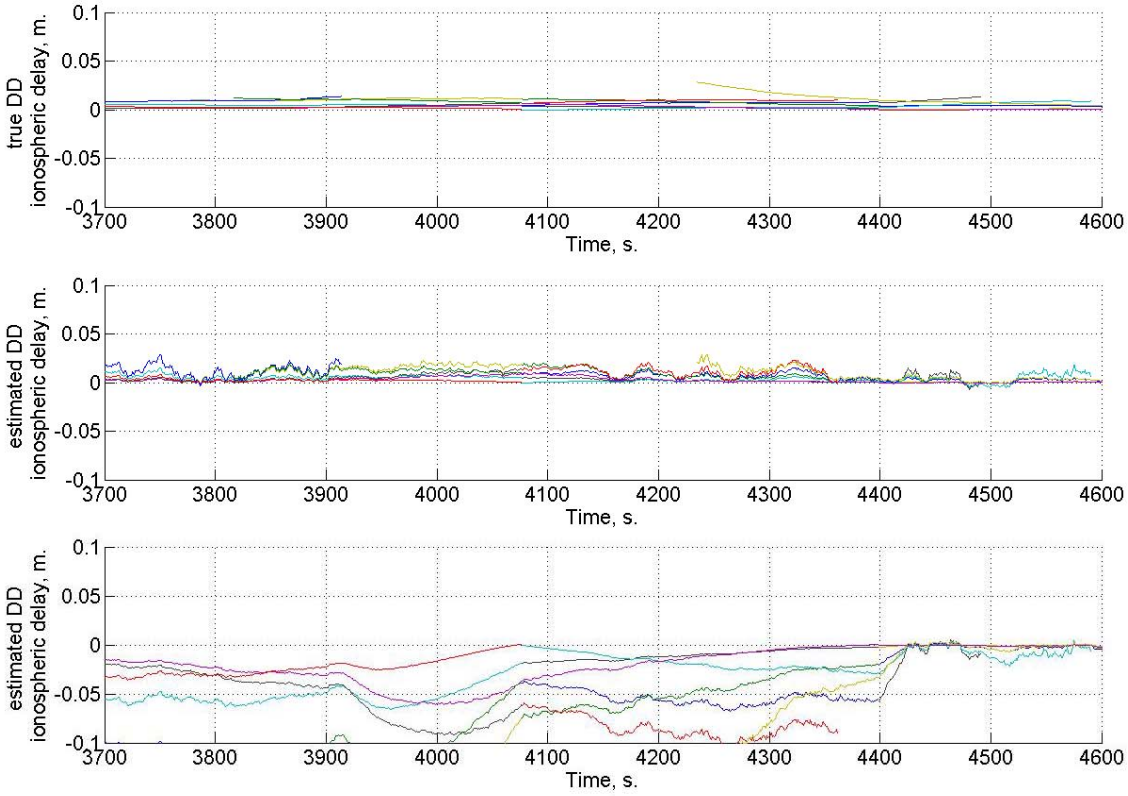


Fig. 10 True DD ionospheric delays (top) versus estimated DD ionospheric delay by CDGPS + ISL (middle) and CDGPS only (bottom).

Finally, Fig.10 shows the DD ionospheric delays as estimated by the considered filters. When wrong ambiguities are computed the CDGPS-only filter overestimates the ionospheric delays, thus preventing the filter to correctly estimates the following ambiguities. The availability of ranging measurements, instead, allows the filter to adequately compute

DD ionospheric delays, even though their order of magnitude is close the carrier-phase noise.

5 Conclusion

In this paper an augmented CDGPS technique has been proposed for the autonomous relative navigation of LEO satellite formations. Indeed, the augmented filter shall exhibit improved robustness to poor GPS geometries and noisy measurements with respect to the CDGPS-only filter. Specifically, the proposed filter integrates double difference code and carrier phase GPS measurements on two frequencies with local ranging measurements provided by an inter-satellite radio link with decimeter-level accuracy. Augmented filter performance have been verified by numerical simulations of LEO satellite formations relevant to next generation SAR missions characterized by short (a few kilometers) and variable satellite separations. Results demonstrate that the augmented filter preserves the CDGPS-only filter capability of correctly computing the double-difference integer ambiguities in presence of poor GDOP. In addition, also when ambiguities are correctly computed, the availability of ranging measurements improve the accuracy in the estimate of the baseline radial component.

References

[1] Krieger, G., Moreira, A., "Spaceborne bi- and multistatic SAR: Potential and challenges," in *Proc. IEE Radar, Sonar Navig.*, 2006, Vol. 153, no. 3, pp. 184–198.

[2] Persson, S., et. Al, "Prisma – An Autonomous Formation Flying Mission", *ESA Small Satellite Systems and Services Symposium* (4S), September 2006, Sardinia, Italy, pp. 25-29.

[3] Leung, S., Montenbruck, O., "Real-Time Navigation of Formation-Flying Spacecraft Using Global Positioning System measurements", *Journal of Guidance, Control and Dynamics*, March-April 2005, Vol.28, No.2, pp. 226-235.

[4] Ebinuma, T., Bishop, R. H., Lightsey, E. G., "Integrated Hardware Investigations of Precision Spacecraft Rendezvous Using the Global Positioning System", *Journal of Guidance, Navigation and Control*, Vol.26, No.3, May-June 2003, pp. 425-433.

[5] Gill, E., Runge, H. , "Tight formation Flying for an along-track SAR interferometer", *Acta Astronautica*, 2004, Vol. 55, pp. 473-485.

[6] Kroes R., Montenbruck, O., Bertiger, W. , Visser, P., "Precise GRACE baseline determination using GPS", *GPS Solutions*, 2005, Vol. 9, no. 1, pp. 21–31.

[7] Krieger, G., et al., "TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry", *IEEE Trans. on Geoscience and Remote Sensing*, 2007, Vol. 45, no. 11, pp. 3317-3341.

[8] Garcìa-Rodriguez, A., et al. "Formation Flight Radio-Frequency Metrology", ESA Technical Note, TEC-ETN/20007.04, Issue 1, May 2007.

[9] Park, C.-W., "Precise Relative Navigation using Augmented CDGPS", Ph.D. Thesis, Stanford University, Department of Mechanical Engineering, Cambridge MA, June 2001.

[10] Persson, S., Veldman, S., Bodin, P., "PRISMA—A formation flying project in implementation phase", *Acta Astronautica*, 2009, Vol. 65, pp. 1360–1374.

[11] Zenick, R. G., Kohlhepp, K., "GPS Micro Navigation and Communication System for Clusters of Micro and Nanosatellites", *IEEE Aerospace Conference*, March 2001, Big Sky MT, USA, Vol. 5, pp. 2515-2522.

[12] Stadter, P. A., Chacos, A. A., Heins, R. J., Moore, G. T., Olsen, E. A., Asher, M. S., "Confluence of Navigation, Communication, and Control in Distributed Spacecraft Systems", *IEEE Aerospace Conference*, March 2001, Big Sky MT, USA, Vol. 2, pp.563-578.

[13] Adams, J. C., Gregorwich, W., Capots, L., Liccardo, D., "Ultra-Wideband for Navigation and Communications", , *IEEE Aerospace Conference*, March 2001, Big Sky MT, USA, Vol. 2, pp.785-792.

[14] Maessen, D., Gill, E., "Relative Orbital Element Estimation and Observability Analysis for Formation Flying Satellites using Inter-Satellite Range Measurements Only", *AIAA Guidance, Navigation, and Control Conference*, Toronto, Ontario, Aug. 2-5, 2010.

[15] Carpenter, J., Alfried, K., "Navigation Accuracy Guidelines for Orbital Formation Flying", *Journal of the Astronautical Sciences*, 2005, Vol. 53, No. 2, pp. 207-219.

[16] Tancredi, U., Renga, A., Grassi, M., "GPS-based Relative Navigation of LEO formations with Varying Baselines", *AIAA Guidance Navigation and Control Conference*, Toronto, Canada, Agosto 2010.

[17] de Jonge, P., Tiberius, C., "The LAMBDA method for integer ambiguity estimation: implementation aspects", LGR-Series Publications of the Delft Geodetic Computing

Centre, Vol. 12, 1996.
[18] Grelier T. et al., "GNSS in Space – Part2 Formation Flying Radio Frequency Techniques and Technology", *InsideGNSS*, January/February 2009, pp. 43-51.
[19] Purcell, G., Kuang, D., Lichten, S., Wu, S.-C., Young, L., "Autonomous formation flyer (AFF) sensor technology development", Technical report, NASA, TMO Progress Report 42-1341998.
[20] Aung, M., Purcell, G. H., Tien, J. Y., Young, L. E., Amaro, L. R., Srinivan, J., Ciminera, M. A., Chong, Y. J., "Autonomous formation flying sensor for the starlight mission", Proceedings of *the International Symposium Formation Flying Missions and Technologies*, October 29–31, Toulouse, 2002.
[21] http://nssdc.gsfc.nasa.gov/multi/tdrs.html
[22] Wang, S., Zhang, E., "Inter-satellite Radio Links and Spread-spectrum Ranging for Satellite Formation Flying", *3rd International Conference on Microwave and Millimeter Wave Technology*, August 2002, pp. 233-236.
[23] Adams, J. C., Gregorwich, W., Capots, L., Liccardo, D., "Ultra-Wideband for Navigation and Communications", , *IEEE Aerospace Conference*, March 2001, Big Sky MT, USA, Vol. 2, pp.785-792.
[24] Klobuchar, J. A., "Ionospheric effects on GPS", in Parkinson B.W., Spilker, J.J. (eds) *Global positioning system: theory and applications*, vol. 1, American Institute of Aeronautics and Astronautics Inc., Washington, pp. 485–515, 1996.
[25] http://www.gpsoftnav.com/satnav.html
[26] ftp://cddisa.gsfc.nasa.gov/gps/products/ionex.
[27] Montenbruck, O., Gill, E., "Ionospheric Correction for GPS Tracking of LEO satellites", *The Journal of Navigation*, 2002, Vol. 55, pp. 293-304.
[28] Renga, A., Tancredi, U., Grassi, M., "GPS-based Relative Navigation in Earth Observation Missions Relying on Cooperative Satellites", in R. Sandau et al. (eds.), Small Satellite Missions for Earth Observation, 2010, pp. 467-475.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Model Predictive Control Strategies for Spacecraft Formation Flying Applications

A. Valmorbida and G. Baù
Center of Studies and Activities for Space (CISAS) "G. Colombo"
Via Venezia 15, 35131 Padova, Italy

E. C. Lorenzini
Department of Mechanical Engineering, University of Padova
Via Venezia 1, 35131 Padova, Italy

G. Sechi
Thales Alenia Space Italia S.p.A.
Strada Antica di Collegno 253, 10146 Torino, Italy

Keywords: *formation control strategies, MPC, LQR, spacecraft formation flying.*

Abstract

Formation flying missions require high performance orbital control strategies, which are based on the accurate analysis of the spacecraft relative dynamics and on efficient control techniques.

Model Predictive Control (MPC) is a modern optimal control strategy that is very attractive for formation flying and rendezvous missions, mainly because of its ability to cope with constraints on control actions and dynamic state of the system. Every real space mission indeed imposes constraints, as for example the maximum control force a thruster system can apply, or the relative positions and/or velocities that satellites flying in formation should have to guarantee formation safety. MPC can take into account both control and state constraints directly in the optimal control action computation, resulting in a more efficient control system than other control strategies.

In this paper we present a study on the MPC application to a formation acquisition maneuver for two space vehicles, taking into account two scenarios: a Leader-Follower formation and a Projected Circular Orbit formation. The performances of an MPC controller are compared with those of a LQR controller in carrying out the same maneuver, evaluating also the effects of the gravitational harmonics J_2, J_3 and drag perturbations on the proposed maneuvers.

1 Introduction

Spacecraft Formation Flying has attracted much attention in recent years and several space agencies have proposed various missions for Earth and astronomical observations using this technique.

The design of an efficient and reliable Guidance Navigation and Control (GNC) system, together with a proper selection of actuators and sensors, is important and can deeply affect the feasibility of a space mission. Linear systems theory is a very effective and commonly used tool in designing GNC systems for formation flying systems, as for instance Linear Quadratic Regulator (LQR) or Linear Kalman Filter. Model Predictive Control is an interesting tool for formation flying and

rendezvous GNC applications, mainly because of its re-planning nature and constraint-handling ability. All these tools are based on linearized models of the relative formation dynamics and reliable linear model are therefore necessary for the implementation of these GNC approaches.

On the other hand, testing the robustness of GNC strategies and actuators/sensors systems is another significant aspect, which in turns needs both a reliable software simulator with an accurate and fast dynamic propagator, and a formation flying hardware simulator.

The final aim of this work is to analyze and compare the performances of two types of controller, LQR-based and MPC-based, while carrying out a formation acquisition maneuver for a Leader-Follower and a Projected Circular Orbit spacecraft formations in a perturbed environment.

The paper is organized as follows. Section 2 describes the main features of the software simulator used to test the formation flying control strategies under investigation. Section 3 presents a review of LQR and MPC is with a brief discussion of the tuning of those parameters that affect the performances of both controllers. In the next two sections we analyze and compare the behavior of a LQR-based controller and a MPC-based controller while performing a formation acquisition maneuver for a Leader-Follower formation in Section 4, and a Projected Circular Orbit formation in Section 5, showing the effects of actuator saturation and of perturbations. In both the scenarios, only one spacecraft, the Deputy, performs an autonomous control by means of continuous thrusters, while the other spacecraft, the Chief, represents the free-flying reference spacecraft.

2 Satellite Formation Flying Simulator

We developed a software simulator for Spacecraft Formation Flying (SFF) applications that enables the user to design and validate GNC strategies with various actuator and/or sensor architectures. The simulator presents a modular structure so that the user can first design a particular element of the simulator, and then evaluate its efficiency and robustness taking into account the environment in which the satellite formation operate and the behavior of the sensor and/or actuator systems.

The simulator was developed in Matlab® using Object Oriented Programming (OOP) and is organized in the following eight main modules (see Fig. 1): a) a nonlinear *DYNAMICS PROPAGATOR* of the system dynamics; b) a *SENSORS* module that simulates the sensors system while producing measurements from the system dynamic state, taking into account noise, resolution and bias; c) a *NAVIGATION* module that performs an (optimal) estimation (via Kalman filtering) of the current state of the system using outputs from the SENSORS module; d) a *GUIDANCE* module that computes the desired trajectory that each satellite of the formation has to follow e) a *CONTROL* module that computes the (optimal) control action that should be performed by the actuators in order to track the desired trajectory; e) an *ACTUATOR* module that models the behavior of the actuators while carrying out the control actions provided by the CONTROL module; f) an *ENVIRONMENT* module that computes the perturbative effects on the system dynamics; g) a *MODE MANAGER* module that manages the transition between different formation flying modes.

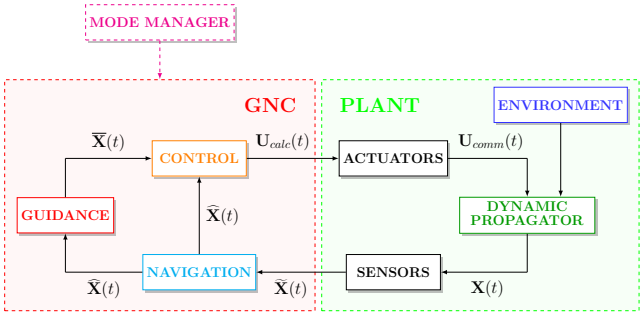


Fig. 1 Main modules of the SFF software simulator

The nonlinear *DYNAMICS PROPAGATOR* is capable of propagating the absolute orbital dynamics of each satellite in formation independently, and the user can choose between the classical Cowell's method or DROMO-FF. The latter is a recently proposed[2] propagator for formation flying orbital motion, which exploits a very efficient regularized method for

the perturbed two-body problem, named DROMO, developed by Peláez in 2006[1]. The relative dynamics of each spacecraft with respect to a reference mass is then properly computed from the absolute dynamics.

DROMO consists in a new formulation based on the decomposition of the two-body problem into the perturbed harmonic oscillation of the orbital radius and the dynamics of the osculating orbital plane governed by a unit quaternion. The method employs regularized elements, whose differential equations are derived through the variation of parameter technique. The method shows to be sensibly more accurate and faster in computing time than the classical Cowell's and Encke's methods and competitive with other very efficient regularized schemes (see Ref. [1]).

Two types of control strategies are designed and used in the *CONTROL* module to perform a formation flying maneuver: the first one is based on Linear Quadratic Regulator (LQR)[3], the second one is based on Model Predictive Control (MPC)[4]. In the following Section 3 we briefly introduce both control types, restricted to the form of the problem solved here, discussing the tuning of those parameters that mainly affect the closed-loop performances.

3 Optimal Control Strategies

The use of optimal linear control and estimation theories in the design of the GNC system is a common strategy in formation flying missions.

In the synthesis of both MPC and LQR controller we adopt the Hill-Clohessey-Wiltshire (HCW) relative motion model, which can be written in state-space form as follows[5]:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\,\mathbf{x}(t) \end{cases} \qquad (1)$$

where $\mathbf{x}(t) = [\,x, y, z, \dot{x}, \dot{y}, \dot{z}\,]^T(t) \in R^6$ is the state vector, $\mathbf{u}(t) \in R^3$ the input or control vector, $\mathbf{y}(t) \in R^6$ the output vector. The state matrix $\mathbf{A} \in R^{6x6}$, the control matrix $\mathbf{B} \in R^{6x3}$ and the output matrix $\mathbf{C} \in R^{6x6}$ are as follows:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3n^2 & 0 & 0 & 0 & 2n & 0 \\ 0 & 0 & 0 & -2n & 0 & 0 \\ 0 & 0 & -n^2 & 0 & 0 & 0 \end{bmatrix} \qquad (2)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{0}_{3x3} \\ \mathbf{I}_{3x3} \end{bmatrix} \qquad (3)$$

$$\mathbf{C} = \mathbf{I}_{6x6} \qquad (4)$$

where n is the Chief's mean motion. The relative dynamics is expressed in the Hill's Local Vertical Local Horizontal (LVLH) rotating cartesian reference frame, with x axis in the radial direction, y axis in the along-track direction and z axis in the cross-track direction.

The dynamic model used in both controllers is a discrete time version of the above Continuous-time model, resulting in a Discrete-time Linear Time Invariant (DLTI) model:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_d\,\mathbf{x}_k + \mathbf{B}_d\,\mathbf{u}_k \\ \mathbf{y}_k = \mathbf{C}_d\,\mathbf{x}_k \end{cases} \qquad (5)$$

where $\mathbf{A}_d$, $\mathbf{B}_d$ and $\mathbf{C}_d$ are the discrete-time matrices of the system obtained from the continuous ones as follows:

$$\mathbf{A}_d = e^{\mathbf{A}t}\Big|_{t=T_s} = \mathbf{\Phi}(t)\Big|_{t=T_s} = \qquad (6)$$

$$= \begin{bmatrix} 4-3c & 0 & 0 & \frac{s}{n} & \frac{2}{n}(1-c) & 0 \\ 6(s-nT_s) & 1 & 0 & -\frac{2}{n}(1-c) & \frac{1}{n}(4s-3nT_s) & 0 \\ 0 & 0 & c & 0 & 0 & \frac{s}{n} \\ 3ns & 0 & 0 & c & 2s & 0 \\ -6n(1-c) & 0 & 0 & -2s & 4c-3 & 0 \\ 0 & 0 & -ns & 0 & 0 & c \end{bmatrix}$$

$$\mathbf{B}_d = \int_0^{T_s} \mathbf{\Phi}(t)\,\mathbf{B}\,dt = \qquad (7)$$

$$= \begin{bmatrix} \frac{1}{n^2}(1-c) & \frac{2}{n}(T_s - \frac{s}{n}) & 0 \\ -\frac{2}{n}(T_s - \frac{s}{n}) & \frac{4}{n^2}(1-c) - \frac{3}{2}T_s & 0 \\ 0 & 0 & \frac{1}{n^2}(1-c) \\ \frac{s}{n} & \frac{2}{n}(1-c) & 0 \\ -\frac{2}{n}(1-c) & \frac{4s}{n} - 3T_s & 0 \\ 0 & 0 & \frac{s}{n} \end{bmatrix}$$

$$\mathbf{C}_d = \mathbf{C} \qquad (8)$$

with $\mathbf{\Phi}(t)$ the HCW state transition matrix, T_s the sampling period of the control, also called *control interval*, $c = \cos(nT_s)$ and $s = \sin(nT_s)$. We assume that the control action is piecewise constant, i.e. it remains constant over the control interval and can change only between two consecutive intervals.

The trajectory tracking problem is seen in the light of a regulation problem for the current state error vector $\mathbf{e}_k = \mathbf{x}_k - \bar{\mathbf{x}}_k$, i.e. the instantaneous deviation between the current (estimated) state vector $\mathbf{x}_k$ and the current reference state vector $\bar{\mathbf{x}}_k$.

Both controllers use a state *feedback control law*; however, for the problem under investigation, that is LIT, while the LQR control law can be computed *off-line* and the control acceleration constraints (i.e., saturations) are considered *a posteriori* (in the *ACTUATORS* module), the MPC control law is obtained by solving *on-line*, at *each* control step, a finite horizon open-loop optimal control problem that takes into account the control constraints.

3.1 Linear Quadratic Regulator

The purpose of the Linear Quadratic Regulator (LQR) synthesis is to find a state feedback control law in the form $\mathbf{u}_k = -\mathbf{K}(\mathbf{x}_k - \bar{\mathbf{x}}_k)$, where $\mathbf{K}$ is the feedback gain matrix, with the goal of minimizing the following linear quadratic performance index[6]:

$$J = \sum_{k=0}^{\infty} [\mathbf{e}_k{}^T \mathbf{Q} \mathbf{e}_k + \mathbf{u}_k^T \mathbf{R} \mathbf{u}_k] \qquad (9)$$

where $\mathbf{e}$ is the state error vector, and $\mathbf{Q} \in R^{6x6}$ and $\mathbf{R} \in R^{3x3}$ are positive-definite matrices representing the state error weighting matrix and the control weighting matrix, respectively. The performance index J represents the global energy of the system and is the sum of two terms representing the error system energy and the control energy, respectively.

Since (A,B) is controllable and (A,C) is observable, there is one and only one optimal controller, whose gain matrix is:

$$\mathbf{K} = \left(\mathbf{R} + \mathbf{B}_d{}^T \mathbf{P} \mathbf{B}_d\right)^{-1} \mathbf{B}_d{}^T \mathbf{P} \mathbf{A}_d \qquad (10)$$

where $\mathbf{P}$ is the unique positive definite solution to the Discrete time Algebraic Riccati Equation (DARE):

$$\mathbf{A}_d{}^T \mathbf{P} \mathbf{A}_d - \mathbf{P} + - \left(\mathbf{A}_d{}^T \mathbf{P} \mathbf{B}_d\right)\left(\mathbf{R} + \mathbf{B}_d{}^T \mathbf{P} \mathbf{B}_d\right)^{-1}\left(\mathbf{B}_d{}^T \mathbf{P} \mathbf{A}_d\right) + + \mathbf{Q} = \mathbf{0} \qquad (11)$$

Matrices $\mathbf{Q}$ and $\mathbf{R}$ are parameters that affect the reference trajectory tracking performances and fuel consumption: higher values of $\mathbf{Q}$ elements with respect to $\mathbf{R}$ elements mean higher weights on state error with respect to control action, that leads to a more aggressive control, a better performance in tracking the desired trajectory but a higher fuel consumption.

3.2 Model Predictive Control

Model Predictive Control (MPC)[7][8], also referred to as Receding Horizon Control, is a modern control technique widely adopted in industry as an effective means to deal with multivariable constrained control problems. MPC is very attractive also for formation flying and rendezvous guidance and control applications, because of its re-planning nature and constraint-handling ability.

Model Predictive Control is based on the idea, illustrated in Fig. 2, of employing an explicit model of the system to be controlled which is used to predict and optimize its future behavior. This prediction capability allows solving optimal control problems on-line, where the tracking error is minimized over a future horizon, called *prediction horizon*, possibly subject to constraints on the manipulated inputs and outputs. The result of the optimization is applied according to a *receding horizon philosophy*: at time k only the first input of the optimal command sequence is actually applied to the plant. The remaining optimal inputs are discarded, and a new optimal control problem is solved at time $k+1$. As new measurements are collected from the plant at each time k, the

receding horizon mechanism provides the controller with the desired feedback characteristics.

The MPC control action at time k is obtained by minimizing the following performance index with respect to the sequence of the input increments $\Delta u(k\,|\,k)$, ... , $\Delta u(m-1+k\,|\,k)$:

$$J_k = \sum_{i=0}^{p-1} \left\| \mathbf{x}(k+i+1\,|\,k) - \bar{\mathbf{x}}(k+i+1) \right\|^2_{\mathbf{Q}(i)} + \\ + \left\| \Delta\mathbf{u}(k+i\,|\,k) \right\|^2_{\mathbf{R}_{\Delta u}(i)} + \left\| \mathbf{u}(k+i\,|\,k) - \mathbf{u}_{\text{target}}(k+i) \right\|^2_{\mathbf{R}_u(i)} \qquad (12)$$

subject to the following constraints:

$$\begin{aligned} &\mathbf{u}_{\min}(i) \le \mathbf{u}(k+i\,|\,k) \le \mathbf{u}_{\max}(i) \\ &\Delta\mathbf{u}_{\min}(i) \le \Delta\mathbf{u}(k+i\,|\,k) \le \Delta\mathbf{u}_{\max}(i) \\ &\mathbf{y}_{\min}(i) \le \mathbf{y}(k+i\,|\,k) \le \mathbf{y}_{\max}(i) \\ &\Delta\mathbf{u}(k+h\,|\,k) = \mathbf{0} \\ &i = 0,\ldots,p-1 \\ &h = m,\ldots,p-1 \end{aligned} \qquad (13)$$

where '$(k+i\,|\,k)$' denotes the value predicted for time $k+i$ based on the information available at time k, p is the *prediction horizon*, m is the *control horizon* and $\mathbf{Q}$, $\mathbf{R}_{\Delta u}$ and $\mathbf{R}_u$ are output error, input increment and input weighting matrices, respectively.

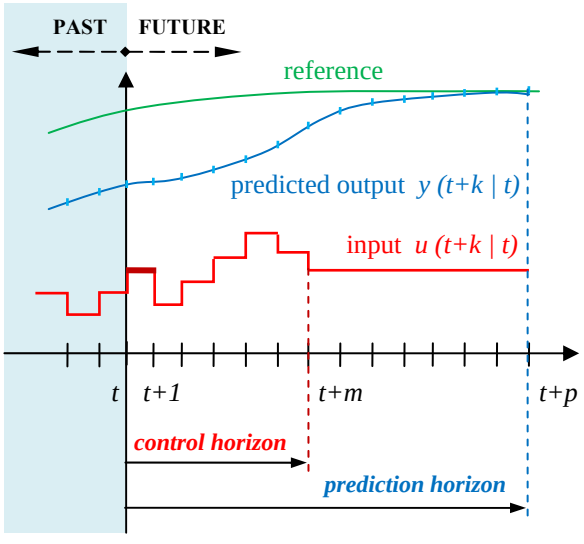


Fig. 2 Idea behind Model Predictive Control

The difference of a MPC with respect to a LQR is clear: a MPC solves an optimal control problem *on-line* for the *current state* of the plant, rather than determining *off-line* a feedback policy that provides the optimal control for *all states*.

When designing a MPC controller one should take into account the following: a more aggressive controller with higher fuel consumption and better tracking performances is obtained with smaller prediction horizon p, tighter constraints on control variables and higher ratio between state error weight and control action weight; a larger control horizon m leads to an optimization problem with more degrees of freedom, thus a "more" optimal (better performance) but more complex (computationally demanding) control.

4 Leader-Follower formation

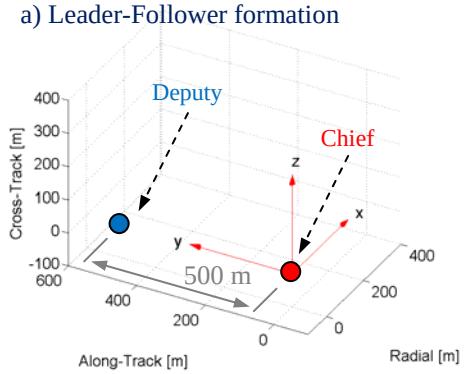
In this section we compare and discuss the performances of a MPC controller and a LQR controller in carrying out the same formation acquisition maneuver for a two spacecraft, $500m$ apart, Leader-Follower formation (see Fig. 3a), with active constraints on the maximum control acceleration that the propulsion system can actuate.

The initial dynamic state of the Chief is given in terms of orbital elements of the osculating keplerian orbit: semi-major axis $a = 6700\,km$, eccentricity $e = 0$, right ascension of the ascending node (RAAN) $\Omega = 0\,\text{deg}$, inclination $i = 97.87\,\text{deg}$, argument of perigee $\omega = 0\,\text{deg}$ and true anomaly $\vartheta = 0\,\text{deg}$. The initial dynamic state of the Deputy is given in Table 1 as relative position $\mathbf{r}_i$ and velocity $\mathbf{v}_i$ vectors with respect to the Hill's frame attached to the chief spacecraft. The previous relative conditions are computed using the Hill-Clohessey-Wiltshire (HCW) solution with the 1st-order bounded relative motion condition $\dot{y}_0 = -2\,n\,x_0$, that is[5]:

$$\begin{cases} x(t) = \rho_x \sin(nt + \varphi_{xy}) \\ y(t) = 2\rho_x \cos(nt + \varphi_{xy}) + \Delta y \\ z(t) = \rho_z \sin(nt + \varphi_z) \end{cases} \qquad (14)$$

with amplitudes of the in-plane and out-of-plane bounded motion $\rho_x = 8\,m$ and $\rho_z = 8\,m$, respectively, along-track offset $\Delta y = 500\,m$, and in-plane and out-of-plane phase angles $\varphi_{xy} = 45\,\text{deg}$ and $\varphi_z = 90\,\text{deg}$, respectively. Therefore, we assume that before the formation acquisition maneuver the Deputy follows a 1st-order bounded parking trajectory near the reference relative state $\bar{\mathbf{x}} = [0, 500m, 0, 0, 0, 0]^T$.

a) Leader-Follower formation



b) Projected Circular Orbit formation

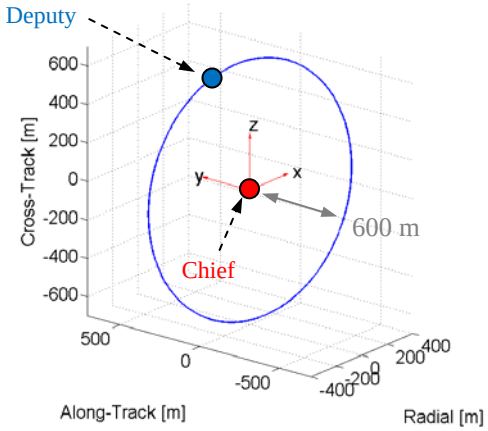


Fig. 3 (a) Leader-Follower formation (b) Projected Circular Orbit formation

Table 1 Initial relative dynamic state of the Deputy

$\mathbf{r}_i\ [m]$	$\mathbf{v}_i\ [m/s]$
$\begin{bmatrix} 5.657 \\ 51.314 \\ 8.000 \end{bmatrix}$	$\begin{bmatrix} 6.512 \\ -13.025 \\ 0.000 \end{bmatrix} \cdot 10^{-3}$

The following simulation results refer to a 10 minutes acquisition maneuver, with a control interval $T_s = 5\,s$ and a maximum control acceleration $u_{\max} = 5 \cdot 10^{-3} m/s^2$, so that:

$$|u_j| \le u_{\max}, \quad j = x, y, z \qquad (15)$$

The parameters of the LQR and MPC controllers are listed in Table 2 and Table 3, respectively. The parameters of both controllers are tuned in order that the components of the position error vector present the same settling time $ST \approx 2\,\text{min}$, i.e. the two controllers presents similar performances in reaching the same reference relative state.

Table 2 LQR parameters

${}^{LQR}\mathbf{Q}$	$\begin{bmatrix} 1.5 \cdot \mathbf{I}_{3x3} & \mathbf{0}_{3x3} \\ \mathbf{0}_{3x3} & 10^{-1} \cdot \mathbf{I}_{3x3} \end{bmatrix}$
${}^{LQR}\mathbf{R}$	$1.1 \cdot 10^{11} \cdot n^2 \cdot \mathbf{I}_{3x3}$

Table 3 MPC parameters

${}^{MPC}\mathbf{Q}$	${}^{LQR}\mathbf{Q}$
${}^{MPC}\mathbf{R}$	$1.25 \cdot 10^{9} \cdot n^2 \cdot \mathbf{I}_{3x3}$
control horizon m	4
prediction horizon p	20

The position error trajectory in Hill's LVLH reference frame is shown in Fig. 4, while the position error components are plotted in Fig. 5. In particular, for both controller types under investigation, the magnitude of the position error vector is less that 0.382 m after 2.00 min from the beginning of the maneuver.

Figure 6 shows the control acceleration components actually performed by the propulsion system $\mathrm{U}_{comm}(t)$ (see Fig. 1), in blue for the LQR controller and in red for the MPC controller. As can be seen from this figure, the control acceleration components of both controllers sometimes reach the maximum values $\pm 5 \cdot 10^{-3} m/s^2$. However, the MPC controller takes the control acceleration constraints into account while performing the optimal control computation, so that the output

of the MPC *CONTROL* module $U_{calc}(t)$ always satisfies the saturation constraint given by Eq. (15). On the other hand, the LQR optimal control computation is performed without taking into account the saturation constraint, and for this reason $U_{calc}(t)$ from the LQR *CONTROL* module can be greater than the maximum control acceleration that can be carried out by the propulsion system.

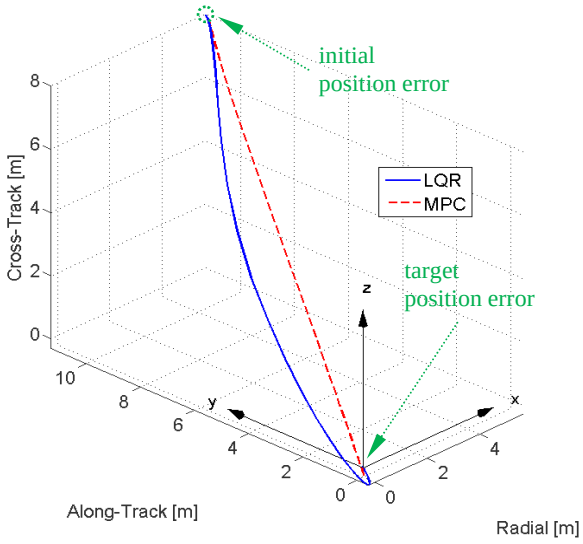


Fig. 4 3D position error trajectory

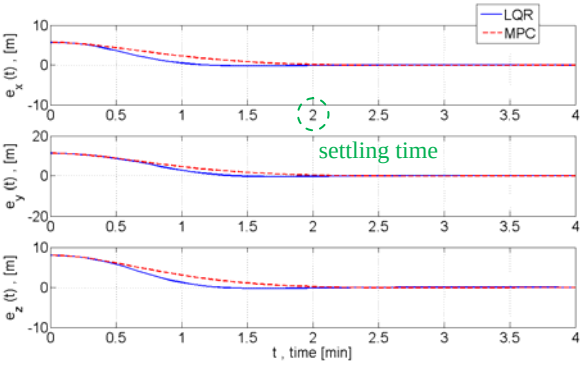


Fig. 5 Position error components

The required Δv is taken as index of the control system performance, defined as:

$$\Delta v(t) = \int_0^t \| \mathbf{u}(\tau) \| d\tau \qquad (16)$$

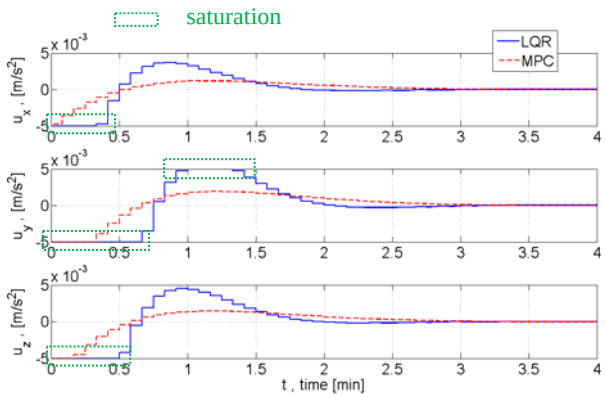
This index is used to compare the performances of the two controllers LQR and MPC. Table 4 presents LQR and MPC Δv requirement over 10 minutes of simulation and their percentage difference, which is defined as:

$$D_{\Delta v} = 100 \frac{\Delta v_{MPC} - \Delta v_{LQR}}{\Delta v_{LQR}} \qquad (17)$$

The profiles of the Δv requirement and the correspondent percentage differences are depicted in Fig. 7 and Fig. 8, respectively.

Table 4 Δv requirements of LQR and MPC controllers

Δv_{LQR} [m/s]	0.68187
Δv_{MPC} [m/s]	0.39995
$D_{\Delta v}$ [%]	– 41.34



Note. Steps are the results of the control frequency

Fig. 6 Control acceleration components of LQR and MPC controllers

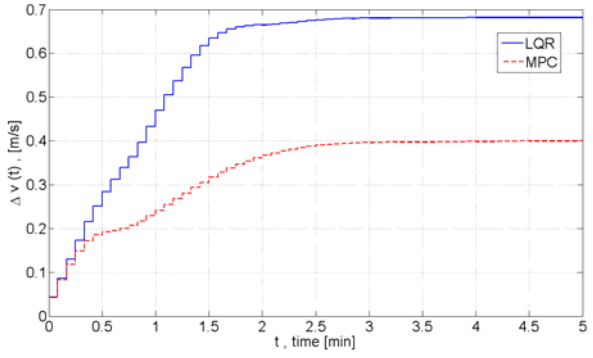


Fig. 7 Δv requirements of LQR and MPC controllers

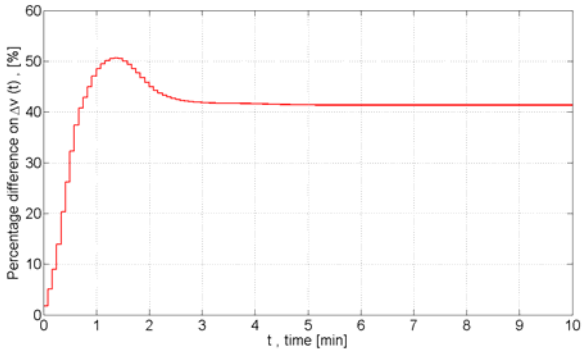


Fig. 8 Percentage difference on Δv requirement

The previous results are obtained assuming a dynamic evolution of the formation with no perturbations. In order to evaluate the effect of perturbations on the previous formation acquisition maneuver, other two simulations were carried out taking into account J_2, J_3 and drag perturbations.

The perturbative acceleration acting on a single spacecraft in the Earth Centered Inertial (ECI) reference frame is modeled as:

$$\mathbf{f}_{drag} = -\frac{1}{2}\rho\frac{C_D S}{M}V^2\hat{\mathbf{V}} \qquad (18)$$

where ρ is the air density, computed as function of the satellite altitude following a classical Jacchia's 1977 model for average solar activity ($\rho \approx 1.9\cdot 10^{-11}\, kg/m^3$ at the initial formation altitude), M and S are the mass and the cross section, respectively, of the spacecraft, C_D is the drag coefficient (equal to 2.2), $\mathbf{V}$ is the absolute velocity vector in the ECI frame. The values of the perturbative acceleration vector due to J_2 and J_3 are computed as reported, for example, in Ref. [9].

We assumed a value of $300\,\text{kg}$ for the mass of both spacecraft and a value of $5\,\text{m}^2$ and $10\,\text{m}^2$ for the cross sections of the Chief and the Deputy, respectively. The other simulation parameters used in the perturbed scenarios are the same as the one used in the unperturbed scenario.

Table 5 lists the Δv requirements of both LQR and MPC controllers for unperturbed and perturbed cases. The position error components profile and the control acceleration components profile are similar to those obtained in the unperturbed case. Also in the perturbed scenarios the use of an MPC controller can reduce significantly the total Δv requirement with respect to an LQR controller. However, the percentage difference $D_{\Delta v}$ between LQR and MPC Δv requirements for the case with J_2, J_3 and drag perturbations is reduced by 0.49% with respect to the unperturbed case.

Table 5 Δv requirements of LQR and MPC controllers for perturbed and unperturbed cases

perturbations	Δv_{LQR} [m/s]	Δv_{MPC} [m/s]	$D_{\Delta v}$ [%]
/	0.68187	0.39995	– 41.34
$J_2 + J_3$	0.68321	0.40118	– 41.28
$J_2 + J_3 + drag$	0.69126	0.40886	– 40.85

5 Projected Circular formation

In this section we compare the performances of the two controllers in carrying out a formation acquisition maneuver for a two spacecraft Projected Circular Orbit (PCO) formation (see Fig. 3b).

The Chief's initial orbital elements of the osculating keplerian orbit are as follows: $a = 6700\,km$, $e = 0$, $\Omega = 0\,\text{deg}$, $i = 97.87\,\text{deg}$, $\omega = 0\,\text{deg}$ and $\vartheta = 0\,\text{deg}$. The Deputy presents an initial relative dynamic state, listed in Table 6 and computed using Eq. (14) with $\rho_x = 295\,m$, $\rho_z = 590\,m$, $\Delta y = 0\,m$, $\varphi_{xy} = \varphi_z = 45\,\text{deg}$, while the reference relative trajectory is given by Eq. (14) with $\rho_x = 300\,m$, $\rho_z = 600\,m$, $\Delta y = 0\,m$, $\varphi_{xy} = \varphi_z = 45\,\text{deg}$, i.e. a PCO with a radius of $600\,m$ in the along-track – cross-track plane.

As in the Leader-Follower formation, we consider a 10 minutes maneuver, with a control interval $T_s = 5\,s$ and a maximum control acceleration $u_{\max} = 5\cdot 10^{-3}\, m/s^2$. The parameters of the LQR and the MPC controllers, listed in Table 7 and Table 8 respectively, are designed to obtain the same settling time $ST \approx 2\,\text{min}$ for the components of the position error vector.

Table 6 Initial relative dynamic state of the Deputy

$\mathbf{r}_i\ [m]$	$\mathbf{v}_i\ [m/s]$
$\begin{bmatrix} 208.597 \\ 417.193 \\ 417.193 \end{bmatrix}$	$\begin{bmatrix} 2.401 \\ -4.803 \\ 4.802 \end{bmatrix} \cdot 10^{-1}$

Table 7 LQR parameters

${}^{LQR}\mathbf{Q}$	$\begin{bmatrix} 1.5 \cdot \mathbf{I}_{3x3} & \mathbf{0}_{3x3} \\ \mathbf{0}_{3x3} & 10^{-1} \cdot \mathbf{I}_{3x3} \end{bmatrix}$
${}^{LQR}\mathbf{R}$	$1.1 \cdot 10^{11} \cdot n^2 \cdot \mathbf{I}_{3x3}$

Table 8 MPC parameters

${}^{MPC}\mathbf{Q}$	${}^{LQR}\mathbf{Q}$
${}^{MPC}\mathbf{R}$	$6.50 \cdot 10^{8} \cdot n^2 \cdot \mathbf{I}_{3x3}$
control horizon m	4
prediction horizon p	30

Table 9 Δv requirements of LQR and MPC controllers

Δv_{LQR} [m/s]	0.52151
Δv_{MPC} [m/s]	0.36522
$D_{\Delta v}$ [%]	– 29.97

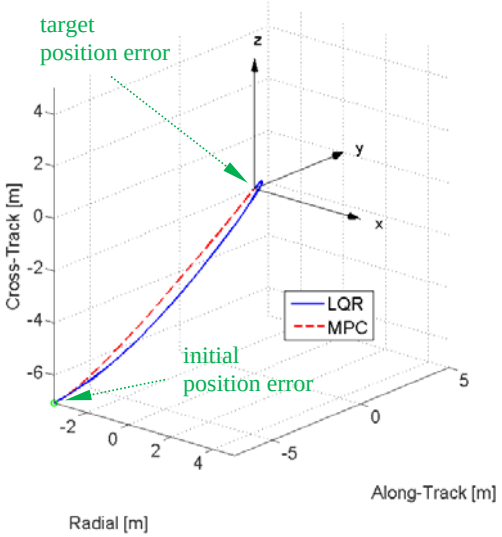


Fig. 9 3D position error trajectory

The position error trajectory in the Hill's LVLH reference frame is shown in Fig. 9, while the position error components are depicted in Fig. 10. Table 9 shows LQR and MPC Δv requirements over 10 minutes of simulation and their percentage differences. Figure 11 shows the control acceleration components produced by the propulsion system, while the profiles of the Δv requirement and the corresponding percentage differences are depicted in Fig. 12 and Fig. 13, respectively.

Also in this scenario, the use of a MPC controller can reduce significantly the total Δv requirement with respect to an LQR controller.

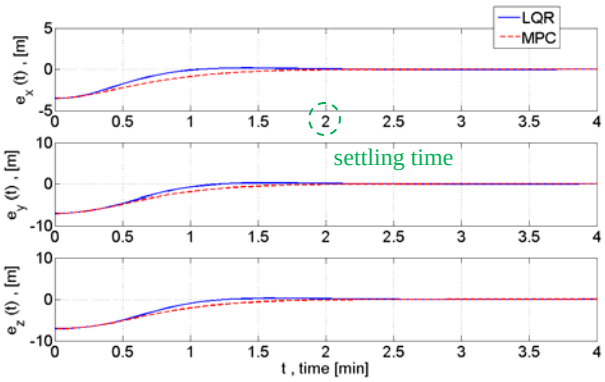


Fig. 10 Position error components

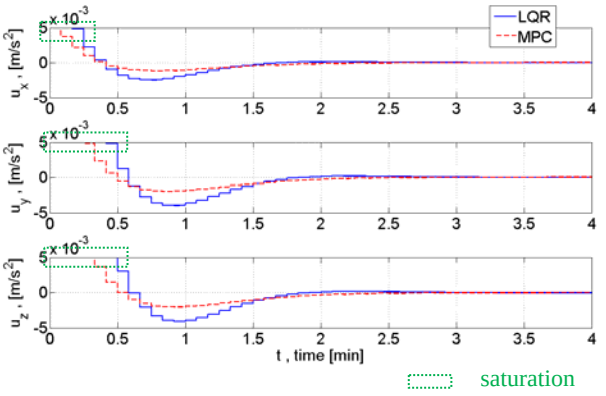


Fig. 11 Control acceleration components of LQR and MPC controllers

The effect of J_2, J_3 and drag perturbations on the PCO formation acquisition maneuver for both controllers are reported in Table 10. Simulation results also confirm that in the perturbed scenarios the MPC controller is more efficient (requires less Δv) than the LQR

controller, with a small reduction of $D_{\Delta v}$ (by 0.16%) between the unperturbed case and the one affected by $J_2 + J_3$ + drag perturbations.

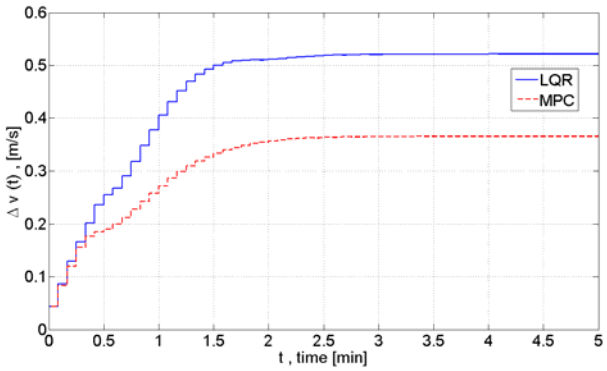


Fig. 12 Δv requirements of LQR and MPC controllers

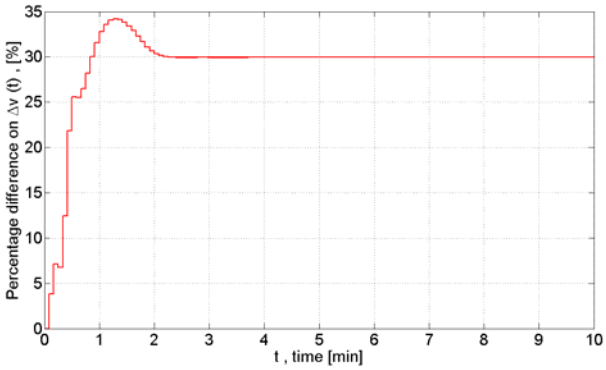


Fig. 13 Percentage difference on Δv requirement

Table 10 Δv requirements of LQR and MPC controllers for perturbed and unperturbed cases

perturbations	Δv_{LQR} [m/s]	Δv_{MPC} [m/s]	$D_{\Delta v}$ [%]
/	0.52151	0.36522	– 29.97
$J_2 + J_3$	0.52239	0.36607	– 29.92
$J_2 + J_3 + drag$	0.52906	0.37130	– 29.81

6 Conclusion

This paper analyzes and compares the performances of an MPC-based controller and a LQR-based one by carrying out a 10 minutes formation acquisition maneuver for two spacecraft flying in two formation configurations: a Leader-Follower and a Projected Circular Orbit in a perturbed orbital environment.

Simulation results show that in the presence of active constraints on the maximum control acceleration, the use of an MPC controller can reduce significantly the total Δv requirement with respect to an LQR controller. In particular, when the parameters of both controllers are tuned to have the same settling time for all position error components, we obtain a Δv reduction of about 40% and 30% for the Leader-Follower and the Projected Circular Orbit scenarios, respectively. This is because the LQR feedback strategy is determined *off-line* providing the optimal control for *all states* without considering actuator saturations, while MPC solves an optimal control problem *on-line* for the *current state* of the plant accounting for the control action constraints.

References

[1] Peláez J., Hedo J. M. and de Andréz P. R., "A special perturbation method in orbital dynamics", *Celestial Mechanics and Dynamical Astronomy*, Vol. 97, No. 2, 2007, pp. 131–150.

[2] Baù, G., Bombardelli, C., Peláez, J., Lorenzini, E., and Francesconi, A., "A variational method for the propagation of spacecraft relative motion", *20th AAS/AIAA Space Flight Mechanics Meeting*, Vol. 136 of Advances in the Astronautical Sciences, San Diego, CA, February 14-17 2010, pp. 2459–2474. Paper AAS 10-266.

[3] "Control System Toolbox™ 8". The MathWorks, Inc., 2009.

[4] Bemporad A., Morari M. and Ricker N. L., "Model Predictive Control Toolbox™ 3. User's Guide". The MathWorks, Inc., 2009.

[5] Alfriend K. T., Vadali S. R., Gurfil P., How J. P. and Breger L. S., *Spacecraft Formation Flying. Dynamics, control and navigation*, 1st ed., Butterworth-Heinemann, Elsevier Astrodynamics Series, 2010.

[6] Friedland B., *Control System Design: An introduction to State-Space Methods*, McGraw Hill, Inc., 1985.

[7] Maciejowski J. M., *Predictive Control with Constraints*, 1st ed., Prentice Hall, 2000.

[8] Mayne D. Q., Rawlings J. B., Rao C. V. and Scokaert P. O. M., "Constrained model predictive control: Stability and Optimality", *Automatica*, Vol. 36, No. 6, 2000, pp. 789-814.

[9] Schaub H. and Junkins J. L., *Analytical Mechanics of Space Systems*, 1st ed., AIAA Education Series, Reston, VA, 2003.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Guidance and Control Strategies for the Collision-Avoidance Mode of a Satellite Pair Flying in Formation

A. Valmorbida
Center of Studies and Activities for Space (CISAS) "G. Colombo"
Via Venezia 15, 35131 Padova, Italy

E. C. Lorenzini
Department of Mechanical Engineering, University of Padova
Via Venezia 1, 35131 Padova, Italy

G. Sechi
Thales Alenia Space Italia S.p.A.
Strada Antica di Collegno 253, 10146 Torino, Italy

Keywords: *guidance and control strategies, collision-avoidance, spacecraft formation flying.*

Abstract

In this paper we present a study on the Collision-Avoidance mode of a formation flying satellite pair in which a free-flying chief spacecraft follows temporary off-nominal conditions and a controlled deputy spacecraft performs collision avoidance maneuvers.

The Collision-Avoidance strategy consists in a Separation Guidance and a Nominal Guidance. The Separation Guidance is in charge of the avoidance of a predicted collision soon to occur, and it is based on a computationally simple, deterministic and closed-form algorithm, so that a valid solution is always available without delay. The Nominal Guidance uses Genetic Algorithms and it takes charge of placing the deputy spacecraft in a bounded safe or "parking" trajectory, while minimizing the propellant consumption and avoiding the evaporation of the formation.

The performances of a Linear Quadratic Regulator and a Model Predictive Control are also compared in tracking the reference trajectories provided by the Collision-Avoidance strategy.

1 Introduction

Formation flying missions that involve the coordinated control of several autonomous vehicles have generated increased interest in later years. This is due to the many inherent advantages that the distributed design adds to a mission.

However, the use of two or more spacecraft flying in formation require high levels of on-orbit autonomy and pose several control challenges. In formation flight, space vehicles operate in close proximity to accomplish common goals, and this poses a great importance on Guidance, Navigation and Control (GNC) system, whose requirements include: use of new sensors and actuators able respectively to measure and maintain relative position and/or velocity with high accuracy; critical dependence on fuel minimization to ensure mission feasibility; the need to prevent relative drift between vehicles; use of relative orbital dynamics for the control design; and reliable collision-avoidance strategies to avoid possible collisions between vehicles in temporary off-nominal conditions.

In this paper, we present a study on the Collision-Avoidance mode of a satellite pair flying in formation. The prime objective of the Collision-Avoidance mode is to ensure the immediate and long-term safety of the formation, in terms of avoiding collision, preventing formation evaporation and, as a secondary objective, saving fuel.

2 Collision-Avoidance Strategies

We define as Chief the free-flying spacecraft following temporary off-nominal conditions and as Deputy the spacecraft performing collision avoidance maneuvers. We also define two ellipsoidal regions centered at the Chief: the Avoidance Region (AR), with semi-axes a_{AR}, $b_{AR} = 2a_{AR}$ and $c_{AR} = a_{AR}$ (a 1x2x1 ellipsoid) in x, y and z directions of the Hill's reference frame, respectively, and the Nominal Boundary (NB) with semi-axes $a_{NB} = k\,a_{AR}$, $b_{NB} = k\,b_{AR}$ and $c_{NB} = k\,c_{AR}$, where k is a safety margin.

The Collision-Avoidance strategy is based on a Separation Guidance and on a Nominal Guidance[1],[2], summarized in the following (see Fig. 1). If a collision is predicted and the estimated relative position between the Deputy and the Chief is within the Avoidance Region, then a separation maneuver is performed (Separation Guidance): the distance between the two spacecrafts has to increase monotonically and the Deputy has to exit the Nominal Boundary within a prescribed time. When the Deputy reaches a relative position outside the Nominal Boundary, an evasive maneuver is planned (Nominal Guidance), allowing the Deputy to reach a safe or "parking" trajectory: the evasive trajectory, from the actual position to the safe trajectory, must not intersect the Nominal Boundary, and the safe trajectory must remain inside a prescribed Safe Region that does not intersect the Nominal Boundary and is far enough from the Chief. A Nominal Guidance maneuver can also be planned for cases in which the sensed relative position is already outside the Nominal Boundary while the Collision-Avoidance mode is activated only when a collision is predicted.

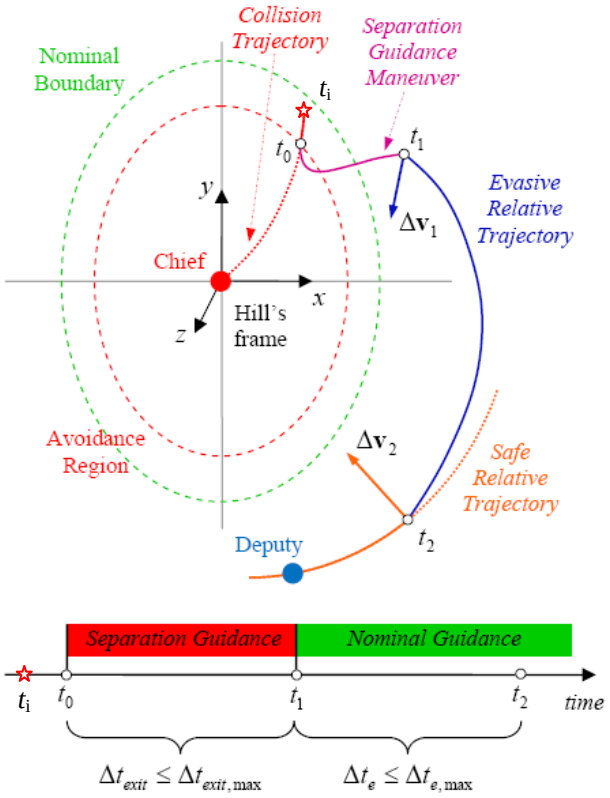


Fig. 1 Collision Avoidance strategy and time sequence

3 Separation Guidance

The Separation Guidance is more critical for the formation safety than the Nominal Guidance, and a reliable solution must be obtained with minimal computation time. For this reason, the Separation Guidance algorithm is a computationally simple, deterministic and closed-form algorithm, so that a valid solution is always available without delay. The algorithm must also be robust to practical levels of uncertainty in the initial estimated relative position and velocity used to plan the maneuver. The actual control action computation is indeed performed taking into account only the distance of the Deputy from the Nominal Boundary and the maximum time needed for the Deputy to exit the Nominal Boundary, $\Delta t_{exit,\max}$. $\Delta t_{exit,\max}$ is a tunable parameter: smaller exit times lead to higher delta-v and therefore higher fuel consumption, but a safe condition is reached in a shorter time.

We define Current Ellipsoid (CE) as a 1x2x1 ellipsoid centered at the Chief, with x semi-axis

a_k and passing through the current relative position of the Deputy, $\mathbf{r}_k$ (Fig. 2).

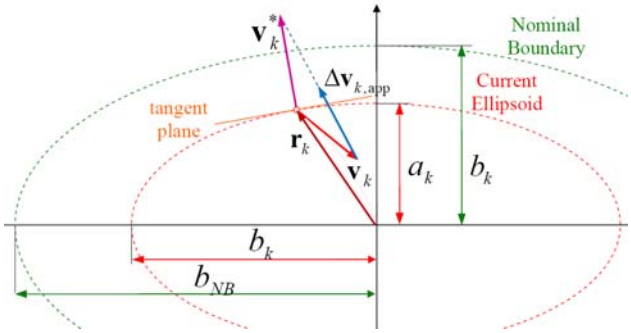


Fig. 2 Separation Guidance strategy

A desired current separation velocity $\mathbf{v}_k^*$ is first computed based only on the current relative position vector $\mathbf{r}_k$. The magnitude of $\mathbf{v}_k^*$ is proportional to the difference between the semi-axes in y direction of the Nominal Boundary, b_{NB}, and the Current Ellipsoid, b_k, and inversely proportional to the current separation time Δt_k:

$$v_k^* = \frac{b_{NB} - b_k}{\Delta t_k} \qquad (1)$$

where $\Delta t_k = t_f - t_k$, with t_k the current time and t_f the time within which the Deputy must be outside the Nominal Boundary. $\mathbf{v}_k^*$ is aligned with the external unit vector orthogonal to the plane that is tangent in $\mathbf{r}_k$ to the Current Ellipsoid. The current delta-v vector $\Delta\mathbf{v}_k$ is:

$$\Delta\mathbf{v}_k = \mathbf{v}_k^* - \mathbf{v}_k \qquad (2)$$

The thruster system may not be able to perform the current delta-v vector, and the current control acceleration vector $\mathbf{u}_k$ is therefore computed as:

$$\mathbf{u}_k = \frac{1}{T_s}\min(\mathbf{u}_{\max}, \Delta\mathbf{v}_k) \qquad (3)$$

where T_s is the control interval of the separation control, and $\mathbf{u}_{\max}$ is the maximum acceleration that the thruster system can apply in the $\Delta\mathbf{v}_k$ direction.

4 Nominal Guidance

The Nominal Guidance starts if the sensed relative position between the Deputy and the Chief is outside the Nominal Boundary and the Collision-Avoidance mode is activated. The Nominal Guidance algorithm compute two basic relative trajectories (see Fig. 1): an Evasive Relative Trajectory (ERT) and a Safe Relative Trajectory (SRT). A first delta-v $\Delta\mathbf{v}_1$ is planned at time t_1 so that the Deputy follows the ERT outside the Nominal Boundary; then, a second delta-v $\Delta\mathbf{v}_2$ is planned at time t_2 and the Deputy finally achieve the SRT.

The couple of trajectories SRT-ERT has to satisfy the following safe requirements: (a) the SRT has to evolve within a Safe Region (section 4.1); (b) the minimum distance of the SRT projected in the cross-track plane has to be greater than a safe minimum distance (section 4.2); and (c) the ERT has to be outside the Nominal Boundary (section 4.3). Furthermore, the whole maneuver has to be performed within a maximum time interval $\Delta t_{e,\max}$, requirement (d), and the total delta-v Δv_{tot} used to plan the Nominal Guidance maneuver has to be minimized.

As discussed in the following sections, each of the previous safe requirements is converted in a correspondent constraint condition for those parameters that define the couple of trajectories SRT-ERT.

4.1 Safe Relative Trajectories definition

The Safe Region is designed with the double intent to assure the formation safety, keeping the two spacecrafts sufficiently far apart, and to prevent formation evaporation, keeping the spacecraft sufficiently close for the communication and relative navigation systems to operate.

Two types of Safe Regions are designed taking into account both in-plane and out-of-plane safe constraints: a Fly-Around type, which

defines SRTs that evolve around the Chief, and a No-Fly-Around type, that does not fly around the Chief and whose geometric center is properly placed along track.

In order to reduce both the computational complexity of the Nominal Guidance algorithm and the propellant consumption, the previous two SRT types are designed using the Hill-Clohessey-Wiltshire (HCW) solution with the 1st order bounded relative motion condition $\dot{y}_0 = -2nx_0$. The HCW solution in the Hill's Local-Vertical Local-Horizontal (LVLH) rotating cartesian reference frame (x radial, y along-track and z cross-track), is:

$$\begin{cases} x(t) = \rho_x \sin(nt + \varphi_{xy}) \\ y(t) = \rho_y \cos(nt + \varphi_{xy}) + \Delta y \\ z(t) = \rho_z \sin(nt + \varphi_z) \end{cases} \quad (4)$$

with $\rho_y = 2\rho_x$. In the previous equations, ρ_x and ρ_z define the dimension of the in-plane and out-of-plane bounded motion, Δy is the along-track offset, φ_{xy} and φ_z are the in-plane and out-of-plane phase angles, respectively, and n is the Chief mean motion.

The value of the five independent parameters ρ_y, ρ_z, Δy, φ_{xy} and φ_z have to be properly selected in order for each SRT to evolve inside the corresponding Safe Region.

4.1.1 Fly-Around

The Fly-Around Safe Region is a 1x2x1 elliptical section cylinder with a central empty region (Fig. 3).

The in-plane projection of the Safe Region and the constraint conditions (c/c) for Fly-Around SRTs are presented in Fig. 4. A Fly-Around SRT evolves within the Fly-Around Safe Region if the values of ρ_y and Δy satisfy the following two constraint conditions (c/c 1 and c/c 2):

$$\begin{cases} y_{\min} \geq y^*_{\min} \\ y_{\min} = \rho_y - \Delta y \end{cases} \Rightarrow \rho_y - \Delta y \geq y^*_{\min}; \text{ c/c 1} \quad (5)$$

$$\begin{cases} y_{\max} \leq y^*_{\max} \\ y_{\max} = \rho_y + \Delta y \end{cases} \Rightarrow \rho_y + \Delta y \leq y^*_{\max}; \text{ c/c 2} \quad (6)$$

For what concern the out-of-plane evolution of the SRT, a constraint on the maximum amplitude in this direction is considered that translates into:

$$\rho_z \in [0, z^*_{\max}] \quad (7)$$

where $z^*_{\max}$ is the maximum amplitude of the out-of-plane motion.

The in-plane and out-of-plane phase angles φ_{xy} and φ_z can assume a value between 0 and 2π.

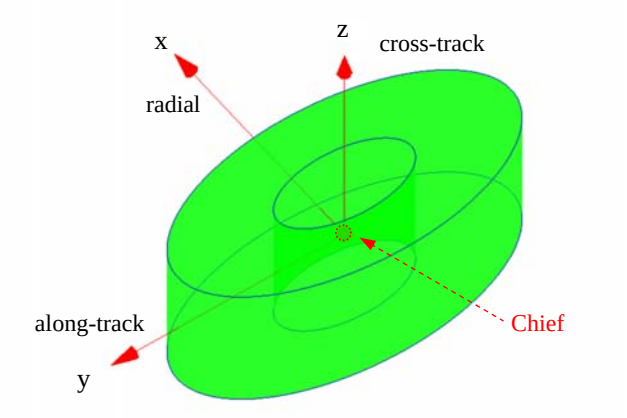


Fig. 3 3D visualization of the Fly-Around Safe Region

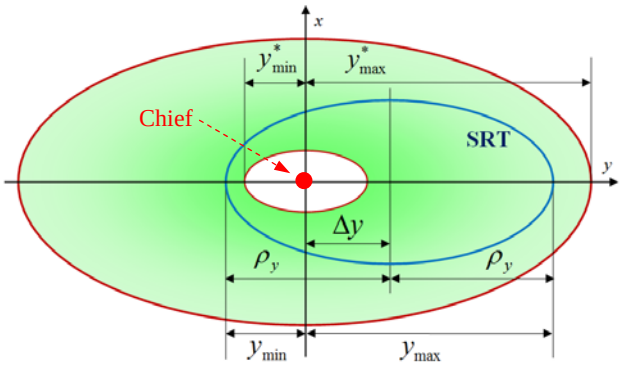


Fig. 4 Constraint conditions for the In-plane motion of the Fly-Around SRTs

4.1.2 No-Fly-Around

The No-Fly-Around Safe Region consists of two parallelograms with zero offset in radial and

cross-track directions and with equal and opposite offset in along-track direction (Fig. 5).

Figure 6 represents the in-plane projection of the Safe Region and the constraint conditions for No-Fly-Around SRTs. A No-Fly-Around SRT evolves within the No-Fly-Around Safe Region if the following two constraint conditions, c/c 3 and c/c 4, are satisfied by ρ_y and Δy:

$$\begin{cases} y_{\min} \geq y^*_{\min} \\ y_{\min} = \Delta y - \rho_y \end{cases} \Rightarrow \Delta y - \rho_y \geq y^*_{\min};\ \text{c/c 3} \qquad (8)$$

$$\begin{cases} y_{\max} \leq y^*_{\max} \\ y_{\max} = \Delta y + \rho_y \end{cases} \Rightarrow \Delta y + \rho_y \leq y^*_{\max};\ \text{c/c 4} \qquad (9)$$

The amplitude of the out-of-plane motion can get any value between 0 and the maximum value $z^*_{\max}$.

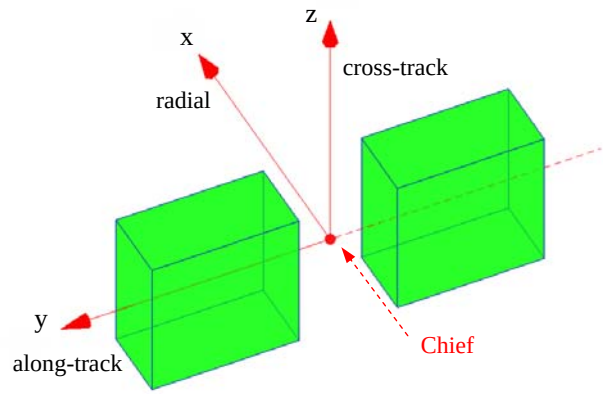


Fig. 5 3D visualization of the No-Fly-Around Safe Region

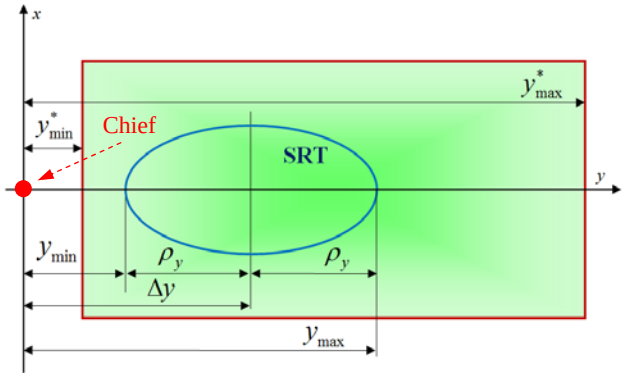


Fig. 6 Constraint conditions for the In-plane motion of the No-Fly-Around SRTs

4.2 Cross-track minimum distance

In some scenarios it is essential to ensure a minimum distance in the cross-track plane to guarantee the safety of the formation flying satellites, even in the presence of along-track drift. Some small amount of along-track drift is always present after a maneuver due to unavoidable thruster inaccuracies, navigation errors and perturbations.

We define cross-track avoidance region as a safe circle in the cross-track plane, centered at the Chief and with radius r_{CT}. The minimum distance of the SRT projected in the cross-track plane is greater than the safe minimum distance r_{CT}, requirement (b), if the semi-minor axis b of the ellipse obtained by projecting the SRT in the cross-track plane is greater than r_{CT}, that is:

$$b = \sqrt{\frac{\rho_x^2 + \rho_z^2 - \sqrt{\rho_x^4 + \rho_z^4 - 2\rho_x^2\rho_z^2\cos 2\alpha}}{2}} \geq r_{CT};\ \text{c/c 5} \qquad (10)$$

where $\alpha = \varphi_{xy} - \varphi_z$.

4.3 Evasive trajectories constraints

To check if the ERT is outside the Nominal Boundary, the ERT is divided into m points $P_{ERT,j}$, with $j = 1,\ldots,m$. Let $d_{ERT,j}$ be the distance of the point $P_{ERT,j}$ from the origin of the Hill's frame, and let $d_{NB,j}$ be the distance from the frame origin of the point $P_{NB,j}$ obtained intersecting the ray from the origin to $P_{ERT,j}$ with the Nominal Boundary (see Fig. 7).

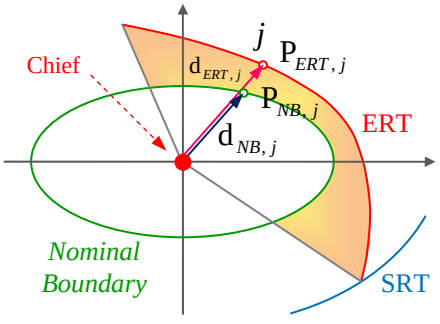


Fig. 7 Evasive Relative Trajectory constraints

The ERT evolve outside the Nominal Boundary if the following constraint condition c/c 6 is satisfied:

$$\min_{j=1,\ldots,m}\left(d_{ERT,j}-d_{NB,j}\right)\geq 0;\quad \text{c/c 6} \qquad (11)$$

4.4 Nominal Guidance Algorithm

The Nominal Guidance algorithm uses Genetic Algorithms[5] (GAs) to find a ERT-SRT couple that satisfies (a) (b) (c) and (d) requirements for both Fly-Around and Not-Fly-Around SRTs. The optimization variables used in GAs are ρ_y, ρ_z, Δy, φ_{xy}, φ_z and Δt_e, whose values are chosen in the following ranges:

- $\rho_y \in [-y^*_{\min}, y^*_{\min}]$;
- $\rho_z \in [0, z^*_{\max}]$;
- $\Delta y \in [-\Delta y^*, \Delta y^*]$ for the Fly-Around case;
- $\Delta y \in [y^*_{\min}, y^*_{\max}]$ or $\Delta y \in [-y^*_{\min}, -y^*_{\max}]$ for the No-Fly-Around case;
- $\varphi_{xy} \in [0, 2\pi]$;
- $\varphi_z \in [0, 2\pi]$;
- $\Delta t_e \in [0, \Delta t_{e,\max}]$.

where $\Delta y^* = \dfrac{y^*_{\max} - y^*_{\min}}{2}$.

Parameters $y^*_{\min}$, $y^*_{\max}$, $z^*_{\max}$, which define dimensions of the Safe Region in the two cases Fly-Around and No-Fly-Around, $\Delta t_{e,\max}$, Fly-Around or No-Fly-Around SRT type and the constraint on the minimum distance of the SRT in the cross-track plane are all elements that have to be properly chosen in accordance with a specific mission scenario.

GAs minimize Δv_{tot} while satisfying the constraint conditions 1-2 for the Fly-Around case or 3-4 for the No-Fly-Around case, 5 and 6. The total delta-v is computed using the HCW model and assuming impulsive maneuvers, with:

$$\Delta v_{tot} = \|\Delta \mathbf{v}_1\| + \|\Delta \mathbf{v}_2\| \qquad (12)$$

To check if each constrain condition is verified, a corresponding penalization term is properly added to Δv_{tot}. Each penalization term is equal to zero only if the corresponding constraint condition is satisfied; on the other hand, it exhibits an exponential rise. Each constraint condition can indeed be rewritten in the form $\mu_i \geq \mu_i^*$, with $i = 1,\ldots,6$. If the ith constraint condition is not verified, then $\mu_i < \mu_i^*$ and the corresponding penalization term is:

$$p_i = e^{k_i\left(1-\frac{\mu_i}{\mu_i^*}\right)} - 1 \qquad (13)$$

where $k_i = \dfrac{\ln(\gamma_i + 1)}{1-\alpha_i}$. This means that if $\mu_i = \alpha_i \mu_i^*$, with $\alpha_i \in [0,1)$, then $p_i = \gamma_i$. k_i is introduced to control the rising rate of each penalization term in relation to the other ones, and it can also be used to weight one constraint violation more than the other ones.

In GAs, both the problem formulation and the genetic parameter selection, such as the population size or the cross-over function, have a big effect on the solution obtained with these methods. Smaller size of population looses the diversity very soon, while higher population size provides a higher diversity and the convergence to better solutions happens sooner than smaller population sizes. Bigger populations needs however more time for the algorithm to run, which should be a problem in collision-avoidance maneuvers. As a consequence, the population size must be a trade off solution between better results and computational time.

Once a solution is obtained with GAs, a Pattern Search (PS) algorithm is run with GA solution as starting point, in order to refine the GAs solution.

5 Control Strategies

The Separation Guidance and the Nominal Guidance strategies has been tested using a Matlab software simulator for spacecraft formation flying applications with a modular structure. The Collision-Avoidance algorithms have been used to develop the following

modules: the Collision Avoidance Separation Guidance (CASG) module, the Collision Avoidance Nominal Guidance (CANG) module and the section of the Mode Manager module that is in charge of the Collision-Avoidance.

Referring to Fig. 8, if the Collision-Avoidance mode is activated, the Mode Manager handles activation/deactivation of the CASG and CANG modules through S_{CASG}^{MM} and S_{CASG}^{MM} commands. When the CASG module is enabled, it computes the control actions that the thrusters system has to execute, based on the estimation of the actual relative position $\hat{\mathbf{x}}$. When the CANG is enabled, it carries out the computation of the reference state trajectory $(\bar{\mathbf{x}},\bar{\mathbf{y}})$ starting from the current estimated relative dynamic state $(\hat{\mathbf{x}},\hat{\mathbf{y}})$. The reference trajectory is utilized in the control module to compute the control acceleration for trajectory tracking.

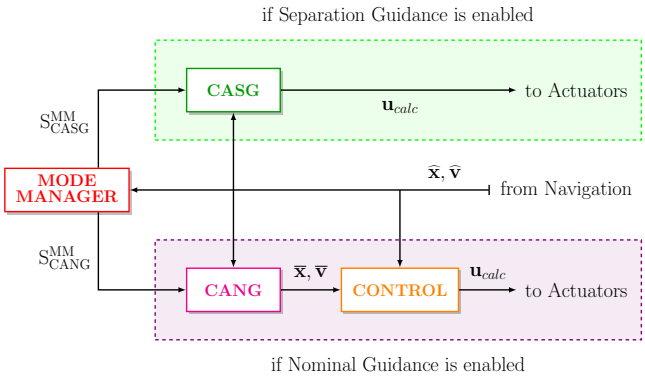


Fig. 8 Guidance and Control scheme for the Collision- Avoidance mode

Two types of control strategies are designed to perform the tracking of the reference trajectories: the first one is based on Linear Quadratic Regulator[3],[6] (LQR), the second one is based on Model Predictive Control[4],[7] (MPC).

Both controllers use a state feedback in the optimal control law computation. However, for the problem under investigation, while the LQR control law can be computed off-line and considering control constraints *a posteriori*, the MPC control law is obtained by solving on-line, at each control step, a finite horizon open-loop optimal control problem that takes into account control and state constraints.

In the synthesis of both controllers, we used the discrete time version of the HCW relative dynamic model in state-space form, that is, a Discrete-time Linear Time Invariant (DLTI) model:

$$\begin{cases} \mathbf{X}_{k+1} = \mathbf{A}\,\mathbf{X}_k + \mathbf{B}\,\mathbf{u}_k \\ \mathbf{y}_k = \mathbf{x}_k \end{cases} \qquad (14)$$

where $\mathbf{X} = [\mathbf{x}^T, \mathbf{v}^T]^T$ is the state vector and $\mathbf{y}$ is the output vector. Furthermore, the trajectory tracking problem is seen in the light of a regulation problem for the instantaneous deviation between the current estimated and the desired trajectories.

The purpose of LQR synthesis is to find a state feedback control law in the form $\mathbf{u}_k = -\mathbf{K}\,(\mathbf{X}_k - \overline{\mathbf{X}}_k)$, where $\mathbf{K}$ is a gain matrix obtained from the solution of the Discrete-time Algebraic Riccati Equation (DARE), in order to minimize the following quadratic performance index:

$$J = \sum_{k=0}^{\infty} [(\mathbf{y}_k - \mathbf{r}_k)^T \mathbf{Q} (\mathbf{y}_k - \mathbf{r}_k) + \mathbf{u}_k^T \mathbf{R}\,\mathbf{u}_k] \qquad (15)$$

where $\mathbf{r}$ is the output reference vector, and $\mathbf{Q}$ and $\mathbf{R}$ are positive-definite matrices representing the state error weighting matrix and the control weighting matrix, respectively.

Matrices $\mathbf{Q}$ and $\mathbf{R}$ are parameters that affect reference trajectory tracking performances and fuel consumption: higher values of $\mathbf{Q}$ elements with respect to $\mathbf{R}$ elements means higher weight on state error with respect to control action, and this lead to a more aggressive control, a better performance in tracking the desired trajectory but a higher fuel consumption.

The MPC control action at time k is obtained by minimizing the following performance index with respect to the sequence of the input increments $\Delta u(k\,|\,k)$, ... , $\Delta u(m-1+k\,|\,k)$:

$$J_k = \sum_{i=0}^{p-1} \left\| \mathbf{y}(k+i+1\,|\,k) - \mathbf{r}(k+i+1) \right\|_{\mathbf{Q}(i)}^2 + \\ + \left\| \Delta\mathbf{u}(k+i\,|\,k) \right\|_{\mathbf{R}_{\Delta u}(i)}^2 + \left\| \mathbf{u}(k+i\,|\,k) - \mathbf{u}_{\text{target}}(k+i) \right\|_{\mathbf{R}_u(i)}^2 \qquad (16)$$

subject to the following constraints:

$$
\begin{aligned}
&\mathbf{u}_{\min}(i) \le \mathbf{u}(k+i \mid k) \le \mathbf{u}_{\max}(i) \\
&\Delta\mathbf{u}_{\min}(i) \le \Delta\mathbf{u}(k+i \mid k) \le \Delta\mathbf{u}_{\max}(i) \\
&\mathbf{y}_{\min}(i) \le \mathbf{y}(k+i \mid k) \le \mathbf{y}_{\max}(i) \\
&\Delta\mathbf{u}(k+h \mid k) = \mathbf{0} \\
&i = 0,\ldots, p-1 \\
&h = m,\ldots, p-1
\end{aligned}
\qquad (17)
$$

where '$(k+i \mid k)$' denotes the value predicted for time $k+i$ based on the information available at time k, p is the prediction horizon, m is the control horizon and $\mathbf{Q}$, $\mathbf{R}_{\Delta u}$ and $\mathbf{R}_u$ are output error, input increment and input weighting matrices, respectively.

When designing a MPC controller one should take into account that: a more aggressive controller with more fuel consumption and better tracking performances is obtained with smaller prediction horizon *p*, tighter constraints on control variables and higher ratio between state error weight and control action weight; larger control horizon *m* leads to an optimization problem with more degrees of freedom, thus a "more" optimal (better performances) but more complex control.

6 Simulation Results

In this section we present results of a Collision-Avoidance maneuver in which both a Separation Guidance maneuver and a Nominal Guidance maneuver are performed.

The initial dynamic state of the Chief is given in terms of orbital elements of the osculating keplerian orbit: semi-major axis $a = 6778\,km$, eccentricity $e = 0$, right ascension of the ascending node $\Omega = 30\,\text{deg}$, inclination $i = 45\,\text{deg}$, argument of perigee $\omega = 20\,\text{deg}$ and true anomaly $\vartheta = 60\,\text{deg}$. The initial dynamic state of the Deputy is given in Table 1 as relative position $\mathbf{r}_i$ and velocity $\mathbf{v}_i$ vectors with respect to the Hill's frame. In order to simulate a collision, the value of $\mathbf{v}_i$ is obtained forcing the Deputy to reach the origin of the Hill's frame in 5 minutes. Separation Guidance parameters are listed in Table 2, while Nominal Guidance parameters are listed in Table 3. $\Delta t_{e,\max} = 69.56\,\text{min}$ is equal to 0.75 times the initial orbital period of the Chief.

Table 1 Initial relative dynamic state of the Deputy

initial relative position $\mathbf{r}_i$ [m]	initial relative velocity $\mathbf{v}_i$ $\cdot 10^{-3}$ [m/s]
$[120.000,\ 50.000,\ -60.000]^T$	$[-374.761,\ -300.006,\ 192.260]^T$

Table 2 Separation Guidance parameters

a_{AR} [m]	a_{NB} [m]	$\Delta t_{exit\,max}$ [min]	T_s [s]	u_{max} [m/s^2]
100.00	200.00	7.00	5.00	3.00e-002

Table 3 Nominal Guidance parameters

$\Delta t_{e,max}$ [min]	y^*_{min} [m]	y^*_{max} [m]	z^*_{max} [m]	Safe Region type	r_{CT} [m]
69.56	450.00	1000.00	700.00	Fly-Around	200.00

Figure 9 shows the profile of the control acceleration components during the Separation Guidance maneuver. Thrusters are switched on for the first 24 seconds of the maneuver to ensure that the Deputy is outside the Nominal Boundary within 7 minutes. Figure 10 shows the Deputy/Chief distance profile: the Deputy is outside the Avoidance Region (AR) and the Nominal Boundary (NB) after 34 seconds and 6.37 minutes, respectively.

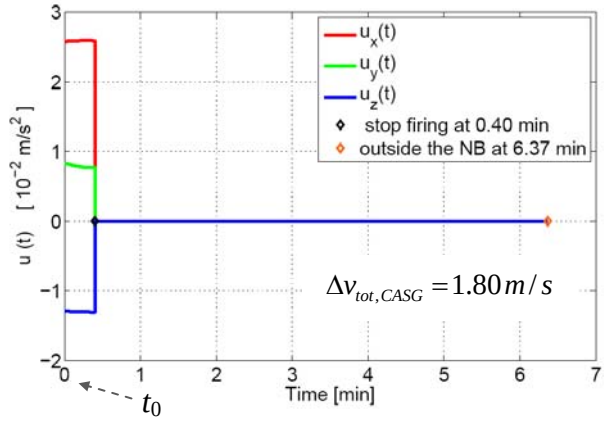


Fig. 9 Control acceleration components – Separation Guidance maneuver

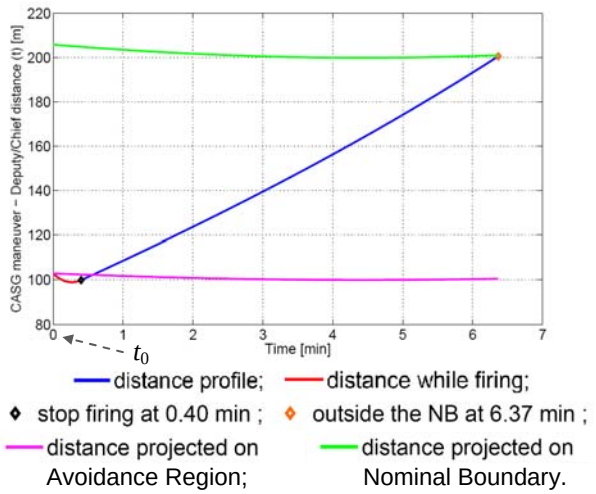


Fig. 10 Distance profiles – Separation Guidance maneuver

Results of Genetic Algorithm and Pattern Search for the Nominal Guidance maneuver are presented in Table 4. These results are obtained using Matlab Genetic Algorithm and Direct Search Toolbox™, with a population size of 4000 individuals and with a maximum computational time for GA and PS algorithms of 1.5 minutes. The corresponding position, velocity and delta-v vectors are listed in Table 5.

Figure 11 and Figure 12 show in-plane and out-of-plane projection, respectively, of the Fly-Around Safe Region and the resulting couple of trajectories SRT-ERT. As can be noted from these figures, all constraint conditions are satisfied.

Table 4 Genetic Algorithms and Pattern Search results – Nominal Guidance maneuver

	ρ_y [m]	ρ_z [m]	Δy [m]	φ_{xy} [rad]
GA	711.300	201.946	−260.883	5.4254
PS	711.300	201.946	−261.300	5.3896
difference [%]	0.00	−0.00	0.16	−0.66

	φ_z [rad]	Δt_e [s]	Δv_{tot} [m/s]
GA	3.7412	3594.86	$3.5593e-001$
PS	3.7040	3620.87	$3.5495e-001$
difference [%]	−0.99	0.72	−0.27

Table 5 Position, velocity and delta-v vectors – Nominal Guidance maneuver

t	$\mathbf{r}(t)$ [m]	$\mathbf{v}(t^-)$ $\cdot 10^{-3}$ [m/s]	$\mathbf{v}(t^+)$ $\cdot 10^{-3}$ [m/s]	$\Delta\mathbf{v}(t)$ $\cdot 10^{-3}$ [m/s]	$\|\Delta\mathbf{v}(t)\|$ $\cdot 10^{-3}$ [m/s]
t_1	$\begin{bmatrix} 183.553 \\ -25.470 \\ -78.728 \end{bmatrix}$	$\begin{bmatrix} 299.887 \\ -255.258 \\ -83.595 \end{bmatrix}$	$\begin{bmatrix} 270.444 \\ -434.602 \\ 49.100 \end{bmatrix}$	$\begin{bmatrix} -29.443 \\ -179.343 \\ 132.696 \end{bmatrix}$	225.031
t_2	$\begin{bmatrix} -354.900 \\ -215.125 \\ 10.052 \end{bmatrix}$	$\begin{bmatrix} 44.747 \\ 783.773 \\ -101.069 \end{bmatrix}$	$\begin{bmatrix} 26.120 \\ 803.043 \\ -228.191 \end{bmatrix}$	$\begin{bmatrix} -18.627 \\ 19.270 \\ -127.122 \end{bmatrix}$	129.917

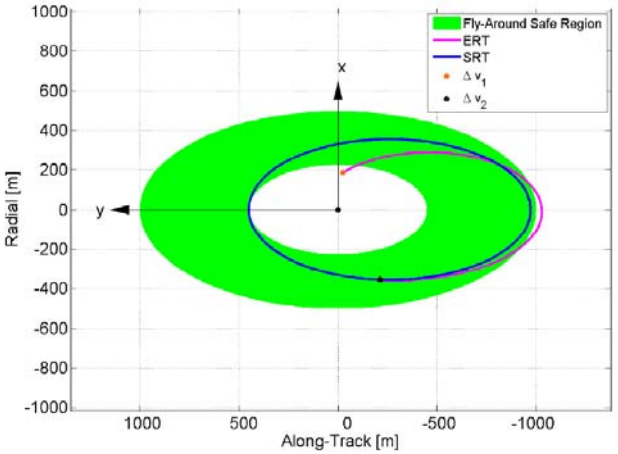


Fig. 11 In-plane projection of the Nominal Guidance maneuver

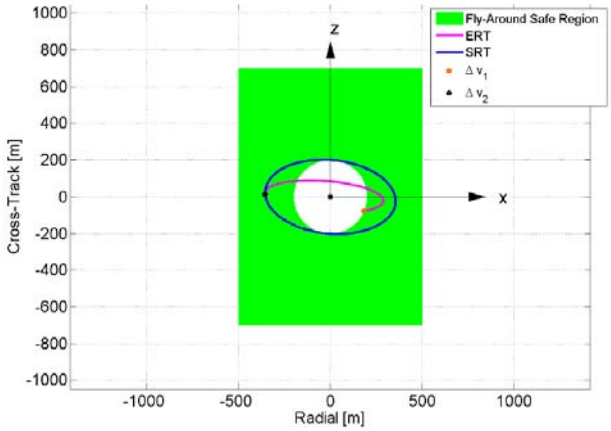


Fig. 12 Out-of-plane projection of the Nominal Guidance maneuver

A comparison of LQR and MPC control acceleration components near times t_1 and t_2 is shown in Fig. 13 and Fig. 14, respectively. These results are obtained with a control horizon $m = 5$, a prediction horizon $p = 25$, a control interval $T_s = 5s$ for both controllers, while the state error and the control weighting

matrices of both controllers are set to obtain similar performances in tracking the same reference trajectories. For the Nominal Guidance maneuver, the total delta-v obtained with LQR and MPC are $\Delta v_{LQR} = 1.4757\, m/s$ and $\Delta v_{MPC} = 1.3914\, m/s$, respectively. In this scenario, the use of an MPC controller reduces the total delta-v by 5.71% with respect to an LQR controller. This is due also to the anticipative action of the MPC, as can be noted in Fig. 14.

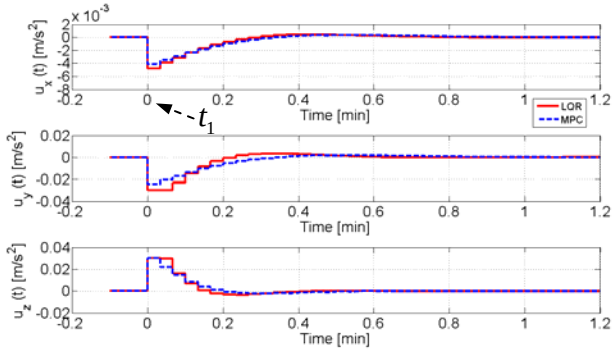


Fig. 13 LQR and MPC control acceleration components near time t_1 (LQR in red and MPC in blue)

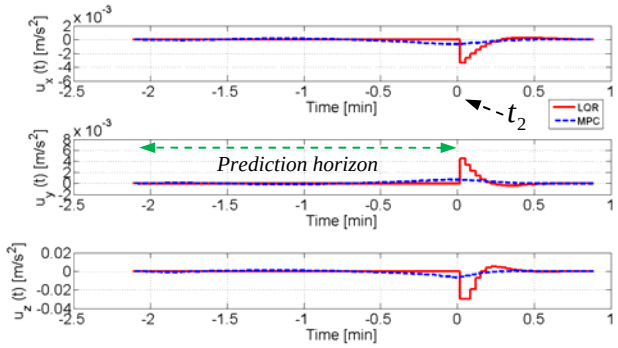


Fig. 14 LQR and MPC control acceleration components near time t_2 (LQR in red and MPC in blue)

7 Conclusions

In this paper we present some collision-avoidance guidance and control strategies for a satellite pair flying in formation. The collision-avoidance strategy consists on a first Separation Guidance maneuver performed to avoid a predicted oncoming collision and on a second Nominal Guidance maneuver that is in charge of placing the deputy spacecraft in a bounded parking trajectory that must fly within a prescribed safe region.

We investigate two types of safe region: a fly-around safe region that evolve around the chief spacecraft and a no-fly-around safe region whose geometric center is properly placed along track. For each of them, we analyze the requirements that the Nominal Guidance maneuver must satisfy to assure the safety of the formation. Furthermore, we describe how Genetic Algorithms can be setup and used to plan a Nominal Guidance maneuver that satisfies the specified safe requirements. In order to track the desired reference trajectories provided by the Nominal Guidance, we also design two types of controllers, a Linear Quadratic Regulator and a Model Predictive Control.

Simulation results show both the effectiveness of the proposed collision-avoidance strategies and the different behavior of the two controllers in tracking the same reference trajectory. In the scenario that we present, the MPC controller can reduce the propellant consumption by 5.71% with respect to the LQR controller.

References

[1] Larsson R., Mueller J., Thomas S., Jakobsson B. and Bodin P., “Orbit Constellation Safety on the PRISMA In-orbit Formation Flying Test Bed”. Proc. 3rd Int. Symp. on Formation Flying, Missions and Technologies, Noordwijk, The Netherlands, 23-25 April 2008 (ESA SP-654, June 2008).

[2] Mueller J. B. and Karsson R. L., “Collision Avoidance Maneuver Planning with Robust Optimization”. 7th International ESA Conference on Guidance, Navigation and Control Systems, 2-5 June 2008, Tralee, Country Kerry, Ireland.

[3] Friedland B., *Control System Design: An introduction to State-Space Methods*, McGraw Hill, Inc., 1985.

[4] Maciejowski J. M., *Predictive Control with Constraints*, 1st ed., Prentice Hall, 2000.

[5] “Genetic Algorithm and Direct Search Toolbox™ 2”. The MathWorks, Inc., 2009.

[6] “Control System Toolbox™ 8”. The MathWorks, Inc., 2009.

[7] Bemporad A., Morari M. and Ricker N. L., “Model Predictive Control Toolbox™ 3. User’s Guide”. The MathWorks, Inc., 2009.

Steering Algorithms for Single Gimbal Control Moment Gyros in Box Configuration

Matteo Meschini and Fabio Curti
DIAEE-Department of Astronautical, Electrical and Energy Engineering, Sapienza University of Rome, e-mail: meschini.matteo@gmail.com – fabio.curti@uniroma1.it, Rome, Italy

Marco Anania
Thales Alenia Space Italia, e-mail: marco.anania@thalesaleniaspace.com, Rome, Italy

Keywords: *Control Moment Gyro Steering Law*

Abstract

A singularity problem for a single gimbal control moment gyros (SGCMGs) system in a Box configuration is analyzed. The Box configuration is affected by some singular states that cannot be escaped by using just a singularity robust inverse approach. In order to avoid this particular singular conditions two alternative steering methods are proposed, based on a real-time change of the Jacobian matrix forcing the CMGs units to exit the singular condition. Numerical simulations are presented in order to evaluate the system performances for a satellite attitude control system during prescribed attitude trajectories tracking maneuvers. The new algorithms are compared to existing steering laws and used with different null-motion methods.

1 Introduction

Control Moment Gyros (CMGs) are momentum exchange actuators, consisting of a flywheel rotating at a constant speed which the angular momentum vector is rotated by one supporting gimbal (single gimbal CMGs or SGCMGs), or two gimbals (double gimbal CMGs or DGCMGs). As a result, a torque is generated in the direction that is orthogonal to both the flywheel angular momentum and the gimbals rotation axes (Fig. 1).

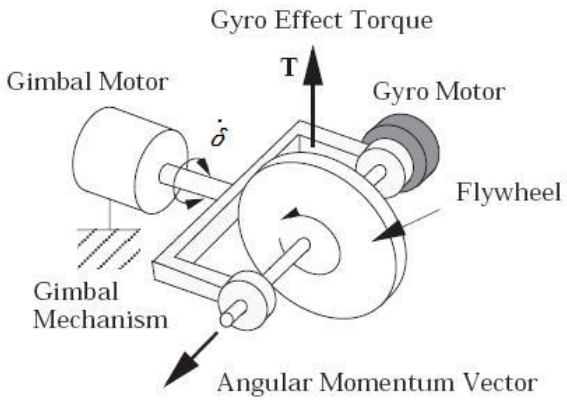


Fig. 1 Single Gimbal CMG

The main feature of the Control Moment Gyros is that they provide a high torque with a quick response with respect to the reaction wheels.

In fact, while reaction wheels can produce a maximum torque of 2 Nm, SGCMGs can reach a 5000 Nm torque capability.

The problem in using CMGs is represented by singularities that arise when a required output torque cannot be exactly provided. The control of a SGCMG cluster is based on the

inversion of the Jacobian matrix, that is a function of each CMG gimbal angle. These singular conditions come up when the classical pseudo-inverse technique is no longer usable because the matrix product between the Jacobian and his transpose results no full-rank. As a consequence the output torque cannot be given along a certain direction, denoted by the so called "singular vector".

The study of CMGs began in the 60's for the application to the Skylab's Apollo telescope mount attitude control system [1], including software studies for the actuator's steering control law.

After this studies, a double gimbal CMG system was chosen, with an approximation for the steering law Jacobian inversion.

In 1978 two independent studies by Margulies and Tokar formulate a geometric theory for singularity. Margulies introduced the gradient method null-motion for singularity avoidance, while Tokar discovered that a system with more than six SGCMG units has no internal singular states.

Based on this theory, a six SGCMG attitude control system for the MIR space station was used.

Since 1980's different approaches have been studied for the singularity avoidance of SGCMG clusters, as the *workspace reduction* [2], consisting in limiting the workspace domain of the total angular momentum or the *off-line planning*, consisting in planning the better angular momentum path before starting the maneuver.

The most studied approach is the *real-time* control, which consists in introducing a mathematical element for avoiding singularity in the Jacobian inversion.

Presently DGCMGs, which have less singularity problems, have been used on the ISS. Moreover these devices are largely studied for the attitude control of satellites thanks to the capability of constructing CMGs of small size and weight [3,4].

This work focus on the SGCMGs *real-time* control approach for a satellite attitude control system of 4 SGCMG units, aiming to analyze the behavior of two proposed steering algorithms, compared to already existing laws.

A particular configuration called Box is considered.

In order to evaluate numerical results a satellite attitude control system have been simulated, performing prescribed attitude trajectories tracking maneuvers.

The required maneuver performances are obtainable only by using CMGs actuators because of S/C high rotation rates and accelerations.

2 Mathematical Modeling

In this section the SGCMGs' (simply called CMGs by now) control theory will be described, introducing the singularity problem.

First of all, we can define the rotational motion equations of the S/C assuming it as a rigid body. Including momentum exchange actuators the equation is simply given by:

$$J\dot{\boldsymbol{\omega}} + \dot{\mathbf{h}} + \boldsymbol{\omega} \times \left(J\boldsymbol{\omega} + \mathbf{h} \right) = \mathbf{T}_{EXT} \qquad (1)$$

Where J is the S/C inertia matrix, including the actuators inertia, $\mathbf{h}$ is the CMGs actuators total angular momentum expressed in the S/C body fixed axes, the vector $\boldsymbol{\omega} = \left(\boldsymbol{\omega}_x, \boldsymbol{\omega}_y, \boldsymbol{\omega}_z \right)$ is the S/C angular velocity, and $\mathbf{T}_{EXT}$ represents the external torques.

Introducing the actuators control torque $\boldsymbol{\tau}$ we can write the Eq. (1) as a system of two equations:

$$J\dot{\boldsymbol{\omega}} + \boldsymbol{\omega} \times \left(J\boldsymbol{\omega} + \mathbf{h} \right) = \boldsymbol{\tau} + \mathbf{T}_{EXT} \qquad (2)$$

$$\dot{\mathbf{h}} = -\boldsymbol{\tau} \qquad (3)$$

The control torque $\boldsymbol{\tau}$ represents the torque given by the CMGs actuators in order to perform an assigned attitude maneuver.

We can define the CMGs total angular momentum as the sum of each unit momentum

vector, and it can be expressed as:

$$\mathbf{h} = \sum_{i=1}^{n} \mathbf{h}_i(\delta_i) \qquad (4)$$

Where $\mathbf{h}_i$ is the angular vector of the *i*-th unit, function of the gimbal angle, δ_i is the gimbal angle of the *i*-th unit representing the orientation of each unit and *n* is the actuators' units number.

To express the time derivative of Eq. (4) in order to resolve the Eq. (3) we must consider the dependence of the angular momentum vector by the gimbal angles, obtaining

$$\dot{\mathbf{h}} = \sum_{i=1}^{n} \frac{\partial \mathbf{h}_i}{\partial \delta_i} \frac{\partial \delta_i}{\partial t} = C\dot{\boldsymbol{\delta}} \qquad (5)$$

Where the *C* matrix is the system Jacobian, defined as:

$$C = \left[\frac{\partial \mathbf{h}_1}{\partial \delta_1}, \frac{\partial \mathbf{h}_2}{\partial \delta_2}, \frac{\partial \mathbf{h}_3}{\partial \delta_3}, ..., \frac{\partial \mathbf{h}_n}{\partial \delta_n} \right] \qquad (7)$$

And $\dot{\boldsymbol{\delta}}$ is the gimbal rates vector $\dot{\boldsymbol{\delta}} = \left[\dot{\delta}_1, \dot{\delta}_2, \dot{\delta}_3, ..., \dot{\delta}_n \right]^T$.

The Jacobian *C* is a $3 \times n$ dimensional matrix, and each column represents the output torque versor of the *i*-th unit, multiplied for the angular momentum module of a single unit.

To invert Eq. (5), since for redundant systems with $n>3$ (or if $n<3$) *C* is a rectangular matrix, the inversion can be made using the well known Moore-Penrose pseudoinverse (PSI), obtaining:

$$\dot{\boldsymbol{\delta}} = C^{\#}\dot{\mathbf{h}} = C^T \left[CC^T \right]^{-1} \dot{\mathbf{h}} \qquad (8)$$

Where $C^{\#}$ indicates the Jacobian matrix inverse. Equation (8) gives an exact, minimum norm, solution of $\dot{\boldsymbol{\delta}}$. On the other hand, the PSI solution exists only if CC^T is full rank. In fact, if $\det\left(CC^T\right) = 0$ the inversion in Eq. (8) is not possible and the system is in a singular state. For each singularity state a vector **u** can be defined as the *singular vector*, along which the system cannot provide torque.

A *real-time* steering control approach tries to solve this problem modifying Eq. (8) in order to give always a solution of $\dot{\boldsymbol{\delta}}$, despite the introduction of a little error while the system pass through a singular state.

2.1 The Box Configuration

The actuators cluster analyzed in this work is the so called Box configuration, that can be recognized also as a Roof type M(2,2) with orthogonal planes or a Pyramid type with a skew angle of 90 deg.

It consists of 4 SGCMGs divided in two pairs of units. Each pair has the same gimbal axis direction and the output torque lies on a plane (Fig. 2).

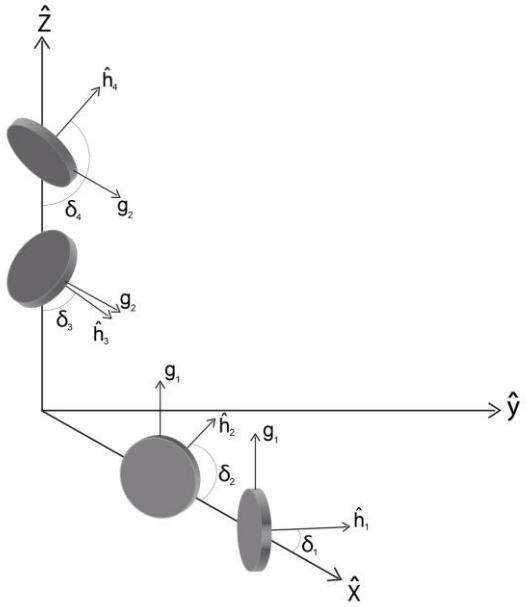


Fig. 2 Box configuration scheme

For the simulations the first units pair (CMGs 1&2) has the gimbal axis $\mathbf{g}_1$ along the body $\hat{\mathbf{z}}$ axis (Yaw axis), the gimbal angles δ_1 and δ_2 are measured starting from the $\hat{\mathbf{x}}$ axis and the generated torque lies on the *xy* plane.

The second pair (CMGs 3&4) has the gimbal axis $\mathbf{g}_2$ along the body $\hat{\mathbf{x}}$ axis (Roll axis), the gimbal angles δ_3 and δ_4 are measured starting from the $-\hat{\mathbf{z}}$ axis and the torque is given on the *yz* plane. The combination of the two pairs allow to generate a torque vector in a 3 dimensional space.

It can be now defined the total angular momentum as:

$$\mathbf{h} = H\begin{bmatrix} \cos\delta_1 + \cos\delta_2 \\ \sin\delta_1 + \sin\delta_2 + \sin\delta_3 + \sin\delta_4 \\ -\cos\delta_3 - \cos\delta_4 \end{bmatrix} \qquad (9)$$

Where H is the constant angular momentum value of each unit. The Jacobian will be:

$$C = H\begin{bmatrix} -\mathrm{s}\delta_1 & -\mathrm{s}\delta_2 & 0 & 0 \\ \mathrm{c}\delta_1 & \mathrm{c}\delta_2 & \mathrm{c}\delta_3 & \mathrm{c}\delta_4 \\ 0 & 0 & \mathrm{s}\delta_3 & \mathrm{s}\delta_4 \end{bmatrix} \qquad (10)$$

Where $\mathrm{s}\delta_i = \sin\delta_i$ and $\mathrm{c}\delta_i = \cos\delta_i$.

The singular states for the Box configuration are defined by the following angles conditions:

- $\delta_1 = \delta_2, \delta_3 = \delta_4$
- $\delta_1 = \delta_2 + \pi, \delta_3 = \delta_4$
- $\delta_1 = \delta_2, \delta_3 = \delta_4 + \pi$
- $\delta_1 = \delta_2 + \pi, \delta_3 = \delta_4 + \pi$
- $\delta_1 = \delta_2 = k\pi$
- $\delta_1 = k\pi, \delta_2 = \delta_1 + \pi$
- $\delta_3 = \delta_4 = k\pi$
- $\delta_3 = k\pi, \delta_4 = \delta_3 + \pi$

(11)

Where k=0, 1.

The comparison with the Pyramid type configuration shows that in case of failure of one CMG, the Box configuration assure the largest singularity free internal workspace, allowing a simpler torque generation with respect to a conventional Pyramid configuration [5].

Moreover, as visible analyzing Eq. (9), the Box configuration angular momentum workspace is near-ellipsoidal with a momentum capacity of 4H along the pitch axis and a capacity of 2H along the roll and yaw axes, in order to improve maneuvering performances for pitch rotations, while for a Pyramid configuration it is possible to obtain a near spherical workspace.

3 Conventional Real-time Steering Laws

Most of the *real-time* steering methods are based on the solution of a mixed, two-norm and least-squares minimization problem:

$$\min_{\dot{\boldsymbol{\delta}}}\left\{\left(C\dot{\boldsymbol{\delta}} - \mathbf{T}\right)^T P\left(C\dot{\boldsymbol{\delta}} - \mathbf{T}\right) + \dot{\boldsymbol{\delta}}^T Q\dot{\boldsymbol{\delta}}\right\} \qquad (12)$$

Subject to $C\dot{\boldsymbol{\delta}} = \mathbf{T}$ (13)

Where $\mathbf{T}$ is the required torque and P and Q are positive definite square matrices. This time the inversion of the Jacobian can be found as:

$$C^{\#} = WC^T\left[CWC^T + V\right]^{-1} \qquad (14)$$

Where W=Q^{-1} and V=P^{-1}.

The first Singularity Robust Inverse (SRI) steering method, proposed by Nakamura and Hanafusa [6], identified W and V as $W = I_{n\times n}$ and $V = \lambda I_{3\times 3}$, where λ is a properly chosen positive scalar needed to avoid the singular state. As a result, Eq. (14) becomes:

$$C^{\#} = C^T\left[CC^T + \lambda I_{3\times 3}\right]^{-1} \qquad (15)$$

The scalar λ is often selected as:

$$\lambda = \lambda_0 \exp\left[-\mu \det\left(CC^T\right)\right] \qquad (16)$$

Where λ and μ are positive scalars to be properly selected. It can be emphasized that in a

singular state $\det\left(CC^T\right)=0$ and λ tends to λ_0, permitting the existence of Eq. (15).

Though the SRI method is effective for the most of internal singularities, when the required torque lies along the singular vector the solution $\dot{\boldsymbol{\delta}}$ is a null vector, and the system cannot escape the singular state.

In order to avoid this problem Wie, Bailey and Heiberg [7] modified the SRI introducing a time dependent matrix able to pass through such a singular state.

This method, called Perturbed Singularity Robust inverse (PSR) by now, is defined by the two matrices $W=I_{n\times n}$ and $V=\lambda E$:

$$C^{\#}=C^T\left[CC^T+\lambda E\right]^{-1} \qquad (17)$$

Where

$$E=\begin{bmatrix}1 & \varepsilon_3 & \varepsilon_2\\ \varepsilon_3 & 1 & \varepsilon_1\\ \varepsilon_2 & \varepsilon_1 & 1\end{bmatrix} \qquad (18)$$

and

$$\varepsilon_i=\varepsilon_0\sin\left(\omega t+\varphi_i\right) \qquad (19)$$

Where t is the time and the amplitude ε_0, the frequency ω and the phases φ_i need to be properly selected.

All the SR based methods act modifying the output torque permitting to quit the singular state while introducing an error.

3.1 Null Motion

In addition to the solution given by the Jacobian inversion another component can be added to improve the singularity avoidance for redundant systems. This element uses the null space of the Jacobian in order to do not produce torque, while adding a proper amount of null motion $\dot{\boldsymbol{\delta}}_N$ that can be used to allow the system to steer away from passable singularities or to maintain the CMGs near a desired position.

The null motion component is defined as:

$$\dot{\boldsymbol{\delta}}_N=\left[I-C^{\#}C\right]\cdot\mathbf{d} \qquad (20)$$

Where **d** is a vector representing the amount of null motion for each unit. In every non singular state it is assured that $C\dot{\boldsymbol{\delta}}_N=0$.

Depending on the selection of **d** two principal null motion types can be defined.

The *gradient method* [2] defines **d** as a gradient of a certain function $f\left(\boldsymbol{\delta}\right)$ permitting its maximization. The function $f\left(\boldsymbol{\delta}\right)$ can be selected as a measure of singularity, such as $f\left(\boldsymbol{\delta}\right)=\det\left(CC^T\right)$.

Another type of null motion, called *preferred gimbal angles* [8], makes possible to maintain the units around a fixed position $\boldsymbol{\delta}_0$, making possible the introduction gimbal rotations constraints and increasing the reliability of the system.

4 Saturation Singularities Escape

The PSR is a steering control law effective for almost every so called internal singularities, but for the saturation singular states can present problems. In fact, when the system reach the angular momentum saturation, though the conventional steering laws produce the CMGs rotation, the system remains in a saturation state preventing actuators desaturation.

For the Box configuration this problem appears when one or more pairs of CMGs get the same orientation resulting aligned.

In this case 2 or 4 rows of $C^{\#}$ are equals and the generated rotation rates $\dot{\delta}_i$ have the same values for each pair of units, preventing the separation.

This causes the permanency in a singular state, or the increase of chances to ran into singularities.

This problem related to saturation singularities was faced by Wie in 2005 [9], proposing a modification of the PSR. He

proposed to change the W matrix of Eq. (14) in a near diagonal matrix obtaining:

$$C^{\#} = WC^{T}\left[CWC^{T} + \lambda E\right]^{-1} \qquad (21)$$

With

$$W = \begin{bmatrix} w_1 & \gamma & \gamma & \gamma \\ \gamma & w_2 & \gamma & \gamma \\ \gamma & \gamma & w_3 & \gamma \\ \gamma & \gamma & \gamma & w_4 \end{bmatrix} \qquad (22)$$

Where the diagonal elements w_i are positive scalars and $\gamma \cong 0$. The λE term is the same as the one used for PSR.

The purpose of the W matrix is to change the gimbal rates of the aligned units in order to separate them while generating near the same PSR torque error. This method has been called *Off-Diagonal SRI* (o-DSR).

Another solution has been proposed by Heiberg in 2011 [10], providing singularity avoidance using virtual array rotations. This steering logic is based on the evaluation of the determinant of CC^{T}: If its value is less than a certain threshold the Jacobian C is recalculated with a virtual gimbal angle displacement, in order to bypass the singular state and generate a torque affected by an error caused by the virtual displacements. The Jacobian evaluation and recalculation is done while the determinant reaches the threshold value.

This method allow to pass even the saturation singularities and in addiction, with respect to the o-DSR, it needs a routine with a decision block to calculate the gimbal rates $\dot{\boldsymbol{\delta}}$.

5 Proposed Steering Control Algorithms

The two proposed steering control logics are based on the PSR method together with a Jacobian change during the alignment of any units pair.

The first method (called PA1 by now) includes an evaluation of the angle between each pair of units and since the saturation singularities are caused by alignment of one or two couples of CMGs, a change of Jacobian is made in order to separate them.

If the relative displacement become less than a threshold value, the Jacobian is changed removing the column associated to the CMG that is needed to stop.

So the non stopped unit become steered like the system has only 3 units (or 2) while the stopped one is not considered in the PSR control law until the relative angle between them exceeds the threshold value.

The logic can be reassumed by the following table:

Tab. 1 PA1 logic

Angles conditions	Jacobian suppressed columns
$\lvert\delta_1-\delta_2\rvert > \Delta\delta_{\lim}$ $\lvert\delta_3-\delta_4\rvert > \Delta\delta_{\lim}$	None
$\lvert\delta_1-\delta_2\rvert \le \Delta\delta_{\lim}$ $\lvert\delta_3-\delta_4\rvert > \Delta\delta_{\lim}$	2nd
$\lvert\delta_1-\delta_2\rvert > \Delta\delta_{\lim}$ $\lvert\delta_3-\delta_4\rvert \le \Delta\delta_{\lim}$	4th
$\lvert\delta_1-\delta_2\rvert \le \Delta\delta_{\lim}$ $\lvert\delta_3-\delta_4\rvert \le \Delta\delta_{\lim}$	2nd and 4th

Where $\Delta\delta_{\lim}$ is the threshold value for the units couples relative displacement.

In Fig. 3 the system behavior is shown: On the left the CMGs couple is in an aligned condition, producing torque on the xy plane, while on the right is shown the response given through the new steering logic.

On the right is also possible to observe that the total torque is near the same as the one given by the PSR method, thanks to the increased rate for the non stopped CMG.

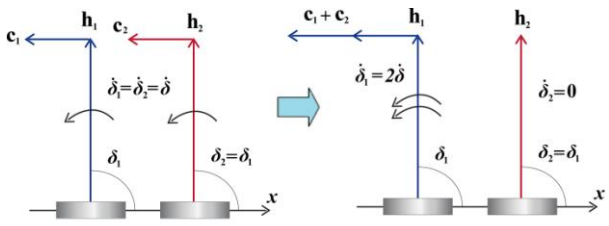


Fig. 3 PA1 Jacobian change effect

The second proposed steering algorithm (called PA2 by now) is based on the first one. The Jacobian change is made in the same way as described before, but a manipulation of the gimbal rates is made. In order to increase the separation velocity, instead of move one CMG stopping the other, a velocity component is added to both the units in opposites directions.

The gimbal rates manipulation can be schematized as follows:

Tab. 2 PA2 rates manipulation

PA1 resultant gimbal rates	New gimbal rates
$\dot{\boldsymbol{\delta}}=\left[\dot{\delta}_1,0,\dot{\delta}_3,\dot{\delta}_4\right]$	$\dot{\boldsymbol{\delta}}_{NEW}=\left[\frac{3}{2}\cdot\dot{\delta}_1,-\frac{1}{2}\cdot\dot{\delta}_1,\dot{\delta}_3,\dot{\delta}_4\right]$
$\dot{\boldsymbol{\delta}}=\left[\dot{\delta}_1,\dot{\delta}_2,\dot{\delta}_3,0\right]$	$\dot{\boldsymbol{\delta}}_{NEW}=\left[\dot{\delta}_1,\dot{\delta}_2,\frac{3}{2}\cdot\dot{\delta}_3,-\frac{1}{2}\cdot\dot{\delta}_3\right]$
$\dot{\boldsymbol{\delta}}=\left[\dot{\delta}_1,0,\dot{\delta}_3,0\right]$	$\dot{\boldsymbol{\delta}}_{NEW}=\left[\frac{3}{2}\cdot\dot{\delta}_1,-\frac{1}{2}\cdot\dot{\delta}_1,\frac{3}{2}\cdot\dot{\delta}_3,-\frac{1}{2}\cdot\dot{\delta}_3\right]$

For example, being $\dot{\boldsymbol{\delta}}=\left[\dot{\delta}_1,0,\dot{\delta}_3,\dot{\delta}_4\right]$ the gimbal rate vector obtained by the first method, the rate vector for the second method can be obtained as $\dot{\boldsymbol{\delta}}_{NEW}=\left[3/2\cdot\dot{\delta}_1,-1/2\cdot\dot{\delta}_1,\dot{\delta}_3,\dot{\delta}_4\right]$.

So the behavior of CMGs 1&2 can be shown in Fig. 4.

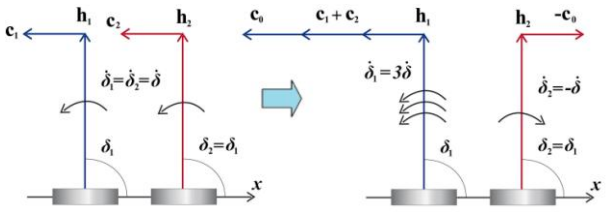


Fig. 4 PA2 solution

And the separation velocity will be $\left|\Delta\dot{\delta}_{1-2}\right|=2\cdot\dot{\delta}_1$, that is the double of the one obtained by the first method. The resultant torque is the same as the one given by the first proposed logic.

5.1 Numerical Example

Consider now a Box configuration in a saturation singularity corresponding to a gimbal angle vector $\boldsymbol{\delta}=\left[70,70,150,150\right]\text{deg}$.

Choosing the angular momentum of each unit as H=1 *Nms*, the Jacobian defined by Eq. (10) becomes:

$$C=\begin{bmatrix}-0.940 & -0.940 & 0 & 0\\ 0.342 & 0.342 & -0.866 & -0.866\\ 0 & 0 & 0.500 & 0.500\end{bmatrix} \quad (23)$$

To calculate the expressions defined by Eq. (16), Eq. (18) and Eq. (19) the following values are chosen: $\lambda_0=0.01$, $\mu=10$, $\varepsilon_0=0.01$, $\omega=\pi/2$, $\boldsymbol{\varphi}=\left[0,\pi/2,\pi\right]$ and t=15 *s*. The W matrix for the o-DSR is selected as W=diag(1,2,3,4).

For this example the desired torque will be $\mathbf{T}=\dot{\mathbf{h}}=\left[-1,0,0\right]^T\ Nm$.

The PSR steering law defined by Eq. (17) provides:

$$C^{\#}=\begin{bmatrix}-0.511 & 0.049 & 0.083\\ -0.511 & 0.049 & 0.083\\ 0.151 & -0.417 & 0.272\\ 0.151 & -0.417 & 0.272\end{bmatrix} \quad (24)$$

And the gimbal rate vector $\dot{\boldsymbol{\delta}}=C^{\#}\mathbf{T}$ becomes:

$$\dot{\boldsymbol{\delta}}_{PSR}=\left[0.51,\ 0.51,\ 0.15,\ 0.15\right]rad/s \quad (25)$$

As can be seen, the CMGs 1&2 gimbal rates are equals, as well as the CMGs 3&4 ones, preventing the singularity escape.

The torque $\mathbf{T}=C\dot{\boldsymbol{\delta}}$ given by PSR is:

$$\mathbf{T}_{PSR}=\left[-0.9611,\ 0.0883,\ 0.1510\right]Nm \quad (26)$$

The o-DSR described by Eq. (21) gives the following rate solution and torque:

$$\dot{\boldsymbol{\delta}}_{o-DSR} = \left[0.34,\ 0.68,\ 0.13,\ 0.17\right] rad/s \quad (27)$$

$$\mathbf{T}_{o-DSR} = \left[-0.9632,\ 0.0875,\ 0.1519\right] Nm \quad (28)$$

For the method proposed by Heiberg, choosing a virtual displacement of 1° for the 2nd and 3rd CMG units we obtain the Jacobian as:

$$C = \begin{bmatrix} -0.940 & -0.946 & 0 & 0 \\ 0.342 & 0.326 & -0.866 & -0.875 \\ 0 & 0 & 0.500 & 0.485 \end{bmatrix} \quad (29)$$

And the rate vector and output torque are:

$$\dot{\boldsymbol{\delta}}_{HEI} = \left[-3.58,\ 4.62,\ -7.73,\ 7.97\right] rad/s \quad (30)$$

$$\mathbf{T}_{HEI} = \left[-0.9731,\ 0.1445,\ 0.1211\right] Nm \quad (31)$$

For the first proposed steering algorithm (PA1) in the new Jacobian the 2nd and 4th columns are suppressed in order to stop CMGs 2&4 and permit the couples separation, obtaining:

$$C = \begin{bmatrix} -0.940 & 0 \\ 0.342 & -0.866 \\ 0 & 0.500 \end{bmatrix} \quad (32)$$

So the rate vector and output torque become:

$$\dot{\boldsymbol{\delta}}_{PA1} = \left[1.017,\ 0,\ 0.299,\ 0\right] rad/s \quad (33)$$

$$\mathbf{T}_{PA1} = \left[-0.9555,\ 0.0891,\ 0.1493\right] Nm \quad (34)$$

The second proposed logic (PA2) provides the following rate vector and torque:

$$\dot{\boldsymbol{\delta}}_{PA2} = \left[1.53,\ -0.51,\ 0.45,\ -0.15\right] rad/s \quad (35)$$

$$\mathbf{T}_{PA2} = \left[-0.9555,\ 0.0891,\ 0.1493\right] Nm \quad (36)$$

The results show that the output torque during the singular state presents comparable errors. The main difference between the different steering laws is represented by the gimbal rates. In fact the separation velocity for CMGs 1&2 are: $\Delta\dot{\delta}_{o-DSR} = \left|\dot{\delta}_1 - \dot{\delta}_2\right| = 0.34\ rad/s$ $\Delta\dot{\delta}_{HEI} = 8.20\ rad/s$, $\Delta\dot{\delta}_{PA1} = 1.017\ rad/s$ and $\Delta\dot{\delta}_{PA2} = 2.034\ rad/s$.

Usually a real gimbal motor can provide gimbal rates often limited to $1 \div 2\ rad/s$ and accelerations limited to few rad/s^2 and in case of gimbal rates saturation, a proportional scaling for all the CMGs rates is made, in order to maintain the output torque along the desired direction even if decreased in value.

So even if in an ideal system the Heiberg steering logic would be preferable with respect to the others because of its higher separation velocity, probably the high rates shown by Eq. (30) would be limited by the motor constraints, affecting the output torque more than the others methods.

6 Simulation Results

In order to evaluate the effectiveness of the presented steering logics a large quantity of simulations have been carried out simulating a typical Earth observation mission in a sun-synchronous orbit. During simulations ideal sensors and ideal gimbal motors have been considered, together with a rate saturation to $\dot{\delta}_{\lim} = \pm 1\ rad/s$ and an acceleration limit of $\ddot{\delta}_{\lim} = \pm 1\ rad/s^2$. The gravity gradient torque has been introduced as external torque.

To calculate the torque needed by CMGs for the prescribed attitude trajectories tracking maneuvers a PID controller was used, together with a feedback-linearization component and a

feed-forward control torque based on the reference S/C accelerations, developed as:

$$\boldsymbol{\tau} = -K_P\mathbf{e} - K_D\left(\boldsymbol{\omega} - \boldsymbol{\omega}_r\right) - K_I\int \mathbf{e}\,dt + \\ + J\dot{\boldsymbol{\omega}}_r + \boldsymbol{\omega}\times J\boldsymbol{\omega} + \boldsymbol{\omega}\times\mathbf{h} \qquad (37)$$

Where the torque $\boldsymbol{\tau}$ has been defined in Eq. (3), K_P, K_D and K_I are diagonal and definite positive matrices, $\boldsymbol{\omega}$ is the measured angular rate of the S/C, $\boldsymbol{\omega}_r$ is the reference S/C rate, **h** is the actuators angular momentum vector, **e** is the quaternion error vector and J is the S/C inertial matrix, that for the analyzed satellite has been taken:

$$J = \begin{bmatrix} 7500 & 375 & 750 \\ 375 & 11250 & 150 \\ 750 & 150 & 9000 \end{bmatrix} kg\cdot m^2 \qquad (38)$$

The following simulation results are referred to a 30 deg pitch-axis reorientation maneuver, depicted in Fig. 5.

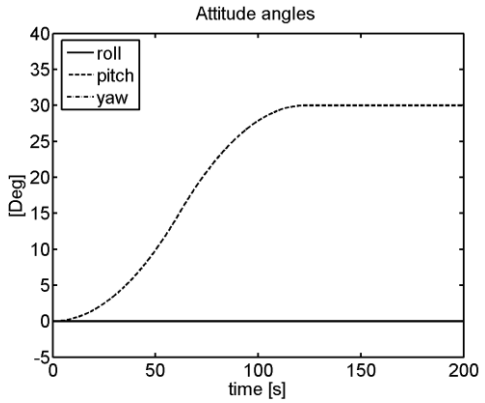


Fig. 5 Pitch-axis reorientation maneuver

For this simulation it will be considered a worst case, starting from a singular state that cannot be escaped by the SRI or PSR as described before.

The set of the initial gimbal angles is chosen as $\boldsymbol{\delta}_0 = [80,80,260,260]$ deg and the angular momentum of each CMG is assumed as H=30 *Nms*. The values of λ, E and W have been already defined in the previous paragraph. The angle threshold value for PA1 and PA2 has been set to $\Delta\delta_{\text{lim}} = 0.1\,\text{deg}$.

In Fig. 6 the attitude errors obtained by each steering law are showed, and as can be seen, the error values are well comparable. At the beginning of the maneuver the escape from the initial singularity produces a larger error for the PA1, while for the o-DSR it becomes very small.

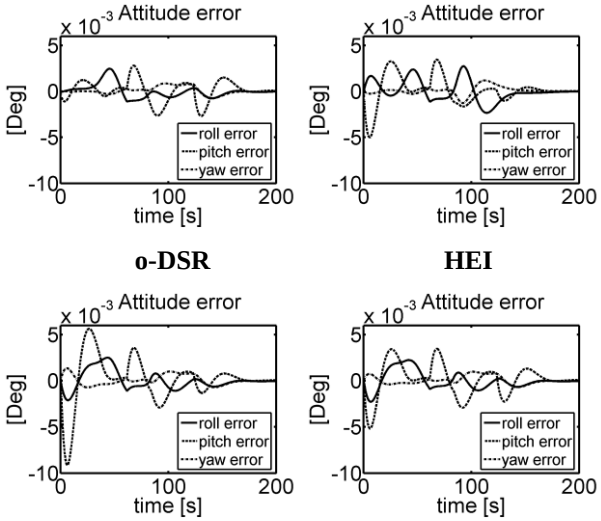


Fig. 6 Attitude angles errors

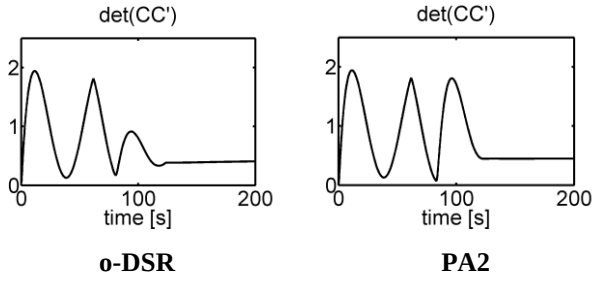


Fig. 7 det(CC^T) time histories

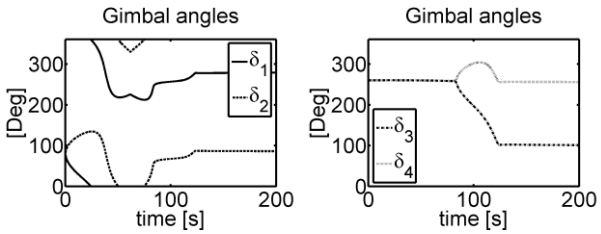


Fig. 8 CMG gimbal angles time history for PA2

In Fig. 7 the determinant time histories for the o-DSR and PA2 are represented, showing an higher average value for PA2, whereas in Fig. 8

it is possible to see the gimbal angles time history of each units couple for the PA2 (CMGs 1&2 on the left and 3&4 on the right)

The results for the PA1 are very similar to the PA2 ones and will not be shown.

To further evaluate the effectiveness of PA1 and PA2, singularity escape simulations have been carried out starting from all type of singular states related to the Box configuration.

Moreover an instrument scanning maneuver has been analyzed, using the two different methods for null motion, of which the results are not shown here. Though the main features revealed by simulations are the meeting of the attitude errors with the scanning phase errors constraints (errors < 3 mdeg) and the possibility to introduce mechanical constraints to the gimbals rotation using a preferred gimbal angles null motion, simplifying the electrical interfaces and increasing the system reliability.

7 Conclusion

Two new steering logics has been proposed for a SGCMG cluster in a Box configuration based on the PSR steering law, together with a real-time Jacobian change made in order to escape saturation singularities, non escapable by the simple PSR.

PA1 and PA2 separation velocity for the aligned units is faster than the o-DSR one, whereas the new laws do not need a routine block with respect to the Heiberg logic.

The test with different null motions has demonstrated that a gradient method can avoid certain type of singularity, while the preferred gimbal angles method makes possible the introduction of mechanical constraints to the gimbal rotations, simplifying the electrical interface and increasing reliability.

References

[1] Chubb W. B. and Seltzer S. M., "Skylab Attitude and Pointing Control System," NASA TN D-6068, 1971.

[2] Kurokawa H., "A Geometric Study of Single Gimbal Control Moment Gyros – Singularity Problems and Steering Law," Ph.D. Dissertation, Aeronautics and Astronautics, Univ. Tokyo, Japan, 1997.

[3] Lappas V. J., Steyn W. H. and Underwood C. I., "Attitude Control for Small Satellites using Control Moment Gyros," *Acta Astronautica*, Vol. 51, No. 1-9, 2002, pp. 101-111.

[4] Gersh J. and Peck M. A., "Violet: A High-Agility Nanosatellite for Demonstrating Small Control-Moment Gyroscope Prototypes and Steering Laws," *AIAA Guidance, Navigation, and Control Conference*, Chicago, IL, 2009.

[5] Sands T. A., Kim J. J. and Agrawal B., "2H Singularity-Free Momentum Generation with Non-Redundant Single Gimbaled Control Moment Gyroscopes," *IEEE Conference on Decision & Control*, San Diego, CA, 2006, pp. 1551-1556.

[6] Nakamura Y. and Hanafusa H., "Inverse Kinematic Solutions with Singularity Robustness for Robot Manipulator Control," *Journal of Dynamic Systems, Measurement, and Control*, Vol. 108, 1986, pp. 163-171.

[7] Wie B., Bailey D. and Heiberg C. J., "Singularity Robust Steering Logic for Redundant Single-Gimbal Control Moment Gyros," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 5, 2001, pp. 865-872.

[8] Vadali S. R., Oh H. S. and Walker S. R., "Preferred Gimbal Angles for Single Gimbal Control Moment Gyros," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 6, 1990, pp. 1090-1095.

[9] Wie B., "Singularity Escape/Avoidance Steering Logic for Control Moment Gyro Systems," U.S. Patent 6,917,862 B2, 12 July 2005.

[10] Heiberg C. J., "Singularity Escape and Avoidance using a Virtual Array Rotation," U.S. Patent 7,904,214 B2, 8 March 2011.

CEAS 2011

GENERAL LECTURE
'What future for the European combat aircraft industry?'

3rd CEAS Air&Space Conference
21st AIDAA Congress

What Future for the European Combat Aircraft Industry?

A Death foretold?

Dr. Georges Bridel
Fellow, Air & Space Academy, France
Member of the Board ALR Aerospace Project Development Group, Zurich, Switzerland

Keywords: *Combat Aircraft, European Industry, Industry, European Defence*

Abstract

Production of the 3 new generation European combat aircraft Eurofighter, Gripen and Rafale lasting some more years, the fundamental question arises: what kind of combat aircraft[1] (manned or unmanned) will be required to guarantee Europe's future security? Focussing on production capability only, the design and development capabilities however erode, if there are no measures taken. Maintaining design and development capability of modern combat aircraft requires therefore continuous work in this technical and operational area on products and operational/technology demonstrators. The paper presents an evolution of the position paper [1] issued by the Air and Space Academy on June 30 2011 to the 26 EU member state ministries of defence, European authorities and industry leaders.

[1] For "combat aircraft", read "aerial combat system", onboard or not, which includes both combat aircraft and combat UAV and encompasses airframe, engine, equipments and weapons systems.

1 Air warfare scenarios without advanced combat aircraft are no longer feasible

In Europe, Defence budgets are shrinking. This fact sends out the wrong signal that European States are generally feeling safer and more comfortable in a world which is itself getting more and more uncertain.

Any projection as regards the next generation of combat air systems is necessarily based on anticipated future conflict scenarios. At present, asymmetric warfare dealing mainly with terrorist organisations and multinational interventions in failed states are in the spotlight. In many such cases, mission aircraft, light aircraft and UAVs are up to the job. However, as can be seen in the current crisis in Libya, even in asymmetric conflicts, sophisticated air combat forces are essential in order to identify targets, ensure strike accuracy, be efficient on mobile targets and avoid collateral damage.

In addition, growing economic competition, dwindling resources and the consequences of climate change might generate multi-polar conflicts between nations and continents going far beyond current conflict scenarios. This is another reason why Europe must take care to maintain an advanced combat aircraft industry.

Finally, diplomatic and economic pressures supported by the presence of military forces and, at the worst, armed conflicts involving mostly sophisticated air power are likely to emerge in out-of-area scenarios where resources are a key to Europe's economic and political survival.

2 Today, Europe runs the risk of losing its air power independence

A state nowadays relies for its independence on its capacity to project power, which itself depends on an advanced air combat force.

Since WWII, European nations have developed and produced several generations of highly competitive combat aircraft. Today three aircraft – Eurofighter, Gripen and Rafale – are currently produced, fulfilling European defence needs and securing industrial and strategic independence.

Of course, this industrial heritage should be optimised to propose a single successor to these three combat aircraft but, so far, no movement is yet visible to secure this crucial European strategic asset in the future.

At the same time, in addition to the USA which has already developed the F-22 and is developing the new JSF/F35 combat aircraft with the financial support of some European countries, emerging powers worldwide are now building up forces equipped with their own advanced systems such as high performance combat aircraft which are developed, procured and soon exported in considerable numbers all over the world.

In the coming decades, the introduction on the market of advanced 5th Generation combat aircraft such as the T-50 (Russia, India), J-20 (China), KFX (Korea, Indonesia) combined with in-flight refuelling capacity allowing for out-of-area interventions, will demonstrate these strategic ambitions.

Europe's strategic independence is therefore at stake, if no action is taken to secure independent air power capabilities.

3 A strong European Industrial Base is mandatory to Support Operations and Upgrades

Another factor of growing importance that must be taken into account is the efficient use of advanced systems in operations. Continuous and close industrial support is absolutely essential during air operations. Such support services are equally needed to maintain air capabilities through future upgrades. This twofold support comes mostly from the industry having produced the systems.

Growing difficulties investing in upgrades of European combat aircraft are a warning that the European air operations capacity is going to shrink.

Acquisition or even participation in a foreign development and production programme – an example is given with the F-35 Joint Strike Fighter JSF - will never be sufficient to maintain overall European airpower capability since key technologies and fundamental air power characteristics will remain exclusively with the leading partner, in the case of F-35 with the USA.

It is now clear that the JSF/F-35 process has significantly weakened the independent capacity of the European combat aircraft industry as a whole. It is urgent for Europe to react.

4 The Combat Aircraft Industry is a driving force for advanced technologies and qualified employment

It is not only European air power capabilities that are at stake: the European combat aircraft industry has become an important factor in employment and technology, enriching the civil aircraft sector as well as many other high technology industries, research institutions and universities.

And these capabilities are sustained by a large spectrum of supplier industries, with engine, electronic and weapons systems manufacturers in pride of place. All over Europe, some 120,000 highly qualified people are employed in the combat aircraft industry.

A complete industrial network is therefore under threat of being destroyed.

5 Fundamentals to secure the European Combat Aircraft Industry

It is common knowledge that maintaining these capabilities requires continuous work on concrete programmes involving the research for new technologies and related know-how in design, development and production.

It goes without saying that European manufacturers cannot maintain such capabilities on a purely national basis, given that no single government will commission for instance a new combat aircraft programme of its industry. The related investments are too high and the single national market is too small. We obviously must think on a European scale.

It is generally acknowledged, in Europe as well as in the USA that such combat aircraft capabilities can only be maintained through affordable demonstrator projects and new programs, initiated roughly every 15 and 30 years respectively, going well beyond existing European TDPs (Technology Demonstrator Programs). It must be remembered that demonstrator variants of today's European combat aircraft have flown for the first time some 25 years ago already.

If no decision is taken on a European level, the industrial capability to design and produce combat aircraft will rapidly be lost and could only be rebuilt by means of a tremendous effort lasting decades.

6 Actions required in order to secure Industrial European Combat Aircraft Capability

The Air and Space Academy strongly believes that, in order to avoid the future death of the independent European combat aircraft industry, the following solutions should be urgently implemented:

Any definition of a future European combat aircraft programme must stem from the expression of joint operational needs and clearly specify the required industrial capabilities; this approach must be reinforced by technical and operational simulations on a global scale.

Urgent actions must be taken in Europe in order to ensure the survival of the avionics (radars, sensors, on-board systems) and weapons industries, the latter of which is currently losing the battle against its competitors. Major weapons programmes will secure the technology level but upgrade programmes are also essential to secure the survival of this industry.

Further consolidation of European industry is required in order to develop new technologies, demonstrators or capabilities. With this aim in mind, it is necessary to go beyond classical cooperation models, and investigate new forms of industrial organisation:

- creation of an industrial integrated structure for each programme
- or the creation of a European integrated company for the air combat industry.

Experience shows that any attempt at industrial consolidation remains artificial and is likely to fail if it is not based on a major programme.

The best way of putting together a successful multinational programme is to include a limited number of cooperating countries (2 or 3) in its launch, all of which should agree on their respective roles and be prepared to encourage other states to join in by providing attractive conditions.

A European operational demonstrator programme with 5th generation technology should be launched immediately, taking advantage of the various technology demonstrators already in development and being tested in Europe, thus bridging 15 years gap mentioned above. At the same time, in order to validate the new operational capabilities envisaged, it would be advisable to develop functional demonstrators.

Lastly, a comprehensive long-term investment plan must be set up, sponsored by industry and European as well as national institutions in order to secure Europe's future

capacity to design, develop, produce and operate a new generation of combat aircraft. The EDA project "Future Air System for Europe[2]" will present in autumn 2011 a comprehensive capability road-map enabling future developments.

7 Conclusion

To-day Europe can still rely on a strong aeronautical industrial base. Tomorrow Europe will be able to secure its future strategic air power independence only if actions are urgently decided and funded. This is the price to pay now if Europe is to maintain its place in the newly emerging multi-polar world.

References

[1] Air and Space Academy., "What future for the European combat aircraft industry?", *2011 (also available in French, German, Italian and Spanish)*

[2] **FAS**4Europe Study: Future Air Systems for Europe

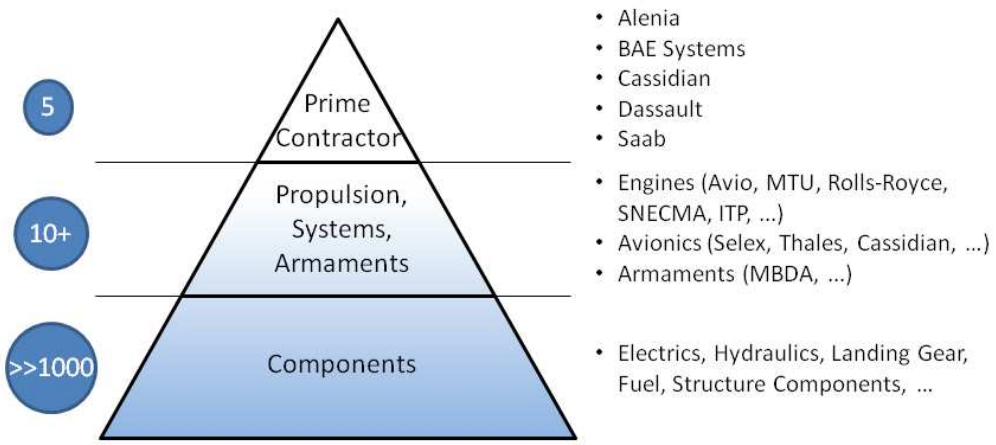


Fig 1. Industrial Network Combat Air Systems

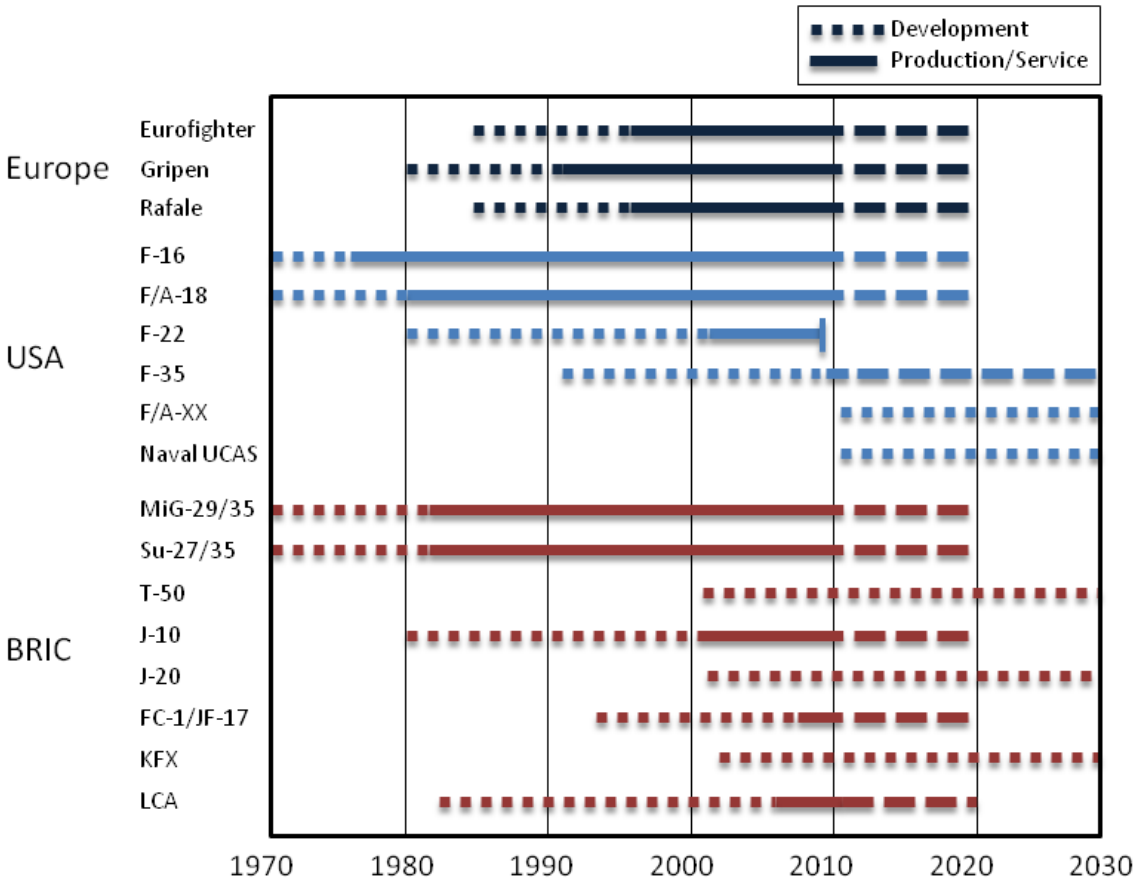


Fig 2. Combat Aircraft Generations, Development and Production

CEAS 2011

ATM and SESAR

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Comparative Study of Complexity Metric for centralised ground based Air Traffic Management

Saam Hasan and Anthony Rossiter
University of Sheffield
Department of Automatic Control & Systems Engineering
Mappin Street, Sheffield, United Kingdom

Keywords: Air Traffic Control, Air Traffic Management, Airspace, Complexity, Sector.

Abstract

This paper outlines an extended literature review on the complexity of centralized ground based Air Traffic Management. The advantages and disadvantages of several complexity metrics are outlined and some promising solutions are described briefly that can act as alternatives to the current approaches in the literature.

1 Introduction to Air Traffic Management Complexity

An Air Traffic Management (ATM) system is a multi agent system, where many aircraft are competing for a common, congestible resource represented by airspace and runways, while trying to optimize their own cost. Co-ordination between different aircraft is needed to avoid conflicts where two or more aircraft get too close one to the other [4]. ATM consists of a set of flight rules and procedures aimed at guaranteeing tactical and safe separation between aircraft and between aircraft and obstacles and to enable airspace users to meet their schedules in accordance to their preferred flight profiles without compromising safety levels [4]. The main goal of ATM is to guarantee safety and to give aircraft optimal trajectories to fly from one airport to another. On the other hand, Air traffic control (ATC) organizes air traffic flows in order to ensure flight safety and to increase the capacity of the route network. In order to manage the airspace safely, the airspace is divided into ATC sectors, such a sector is a region of airspace defined by geographical and height boundaries [6]. Each sector has a team of ATC's responsible for providing air traffic control service to the sector. ATC also carry out Traffic Flow Management (TFM) functions, in which ATC control the aircraft(s) in an efficient way in order to facilitate a unique traffic flow pattern over a long-term time horizon [12]. The sector capacity is defined and allocated in such a way that it doesn't compromise the workload of ATC.

In some aspects a managed airspace can be thought of as a control system where ATC acts as a feedback controller as shown in Fig.1. The lines represent multidimensional signals and the metric block is used to represent some form of complexity indication that can represent a flow pattern of the Air Traffic environment. In the feedback control scheme the Aircraft model, Flight Management System and the flight plan blocks act as a sector depending on the number of aircraft involved.

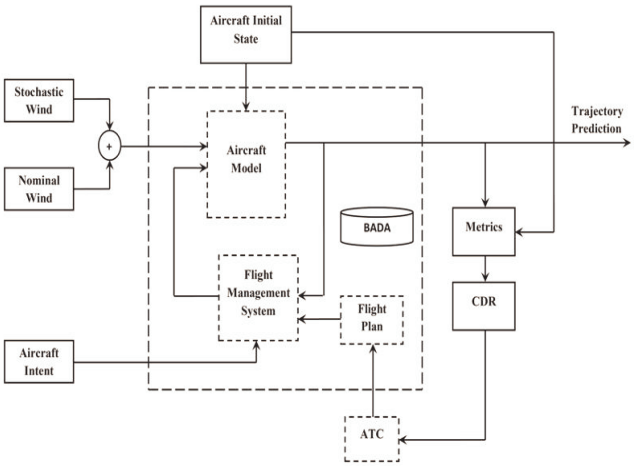


Figure 1: Feedback Control Scheme

For many years, a constant increase of air traffic has induced a large congestion in ATC sectors. Certain countries such as France are considered as a crossroad for most European destinations. This results in a huge amount of ATC workload. In this case, the workload of the ATC is implicitly related to the functioning of the TFM system, hence a complexity measure of ATM is crucial in evaluating TFM workload [14]. The consensus view among ATC research and operational communities is that complexity drives controller workload, which in turn is thought to ultimately limit sector capacity. Although it is very difficult to quantify complexity in order to capture every feature that can assist ATC in making sound and correct decisions. It can however be used as a tool for ATC to identify congestion and problematic situations in the sector in question. For example, sectors can be designed in order to balance congestion. Another example where a metric is required is the traffic assignment [5] for which an optimal time of departure and a route are searched for each aircraft in order to reduce the congestion in the airspace. There is also a need to design air networks which facilitate free flight concepts [12].

This paper is focused on reviewing the current literature on complexity metrics. The main contribution is an extended comparative study of the different approaches in the literature for complexity metrics which is outlined in section 2 followed by a discussion in section 3. A second contribution is to consider alternative approaches which are outlined in section 4. It also gives a summary of the proposals to be presented in the paper as well as indications of where there is likely to be some benefits for future work.

2 Approaches to Air Traffic Complexity within ground-based ATM

Air traffic complexity is defined in [5] as "... measure of difficulty that a particular traffic situation can impose on ATC ...". This can act as a fundamental measure since it is possible to evaluate how the ground based ATM is operating, and to provide some form of approximation to the traffic flow so that guidelines can be implemented on how to effectively manage an airspace. Other uses can include on how to efficiently allocate ATC staff to a pre-defined sector, so that ATC productivity increases.

An excellent survey was carried out in [6] in which the authors outlined an extensive review of some of the complexity measures in the current literature. Some of the measures will be reviewed in this section for sake of completeness. The airspace complexity is related with both the structure of the traffic and the geometry of the airspace. Different efforts are underway to measure the whole complexity of the airspace. Significant research interest in the concept of ATC complexity was generated by the Free Flight operational concept. At the moment, most approaches assume ground based centralized ATM, these can be classified as shown in the following subsections.

2.1 Aircraft Density

The current ATM system uses Aircraft Density as a complexity metric; it is defined as a threshold based on aircraft count to measure sector level capacity and ATC workload [5]. It is

widely recognized, and has been the most cited, studied, and evaluated in terms of its influence on workload [5]. It is considered as the best available indicator of complexity because of its ease of interpretation. However, the number of aircraft divided by the effective volume of airspace (i.e., only the amount of airspace the aircraft are using) within a given sector on its own has significant shortcomings in its ability to accurately measure and predict sector level complexity, it's also criticized for not being able to appropriately characterize what ATC find complex [5]. For example, the observation of controlled sectors shows that sometimes ATC accept aircraft beyond the capacity threshold, while in other situations they refuse traffic before the capacity is reached. This observation clearly shows that the operational metric alone cannot account for ATC's workload [5]. This measure also does not take into account factors such as traffic pattern, traffic predictability, disturbances, etc., which influence the actual workload levels experienced in practice [5].

Even though Aircraft Density has many disadvantages, it has been widely used as a decision making tool and also to carry out some different schemes and strategies. For example, ATC managers and supervisors adapt the traffic demand to the existing capacity (slot-route allocation, collaborative decision making etc...), and sometimes they adapt the capacity to the demand (modification of the air network, sectorization, new airports etc ...). They are also used to split or merge the sectors based on the number of ATC's, flight numbers and sector capacity.

2.2 Dynamic Density

Integral to Free Flight was the notion of dynamic density. Conceptually, dynamic density is a measure of ATC complexity that would be used to define situations that were so complex that centralized control was required [12]. This was introduced within a research program in the U.S. involving Federal Aviation Administration (FAA), NASA, MITRE, and Wyndemere Corporation as main participants. In [12] dynamic density was defined as a measure of control related workload that is a function of the number of aircraft and the complexity of traffic patterns in a volume of airspace. Wyndemere inc. [6] proposed a measure of the perceived complexity of an air traffic situation, which is related to the cognitive workload of the ATC with or without knowledge of the intents of the aircraft.

Mathematically, dynamic density is a single aggregate indicator obtained as a linear combination of traffic density and other controller workload contributors (the number of aircraft undergoing trajectory change and requiring close monitoring due to reduced separation). More precisely, dynamic density is the weighted sum of the number of aircraft and the following aggregate indicators of the aircraft changing geometries during a one-minute sample time interval as outlined in [5]:

- the number of aircraft with heading change > 15 degrees.
- the number of aircraft with speed change $> 0:02$ Mach.
- the number of aircraft with altitude change > 750 ft.
- the number of aircraft with 3-D Euclidean distance between $0-5$ nautical miles excluding violations.
- the number of aircraft with 3-D Euclidean distance between $5-10$ nautical miles excluding violations,
- the number of aircraft with lateral distance between $0-25$ nautical miles and vertical separation $< 2000 = 1000$ feet above/below 29000 ft.
- the number of aircraft with lateral distance between $25-40$ nautical miles and vertical separation $< 2000 = 1000$ feet above/below 29000 ft.
- the number of aircraft with lateral distance between $40-70$ nautical miles and vertical separation $< 2000 = 1000$ feet above/below 29000 ft.

The parameters of the sums have been adjusted by showing different situations of traffic to several ATC's. Validation of the dynamic density measure was carried out in an operational environment and shown to be highly correlated with ATC's activity [12]. Note that the complexity indicators entering the dynamic density metric depend on the aircraft trajectories during a one-minute sample time interval and does not include ATC's physical work such as data entry and radio communication [5].

In this case it is then possible to predict the dynamic density measure over a suitable time horizon so as to forecast future workload levels and use this information for traffic management purposes [5]. This can be carried out by using the speed and aircraft positions from the Center-TRACON Automation System [9] in conjunction with aircraft dynamic models, flight plans, radar tracks that are available from the Air Route Traffic Control Center (ARTCC), and weather data.

There are some disadvantages to be aware of. Firstly, the computed weights need to be recalculated for each sector. In addition, the proposed dynamic density model for air traffic complexity is actually a static model, which does not incorporate explicitly either future predicted aircraft or past state information [5]. It is also impossible to interpret from a single measure on how to solve a high workload situation, which leads to losing important information on which complexity factor has caused the problem [5].

Dynamic density does not correspond to a single metric but a collection of metrics, and is used in a variety of contexts in the literature. Most of the factors used in dynamic density models are dynamic traffic characteristics that are generally useful for real-time decision support [5]. Various linear and nonlinear methods are used to perform the correlation. A human-in-the-loop simulation study with ATC's controlling traffic in a real-time simulation environment was performed in [17] with the aim of introducing a new dynamic density complexity model and for validating dynamic density versus aircraft count. It was outlined that dynamic density is better than a simple aircraft count for both the instantaneous and the predicted complexity measures [5]. In [14], the use of dynamic density as an operationally useful sector workload measure for enabling TFM personnel to prevent overloads was investigated. An approach was proposed in [18] in which the authors defined five capacity estimation methods based on simplified dynamic density. Simulation was carried out in an Airspace Concept Evaluation System (ACES) using historical traffic schedules with the unlimited and actual airport capacity. The sector capacity formula used is based on average flight time, which does not include capacity adjustments based on human judgment. Results based on the delays, aircraft counts and sector capacities were reported and analyzed, and they concluded that their approach produced the shortest average delay time.

Issues such as structural considerations have also been considered since positions and speeds of the traffic itself do not appear to be enough to describe the total complexity associated with an airspace. Including structural consideration has been explicitly identified in work at Eurocontrol [12]. In a study to identify complexity factors using judgement analysis, airspace design was identified as the second most important factor behind traffic volume [12].

It was outlined in [12] that cognitive aspects affecting workload appear to simplify traffic situations. Wyndemere Corporation proposed a dynamic density metric that included a term based on the relationship between aircraft headings and the dominant geometric axis in a sector [5]. Special attention was given on the traffic and airspace characteristics that impact the cognitive and physical demands placed on ATC including an attempt to include the level of knowledge about the intent of the aircraft. A survey was carried out in [5] that describes numerous approaches to air traffic complexity that are characterized by a dynamic density model similar to that in [12]. In [18], the authors have com-

bined sector workload factors in a single workload metric which are used as a classification into airspace design factors, dynamic traffic characteristics, and operational factors. Other literature also consider the level of importance on representing complexity in such a a way that can outline actions and decisions that affect ATC's workload [5].

2.3 Interval complexity

The interval complexity of a given sector "s" over a given time interval $[m, ..., m + kL]$ is the average of the moment complexities of s at the $k + 1$ sampled moments $m + iL$, for $0 \leq i \leq k$:

$$C(s, m, k, L) = \frac{\sum_{i=0}^{k} C(s, m + iL)}{k + 1} \quad (1)$$

where k is called the smoothing degree, L is the time step between the sampled moments, and "s" is the sector in question 5.

Interval complexity is in a way similar to dynamic density, but can be considered as a smoothed version of a dynamic density-like complexity measure. The prediction of this measure of complexity is long term (time horizon of 20 to 90 minutes) [5]. Its main use is for selecting appropriate complexity resolution actions with the aim of minimizing and balancing traffic complexities between adjacent sectors of a certain airspace region.

2.4 Geometric Technique based on Fractal Dimensions

A geometric approach based on fractal dimension has been outlined in [5] and explained in [7]. The fractal dimension is a unique metric for the evaluation and comparison of alternative traffic patterns emerging from differing operational concepts. One of the interesting aspects of fractal geometries, is that one can often compute a non-integer fractal dimension. The fractal geometry is often defined as the ratio of the log of the number of similar copies to the log of the scale. For example doubling a square ($scale = 2$) will give four copies of the square. This gives $\log(4)/\log(2))$ a fractal dimension of 2 for a square. It also allows some decoupling of the complexity due to airspace partitioning in sectors from the complexity due to traffic flow features [6]. The dimension of geometrical figures is 1 for a curve, a surface has dimension 2, and so on. It is quite simple to derive those integers from a covering measure since the minimal number of balls of radius ϵ needed to cover the object will evolve roughly as $(\frac{1}{\epsilon})^d$ as $\epsilon \to 0$ and d being the dimension.

Fractal dimension can be considered as an extension of this concept to more complicated figures in which their dimensions may not be an integer [7]. The block count approach is used to compute fractal dimensions which is a simpler procedure for computing a geometrical entity. It consists of a rectangular grid of size d and then counting the number N of blocks of linear dimension d covering the given geometrical entity. Then, the fractal dimension of the geometric entity is defined as [6]:

$$D_0 = \lim_{d \to 0} \frac{\log N}{\log d} \quad (2)$$

The application of this concept is used in airspace analysis by computing the fractal dimension of the geometrical figure composed of existing air routes. Currently, aircraft cruise on linear routes at specified altitudes, corresponding to a geometrical dimension of 1. If all of the airspace was covered by routes, the fractal dimension of the future route structure would be 3. However, there will still be preferred routes (due to the position of connected airports, or to wind currents, etc.), thereby decreasing the actual dimension of the route structure [5].

Fractal dimension also provides information on the number of aggregate degrees-of-freedom present in a collection of traffic being analyzed [7]. A higher fractal dimension indicates more degrees of freedom. This information is independent of sectorization and does not scale with traffic volume. Fractal dimension is an example of complexity metric independent of workload aspects. It must be thought of as a geometrical feature of a limit shape obtained by observing trajectories on an infinite time period. The fact that timing information is lost which is a main limitation of fractal dimension as a complexity

measure [5]. In addition, higher fractal dimension indicates preferable scalability in terms of the number of conflicts and it also indicates a higher degree of freedom, but on the other hand a higher fractal dimension may also indicate a higher traffic complexity measure.

2.5 Input-output model

In [8], air traffic complexity is defined in terms of the control effort required to avoid the occurrence of conflicts. The authors introduce an input-output system consisting of the air traffic within the region of the airspace under consideration and a feedback controller. ATC is replaced by an automatic solver and the airspace sector by an airspace region [9]. The input to the closed-loop system is represented by a (fictitious) aircraft entering the airspace region, and the output is given by the deviation from their original flight plans issued by the feedback controller to the aircraft already present in the airspace so as to avoid conflicts. The deviation imposed by the controller is taken as measure of the air traffic complexity. Each aircraft is described by a very simple 2D kinematic model.

$$\dot{x} = V_i \cos\theta_i \quad \dot{y} = V_i \sin\theta_i \tag{3}$$

Where $\dot{x}$ and $\dot{y}$ denote the aircraft position at a fixed altitude, V_i is the speed and θ_i is the heading. The authors assume that a conflict solver is available and that the speed is constant and only the heading can be used a resolution maneuver. The solver is based on mixed integer linear programming. Minimum control activity required is the complexity measure and is described by the following cost function

$$J = \sum_{i=1}^{i=n} |\theta_{in} - \theta i| \tag{4}$$

Where θ_{in} in is the new heading angle for each aircraft.

Another interesting approach outlined in [13] is to evaluate mid-term air traffic complexity in 3-D airspace through a probabilistic measure of the airspace occupancy. The key feature of this approach is that uncertainty in the future aircraft positions is explicitly accounted for when evaluating complexity and the complexity measure was used to design trajectories by minimizing a cost function that is related to the complexity measure.

2.6 Intrinsic complexity measures

Intrinsic complexity measures aims to quantify the level of disorder as well the organization structure of the air traffic distribution, irrespective of its effect on the ATC workload. The previous metrics defined above do not take into account the intrinsic traffic disorder which is related to the complexity. The first efforts related to traffic disorder can be found in [7]. This paper introduces two classes of metrics which measure the disorder of a traffic pattern, both based on the measurements of the aircraft velocities and positions. The first class consists of a geometrical approach where complexity is a function of the relative position vectors and relative velocity vectors of the aircraft. This metric can extract features on the traffic complexity such as proximity (measures the level of aggregation of aircraft in the airspace), convergence (for close aircraft, this metric measures how strongly aircraft are closer to each other) and sensitivity (this metric measure how the relative distance between aircraft is sensible to the control manoeuver). The second class is a dynamic system model of the air traffic and use the topological Kolmogorov entropy as a measure of disorder of the traffic pattern. It can outline traffic flow organization [6].

Topological entropy was further extended where linear models were used to extract the topological entropies measures of the complexity. The limitations of the linear modelling based approach is that it provides only a measure of the global tendency of the traffic, and does not fit exactly with all traffic situations [6]. Nonlinear extension can be used to produce maps of local complexity thus allowing for the identification of critical air traffic areas [3].

Inspired by [6], the key idea is to find a dynamical system which models a vector field as closely as possible to the observations given by the aircraft positions and speeds. Based on this dynamical system model, a trajectory disorder metric can be easily computed. The approach is to interpolate a velocity vector field which is determined based on a snapshot of the air traffic, with each aircraft represented by a point at a certain position and with a certain velocity [3]. The interpolating vector field should satisfy some constraints related to maneuvers feasibility (minimum and maximum speed, and continuity to limit acceleration and turn rate). If a smooth vector field is found, aircraft can follow non intersecting trajectories and the introduction of an additional aircraft causes a marginal increase in complexity. If no smooth solution is found, then the continuity constraint is relaxed, which leads to the introduction of a separation boundary where the vector field loses continuity.

In [3], complexity was evaluated based on the representation of the resulting vector field. The location of the separation boundary corresponds to critical areas. The main challenge of the approach is computing the separation boundary in real-time. On the other hand, the authors extended their work to a non linear representation of their previous dynamical system model. Such a model is exact and can fit the observations without error. The nonlinear model can be used to produce maps of complexity by identifying areas with high and low complexity. This extension can also address time extension of the model and can then work on the trajectory segments instead of speed vectors. A collocation techniques has been used to speed up the computation of the associated complexity, mainly when such a computation is done on large areas like countries [3].

Other approaches considered was using a nonlinear point mass model for an aircraft and linearizing the model under certain operating points and deriving the complexity measure using lyapunov exponents. The main advantage is that it only requires to $n(n+1)/2$ ordinary differential equations in comparison to the approaches that used Singular Value Decomposition (SVD) proposed in [11] which is computationally more complex. The main issue with the Lyapunov exponent theory is that positive largest Lyapunov Exponent doesnt, in general, indicate chaos and negative largest Lyapunov Exponent doesnt, in general, indicate stability.

3 Discussion

In the above sections a brief description was given on complexity measures in the literature. It is mainly based on the current ground based ATM system. The main difficulty in complexity metric evaluation is that the problem is not well posed, and a that the term "ATC workload" can be very confusing and not very well defined [5]. Bearing in mind that the relationship between complexity and ATC workload is an indirect one and moreover that there is a swath of resolution advisories that ATC can use, but they usually choose the most economical advisories when faced with a high traffic situation. On the other hand, ATC workload is a direct function of the task load and most studies found a positive correlation. Unfortunately, not all factors were considered which makes the comparison somewhat difficult. So, in this case a workload measure is important in order to determine how complexity is evaluated [5]. An indirect approach as outlined in the above techniques can prove to be somewhat sufficient, for example structural dependence on the sector and sector based measures such as interval complexity may give an indication of the perceived workload of ATC. Currently workload measures are carried out in a direct (using questionnaires, interviews, etc...) and indirect (key strokes, radio communication with pilots, frequency of conflicts and resolution advisories, etc...) manner.

Control dependency on the approaches is another factor to be considered. The input-output complexity metric is clearly a control dependent approach since it measures the control effort required. While approaches such as intrin-

sic complexity and fractal dimensions are not. Approaches that are not control dependent do not require knowledge of ATC workload, and are indirectly accounted for. This can be quite useful if free flight concepts are considered or that partial delegation is given to the pilot which eventually will lead to all aircraft communicating with each other. In this case ATC will have a time varying decentralized structure which is difficult to quantify the control effort required. Hence a control-independent approach can be more suited for free flight concepts and partial decentralization of ATC tasks.

4 Proposals and future work

Section 2 provided a brief overview of the current techniques in the literature and a discussion was carried out in section 3. Most of the techniques described above do not consider trajectory flexibility and robustness metrics. They may have the potential to alleviate some of the problems associated with them. Also, very few studies deal with the fact that ATM systems run in closed loop, i.e. generating complexity maps while also relying on some form of conflict resolution algorithm.

Some of the unanswered question are:

- How different is the behaviour of the closed-loop system from the behaviour of the open-loop system when complexity measures are considered and, is there a way to accurately model the influence of the conflict avoidance algorithm on the conflict with respect to complexity?
- Does the complexity metric give an indication on whether there is enough flexibility to proceed safely?
- Can I design my own trajectory to increase my flexibility and not adhere to the rigid structures of waypoints?
- Do I need to avoid this airspace entirely and replan?

Another solution that can prove to be useful in evaluating traffic complexity is using fast multipole methods [1,2,4] combined with Lyapunov exponents [3]. Fast multipole methods treat each aircraft as a particle with each particle having a potential and a force field. This algorithm is useful in systems involving large numbers of particles whose interactions are Coulombic or gravitational in nature. For a system of N aircrafts, an amount of work of the order $O(N^2)$ has traditionally been required to evaluate all pairwise interactions of particles.

Other metrics such as trajectory flexibility and robustness metrics can be considered. Even though these metrics do not consider complexity explicitly, they can enable an aircraft to plan its trajectory such that it preserves a required level of maneuvering flexibility in order to accommodate later disturbances caused, for example, by other traffic and weather activity. The concept is that if each aircraft preserves its own trajectory flexibility, using an air-based or ground-based system, acceptable traffic complexity will naturally be achieved. Flexibility preservation complements separation assurance both within the conflict resolution horizon and beyond it to extend the flexibility planning horizon. Within the conflict resolution horizon, flexibility aids in selecting conflict resolution solutions that afford the aircraft more flexibility with respect to the complexity metrics.

5 Conclusion

In this paper, we have presented an extended literature review and provided a discussion of the complexity measures. Some promising techniques that may be applicable to the measure of complexity in an ATM environment has also been proposed. The main challenge and difficulty is to compare the complexity metrics because of the wide variety of indicators. And each measure incorporates a certain feature, some of them are workload related and some are not, even if it is then some measure of workload has to be assumed if the measure were to make any sense in its implementation. The dynamic density-like metrics, Aircraft Density and interval complexity are clearly workload-oriented and sector-based measures of complexity. A chal-

lenging problem posed is to have a reliable measure of workload, but they have to be dependent on some measure of control as well, while other metrics such as intrinsic measures are not.

References

[1] Berman L., "Grid-Multipole Calculations", *SIAM J. Sci. Comput.*, Vol. 16, 1995, pp. 1082-1091.

[2] Carrier J. , Greengard L., and Rokhlin V., "A Fast Adaptive Multipole Algorithm for Particle Simulations.", *SIAM J. Sci. and Stat. Comput.*, Vol. 9, 1988, pp. 669-686.

[3] Delahaye D., Puechmorel S., Hansman J., "Air traffic complexity map based on non linear dynamical systems.", *Air Traffic Control Quarterly*, Vol. 12, no.4 2004, pp. 367-390.

[4] Greengard L. and Rokhlin V., "On the Evaluation of Electrostatic Interactions in Molecular Modeling.", *Chemica Scripta*, Vol. 29A 1989, pp. 139-144.

[5] Prandini M., Piroddi L., Puechmorel S., Brzdilov S.L., "Complexity metrics applicable to autonomous aircraft Version: Final (1.1).", iFly Deliverable D3.1, 2009.

[6] Puechmorel S. "A short introduction to complexity computation." Draf paper, July 2007.

[7] Mondoloni S. and LiangD. "Airspace fractal dimension and applications." In Eurocontrol-FAA, editor, Fourth USA/EUROPE Air Traffic Management R&D Seminar, 2001.

[8] Lee K., Feron E., and Pritchett A. "Air traffic complexity: An input-outputapproach." In American Control Conference, New York City, USA, July 2007.

[9] Kriwan B., Scaife R., and Kennedy R. "Investigating complexity factors in UK air traffic management." In Human Factors and Aerospace Safety, 2001.

[10] Idris H., Vivona R., and Garcia-Chico J-L. "Trajectory-oriented approach to managing traffic complexity - operational concept and preliminary metrics definition." Ames Contract NNA07BB26C NASA/CR-2008-215121, National Aeronautics and Space Administration, Langley Research Center, Hampton, Virginia, 2008.

[11] Hilburn B. "Cognitive complexity in air traffic control: A literature review." Technical Report 04/04, EUROCONTROL Experimental Centre, 2004.

[12] Dieci L., Russel R.D., and Van Vleck E.S. "On the computation of Lyapunov exponents for continuous dynamical systems." SIAM J. Numer. Anal., Vol.34 No.1 pp. 402-423, February 1997.

[13] Prandini M., Putta V., Hu J. "A probabilistic measure of air traffic complexity in three-dimensional airspace." International Journal of Adaptive Control and Signal Processing, special issue on Air Traffic Management: Challenges and opportunities for advanced control, vol. 24 No.10 pp. 813829, 2010.

[14] Pawlak W.S., Brinton C.S., Crouch K., and Lancaster K.M. "A framework for the evaluation of air traffic control complexity." In AIAA national Conference, 1996.

[15] Flynn G.M. , Benkouar A., and . Christien R. "Adaptation of workload model by optimization algorithms and sector capacity assessment." Technical Report EEC Note No. 07/05, Eurocontrol, 2005.

[16] Bridges T.J. and Reich S. "Computing Lyapunov exponents on a Stiefel manifold." Physica D, 156(3):219-238, August 2001.

[17] Blom H.A.P., Krystul J. , and Bakker G.J. "A particle system for safety verification of free flight in air traffic." In 45th IEEE Conf. on Decision and Control, 2006.

[18] Lai C.F., Zelinski S. "Simplified Dynamic Density Based Capacity Estimation". 28th Digital Avionics Systems Conference, 2009.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Looking towards the Single European Sky: a Tailored Security Assessment for Future ATM Systems

F. Matarese and P. Montefusco
SESM – Sistemi Evoluti per la Sistemistica e i Modelli S.c.a.r.l., A Finmeccanica Company, Italy

J. Fonseca
University of Coimbra/Polytechnic Institute of Guarda, Portugal

Keywords: *Security, Methodology, Threat, Vulnerability, Risk.*

Abstract

The objective of this paper is the definition of a new methodology for carrying out security risk assessment in the Air Traffic Management (ATM) domain. This process is carried out by modelling the system, identifying the assets, threats and vulnerabilities, prioritizing the threats and proposing countermeasures for the weaknesses found.
ATM security is concerned with securing the ATM assets, to prevent threats and limit their effects on the overall aviation network. This effect limitation could be achieved by removing the vulnerability from the system and/or increasing the tolerance in case of component failures due to attacks.
The security risk assessment methodology proposed is based on what is currently being done by the industry and international organisations (International Civil Aviation Organization (ICAO), Common Criteria (CC), International Standard Organisation (ISO), EUROCONTROL Guidance Material, etc.) and comprises five main stages.
For demonstrative purposes, the methodology is applied to a case study on the Flight Data Processing Subsystem (FDPS), which is a component of many ATM systems.

1 Introduction

ATM security is concerned with securing the ATM assets (including services), to prevent threats and limit their effects on the overall aviation network. This effect limitation could be achieved by removing the vulnerability from the system and/or increasing the tolerance in case of component failures due to attacks.

Recent ATM vulnerabilities discovered and attacks executed (e.g. refer to [1]), prove that the security of these systems is always under the spotlight, which is also confirmed by public entities (e.g. refer to [2]). Furthermore, the increasing complexity of ATM system(s) due to the pervasiveness of emerging technologies and growing number of daily flights create the conditions for the rise of unpredicted threats that may potentially turn into dramatic events. This is also driven by the on-going update of legacy systems with new technologies and their connection to innovative systems, which creates a new environment with new threat vectors, for which these systems were not prepared when they were designed. Thereby, given that the ATM plays a critical role in supporting the overall airspace/aviation system, the security risk assessment of ATM should be a major concern and a top priority.

Presently, the ICAO Security Manual for Safeguarding Civil Application, one of the main references for threat and risk assessment in the ATM domain, offers little help in identifying and prioritising threats according to time and budget constraints. On the other side, new guidelines are on the way, such as those currently being developed in the Sub Work Package (SWP) 16.2 of the Single European Sky ATM Research (SESAR) project, which focus on ATM security framework, methodology, tools and best practices.

As a consequence, the security risk assessment methodology proposed in this paper can be seen as generic, as it is not bound to any technological or implementation constraints, so it can be applied to most ATM systems. In fact it is based on an abstract model defining assets, threats and vulnerabilities related to any ATM system. In fact, it addresses the following objectives:

- To be adopted either by state-of-the-art ATM systems as well as legacy systems allowing the assessment of the new risks that their interconnection may (and will) introduce.
- To be based on existing and well established safety standards already in use by the industry, including the ICAO, the CC, the ISO 270xx, etc. and extend them to cover the ATM security scenario. For example, although widely adopted, the CC does not provide the procedures that should be used to assess the security of the system, whereas the risk assessment methodology that we present addresses this aspect.
- To be complementary with the risk assessment methodology currently being developed within SESAR and other EU projects in ATM security.

2 Security Assessment Methodology

The proposed methodology is based on what is currently being done by the industry and it comprises five main stages that should be revisited during the development and periodically, after the deployment of the ATM system. The methodology is the synthesis of SESAR/ICAO ATM security guidelines and of Microsoft Threat Modelling for the software related threats:

1. **Assets identification**. The ATM system is formally decomposed using Use Cases or Data Flow Diagrams (DFD) to obtain the list of assets and their interconnections. The technique used is complemented with information about trust (or privilege) boundaries between entities.
2. **Threat analysis**. This stage involves determining the possible threats to each asset identified in the previous stage. The following groups of security attributes are used to obtain and classify the threats: the widely accepted set consisting of Confidentiality, Integrity and Availability, or a more detailed view consisting of Authentication, Integrity, Non-repudiation, Confidentiality, Availability and Authorization. It is also in this stage where the Fault Tree model of the threats of each asset is built.
3. **Vulnerability assessment**. Closely related to the threats, vulnerabilities also drive the respective countermeasures, which will be implemented in the last stage, according to the risk analysis outcome.
4. **Risk analysis**. This allows prioritizing the threat mitigation by directing the resources to the most critical threats first. Risk is a measure of the threat impacts to the system vs. the probability of that threat to occur. Several schemes to obtain the likelihood of the occurrence of the threat may be used, some of them based on the outputs that can already be obtained from the SESAR project.
5. **Countermeasure identification/risk treatment**. This provides the mitigation procedures that need to be executed in order to eliminate the threat or limit its effect to an acceptable residual level. They are closely related to the specific threat they apply to and to the target vulnerabilities. The set of countermeasures/security controls are the most important output from this security

assessment methodology as they can be seen as the recommendations or the security requirements for the ATM under assessment.

2.1 Assets identification

An asset is something of value to the Organisation. In general, technological assets combine logical and physical assets and can be grouped into the following categories:

- **Information.** Documented (paper or electronic) data or intellectual property used to meet the mission of an Organisation.
- **Software.** Software applications and services (such as operating systems, database applications, networking software, office applications, custom applications, etc.) that process, store, or transmit information.
- **Hardware.** Physical devices needed for the proper functioning of the Organisation (such as workstations, servers, etc.). This asset normally focus solely on the replacement costs for physical devices.
- **People.** The people in an Organisation that possess unique skills, knowledge, and experience and that are difficult to replace.

2.2 Threat analysis

An ATM system consists of a set of hardware, software and communication assets, operated by several users with different operation statuses. Threat assessment and risk management together form the basis of a viable and cost effective security response to threats that could target ATM system. One of the most difficult tasks for security professionals is devising an effective security plan that correlates to the threat. Accurately identifying the threat or threats must be the first step in the process. Our challenge is to perform a quantitative analytical approach will be used to perform threat assessment.

In devising a threat assessment methodology, it is preferable to use a systematic and quantifiable approach. Therefore, the threat assessment proposed to evaluate the threats affecting the ATM system uses a quantitative analytical approach. The structure of this methodology employs three core principles of security: *identify*, *implement* and *sustain*.

In undertaking the task of assessing the threats, there are several sources of empirical evidence and statistical data available in the fields of intelligence and security from which to form an analysis of past trends of acts of unlawful interference. In order to provide decision-makers with a current and credible threat assessment, however, multiple sources of information should be explored. Threat and vulnerability criteria have to be determined before conducting the assessment by deciding on *focal points/hot spots*. Focal points can be defined as those factors or criteria that are estimated to have the most weight or value in a given process.

This methodology utilises two facets of analysis that together form a credible means of assessing the threat and determining a security response through application of risk management measures.

First, it must be understood that a *deliberate* act of unlawful interference must, by definition, be premeditated and carried out with *purpose* by the perpetrators. This means that someone has a reason to conduct an unlawful act and thus proceeds to plan and execute the act. Therefore, before assessing how an act of unlawful interference may be carried out against a target, the analyst should first consider the reasons why an unlawful act would be committed and the probability of its being committed.

The next step would be to create a working tool to assist in the assessment process: the *Vulnerability Matrix*. The Vulnerability Matrix forms the final analysis for a follow-on risk management process. It covers security threat categories, which can be adapted to assess the threat directed at a potential target or to evaluate the security posture of a part of the system.

Security professionals have long recognized that implementing increased preventive measures commensurate with a higher level of

threat has an associated expense that may become a heavy financial burden on the resources of an Organisation. It is therefore considered more effective to deploy defences where and when they are most needed rather than applying them universally. This concept is called risk management.

Standards consist of a minimum set of security control measures that are expected to be applied equally at international level regardless of the threat environment impacting on operations. While these arrangements were established to ensure minimum uniform standards, no specific standards exist to address variable threat conditions. Whenever an Organisation introduces additional security measures to meet a higher threat level, it may find that implementation is difficult to sustain, especially when the extra measures have not been tailored to the specific threat. Therefore, once an Organisation has properly assessed the nature and level of threat within its own territory, it can then apply appropriate enhanced measures. Organisations can profit of a risk management approach whereby enhanced measures are implemented either to prevent an unlawful act from being committed or, at a minimum, to mitigate any consequences resulting from an unlawful act.

2.3 Vulnerability assessment

A vulnerability assessment is a systematic, point-in-time examination of an Organisation's technology base, policies, and procedures. It includes a complete analysis of the security of an internal environment and its vulnerability to internal and external attacks.

Technology-driven assessments generally:

- Use standards for specific IT security activities (such as hardening specific types of platforms).
- Assess the entire computing infrastructure.
- Use (sometimes proprietary) software tools to analyze the infrastructure and all of its components.
- Provide a detailed analysis showing the detected technological vulnerabilities and possibly recommending specific steps to address those vulnerabilities.

2.4 Risk analysis

According to the ISO GUIDE 73:2002, "Risk is the combination of the probability of an event and its consequences". Inversely, an enterprise manager should decide to make a financial effort to harden a specific asset if the cost of securing it is less than the risk of loss of the asset. In other words, the manager must be sure that the cost of security in every transaction involving the asset is less than the risk of loss. This is the foundation of security risk management, as detailed by Dan Geer [5].

In fact security can be seen as risk management, because we do not want to spend too much on security, comparing to what assets we are protecting. Many times, the big questions posed in an enterprise when it needs to calculate the budget is the measure of the potential loss and lack of knowledge of where it is likely to occur.

A Threat is, in a general approach, anything that might trigger a Risk. However, it is important to point out that a Threat is not directly connected to Risks. A Threat is effective only if it is connected to a Vulnerability. The Risk is thus dependant on the Vulnerability rather than on the Threat itself. If there is a Vulnerability but there is no Threat using it, the Risk remains. Hence, Threats are mitigated through Vulnerability Analysis over the Assets. According to the Vulnerability Analysis, the Threats can be eliminated or reduced to a point where the value of the Risk is acceptable. The process of mitigating the Vulnerabilities is on the scope of the Security Policies and it is implemented with the Countermeasures. The Security Framework will define the Security Policy and the Risk Management Process to secure ATM system.

At the system level, the risk deliberated can be defined by the following formula:

$$Risk = Likelihood\ of\ the\ Threat * Vulnerability * Consequences\ of\ the\ Exploitation \quad (1)$$

The assessment of likelihood takes into account statistical analyses. The assessments of the consequences in terms of loss of security will be considered as the consequence on the operational reliability in the sense that each threat scenario will be evaluated regarding the consequence of the loss of a corresponding security criteria and cost of the primary asset on the operational reliability.

2.5 Countermeasure identification/risk treatment

The main purpose of any security countermeasure is prevention. Therefore, after the first step to *identify* the threat or threats is completed, the next task is to devise an appropriate security response commensurate with that threat. This task employs the *implement* principle.

If the assumption is made that potential perpetrators with the intention to interfere can defeat a security system if given enough information, time and opportunity, then the logical objective is how best to deter the perpetrators from carrying out a successful act of unlawful interference. It is therefore essential that the implementation of suitable preventive security measures be considered.

This operational intervention leads to the third principle, *sustain*, which can be described as an Organisation having the political will and accompanying capability to maintain appropriate reliable security practices. Without the commitment to *sustain* effective security measures, the efficacy of the other principles is diminished.

Countermeasures/security controls will be identified for risk management. A countermeasure is any system, *passive* or *active*, aimed at resolving a risk occurrence. By nature it is reactive rather than proactive, and is aimed at mitigating the loss due to the risk occurrence. Depending on the nature of the risk and the kind of countermeasure, the risk outcome can be only partially mitigated or totally mitigated.

The security countermeasures identified will be spread over the ATM system architecture [4].

3 Case Study: FDPS

For demonstrative purposes, we will apply our methodology to a case study on FDPS, which is a component of many ATM systems. FDPS is based on a open architecture that manages the flight plan data accepting, processing, updating and distributing the trajectories and related data, according to the aircraft current position. It supports the air traffic controllers during the planning and progress phases of the flight. The safety solution implemented by FDPS is based on redundancy, duplicating FDPS instances and managing this distributed system. However, this increases the attack surface of FDPS and creates new entry points that may not be so well protected, like those related with the management and synchronization between FDPS instances. By increasing the safety of FDPS the current design may also be affecting its security and this is where our security assessment methodology can be applied, contributing to uncover and mitigate these issues.

FDPS provides the processing of flight plan data and other related information to support air traffic controllers during the planning and progress phases of flights.

FDPS is based on an open architecture which provides the processing of flight plan data and other related information to support air traffic controllers during the planning and progress phases of flights [3].

FDPS is capable of accepting, processing, updating, distributing and displaying flight data and other information, according with ICAO requirements and in compliance with European Air Traffic Management Programme (EATMP) requirements.

3.1 Assets identification

FDPS is formally decomposed using Data Flow Diagrams to obtain the list of assets and their interconnections.

Flight plans data can be received via the Aeronautical Fixed Telecommunication Network (AFTN) lines, by means of Air Traffic Services (ATS) messages from eligible sources (e.g. the Integrated Initial Flight Plan Processing

System (IFPS)), via the connections with adjacent ATS units, or by means of notification and coordination messages according to EUROCONTROL Standard for On-Line Data Interchange (OLDI).

Furthermore, authorised operators can enter the flight plans.

Whenever an updating of one or more System Flight Plan (SFPL) Data is performed, FDPS provides the updated data to Control Working Position (CWP) operators by means of SFPL and other external users, such as ATS units and Airport Report Office (ARO), by OLDI and ATS messages.

FDPS receives the Monitoring Aids data from the Safety Nets, and it receives Air Traffic Flow Management (ATFM) and ATS messages from Central Flow Management Unit (CFMU), whilst Meteo messages are received from the Meteo/Aeronautical Information Service (AIS) units.

Warning and SFPL data are provided to the Flight Data Assistant.

Furthermore, FDPS sends the diagnostic towards the Control Management System (CMS), and Special Service Request (SSR) Codes data, Configuration and Sectorisation data are sent to FDP Technical Supervisor.

The FDP system is provided in a redundant highly reliable configuration consisting of two identical instances and of distributed logics. It is provided in a redundant configuration to ensure the radar data processing continuity in the system.

FDPS function is required to receive data from and/or send data to a number of other functions and systems. The following table reports FDPS interfaces:

System	Data IN	Data OUT
METEO/AIS Units	MET/AIS messages.	-
CFMU/IFPS	ATS messages.	-
CFMU/TACT	ATFM messages.	FSA message
Aircraft Operators/ARO	ATS Messages; Free Text Messages.	ATS Messages; Free Text Messages.
Adjacent ATS Units	Notification and Coordination messages according to Eurocontrol Standard for	Notification and Coordination messages according to Eurocontrol Standard for
	OLDI; ATS messages; Free Text Messages.	OLDI; ATS Messages; Free Text Messages.
SNET	Monitoring Aids data.	SFPL Data including CFL and trajectory.
ODS	SFPL data update; EXE and PLN orders; Co-ordination data update.	SFPL data; Co-ordination data; Environment data; Warnings.
Flight Data Assistant	SFPL data; RPL data.	SFPL data; Warnings.
FDP Technical Supervisor	Sectorisation data; Configuration data; SSR Codes data.	Sectorisation data; Configuration data; SSR Codes data.
CMS	Configuration orders.	Diagnostics.

Table 1 Data Flow in/out for FDPS

3.2 Threat analysis

FDPS *Hardware security threats* will be investigated. First of all, hardware assets will be categorised, then the following main sources of threats will be analysed:

- *"Physical attack"*

An attack aimed at interrupting, disturbing or in any case damaging the infrastructure. The basic key point is that the attack is done in the physical domain rather than in the information domain.

- *"Environmental threats"*

It can be classified as a special case of Physical attacks, whereas the point is that the threat can also arise from natural causes. Typically, this is the case for climatic phenomenon seismic phenomenon, meteorological phenomenon or flood, which can directly lead to physical damages like fire, water, pollution, major accident, destruction of equipment of media, dust, corrosion, freezing. As a secondary consequence, events like loss of power, failure of telecommunication equipment, electromagnetic or thermal radiation may occur that can bring down electronic and computing systems.

ATM *Software security threats* will be investigated. First of all, software functionalities will be categorised, then the following main sources of threats will be analysed:

- *"Intrusion"*

Any form of attack that leads the attacker to gain unauthorized access to one of the ATM subsystems. The attack can be performed in a number of ways, mainly dependant on software and protocols bugs and vulnerabilities.

ATM *Information security threats* will be investigated. First of all, communication assets and functionalities will be categorised, then the following main sources of threats will be analysed:

- *"Data corruption and stealing"*

It can arise from two different events:

a) Communication security failure
b) System security failure

The first is a consequence of an attack aimed at the communication infrastructure, hence on the data being transmitted. The second kind arises from an attack to a working server or client, i.e., an intrusion.

- *"Identity usurpation"*

It is usually the consequence of a successful attack either at communication or system level, i.e., data stealing or system intrusion. The usurper can use the stolen identity to perform actions of systems that, at first, might seems perfectly legit.

3.3 Vulnerability assessment

The following Computer Software Configuration Items (CSCIs) are included within the software architecture of FDPS:

CSCI	Description
FTF	Fault Tolerance Function: provides services that can be used to handle automatic fail over of applications and data files associated even in the event of operating system, services or hardware failures; when the active node fails, its resources are transferred to the standby node.
XSD	Advanced System Message Dispatcher: provides dispatching of commands and diagnostic messages. XSD produces Node Status Messages using information collected from those CSCIs running inside the same node: • SPV sends information about system peripherals and devices status; • other CSCIs send their internal logical status. XSD collects these data and then sends the node status message periodically to the other system nodes.
SPV	Supervisor: supports for node and process start-up/shutdown, provides the information about the peripherals status and the Node Role (Master/Stand-by) and manages the operator system console in order to view diagnostic messages and to issue commands.
CDB	Common Data Base: handles the whole system configuration, manages the dynamic role of the CWPs and constantly aligns a so called Common Data Base (CDB). CDB is also a Data Base Fault Tolerant and this capability is achieved by its replication on the all system nodes. Consistency of the databases is achieved through a Best Node.
AFS	Advanced FDP Server: responsible for the processing of the core functions of FDPS, including the environment data handling, the flight data processing and distribution, the message handling and data exchange with the other subsystems.
AFJ	Advanced Flight Data Processing Java M.M.I.: provides eligible operators with an HMI supporting the following functionality: • Environment data management, where these information are separated in dynamic data, geographical data and configuration of OLDI and ATS units. • SFPL data management. • Message handling including: generation of one message(OLDI and ATS format), management of wrong and rejected message. • RPL management. • System administration (e.g. line configuration) • Archived data/messages inspection • SSR Code configuration
IOL	Input/Output LAN: provides a set of basic point-to-point and multipoint communication services allowing the exchange of messages among CSCIs of different nodes.
IKS	Internal Kernel SPV: provides communication services allowing the exchange of messages from CSCIs toward LAN. The high level protocol used to communicate over the LAN connecting the computers is the User Datagram Protocol (UDP).

Table 2 FDPS' CSCIs

Only the following CSCIs represent a vulnerability for FDPS: AFS, CDB and IKS, considering the criticality and the impact on the system if affected by malicious attacks.

3.4 Risk analysis

There is no statistic data available related to FDPS attacks to justify a likelihood analysis. For this reason, risk will be evaluated considering just the impact of potential threats on the system and assuming the probability equal to 1 (i.e. 100%). Countermeasures are so identified, initially, on the basis of the threat analysis and the architecture of the system.

The following table reports FDPS risk analysis:

CSCI	Threat	Local Effect	System Effect	Severity
AFS	Data corruption and stealing: loss of message coming from external networks (ATS, meteo, OLDI messages)	Inability of communicating with external networks. (ATS, meteo, OLDI messages). ATC controller is aware of this. Increased workload.	Loss of message flight data exchanged with external networks	Significant
AFS	Data corruption and stealing: loss of data coming from IOL CSCI	Loss of tracks data. Inability of updating trajectories. ATC controller is not aware of this and continues working with existing flight data.	Corruption of flight data exchanged with CWP	Major
AFS	Data corruption and stealing: undetected corruption of data coming from CDB CSCI	The corrupted message is not recognised. Incorrect activation of flight plan data; incorrect entry of new flight plan data; incorrect cancellation of flight plan data.	Corruption of flight data exchanged with CWP	Major
AFS	Data corruption and stealing: corruption of message toward external networks (ATS, meteo, OLDI messages)	Corrupted messages are checked and discarded by receivers. Inability of communicating with external networks. (ATS, meteo, OLDI messages). ATC controller is aware of this. Increased workload.	Loss of message flight data exchanged with external networks	Significant
AFS	Intrusion: overload of messages that causes a memory leak	Flight data not available at CWP. Loss of automatic advance warning of active flights. Inability of activation of flight plan data; inability of entry of new flight plan data; inability of cancellation of flight plan data.	Loss of flight data exchanged with CWP	Major
AFS	Intrusion: corruption of static data in internal AFS CSCI database	Corruption of static data contained in AFS CSCI database. Incorrect distribution of static data for display. Incorrect flight data references may result in inappropriate actions being taken by the CWP operator.	Corruption of flight data exchanged with CWP	Major
AFS	Identity usurpation: loss of connection with internal AFS CSCI database or algorithm failure	Flight data not available at CWP. Loss of automatic advance warning of active flights. Inability of activation of flight plan data; inability	Loss of flight data exchanged with CWP	Major

CSCI	Threat	Local Effect	System Effect	Severity
		of entry of new flight plan data; inability of cancellation of flight plan data.		
CDB	Data corruption and stealing: loss of received message from AFS CSCI	The common database is not updated.	No effect	Significant
CDB	Data corruption and stealing: detected corruption of received message from AFS CSCI	The corrupted message is discarded and the common database is not updated.	No effect	Significant
CDB	Data corruption and stealing: undetected corruption of received message from AFS CSCI	The corrupted message is not recognised and the common database is contaminated.	Corruption of flight data exchanged with CWP	Major
CDB	Data corruption and stealing: loss of sent message toward AFS CSCI	AFS doesn't receive data from CDB but is aware of it. Inability to activate flight plan data, to enter new flight plan data and to cancel flight plan data.	Loss of flight data exchanged with CWP	Major
CDB	Data corruption and stealing: loss of sent message toward CWP	ATC controller doesn't receive updated data from FDP system but is aware of it. Flight data not available at CWP. Loss of automatic advance warning of active flights.	Loss of flight data exchanged with CWP	Major incident
CDB	Data corruption and stealing: undetected corruption of sent message toward CWP	The corrupted message is not recognised. Incorrect distribution of data for display. ATC controller is not aware of this. Incorrect flight data may result in inappropriate actions being taken by the CWP operator.	Corruption of flight data exchanged with CWP	Major
CDB	Data corruption and stealing: undetected corruption of sent message toward AFS CSCI	The corrupted message is not recognised. Incorrect activation of flight plan data; incorrect entry of new flight plan data; incorrect cancellation of flight plan data.	Corruption of flight data exchanged with CWP Loss of message flight data exchanged with external networks	Major
CDB	Intrusion: overload of messages that causes a memory leak	Flight data not available at CWP. Loss of automatic advance warning of active flights. Inability of activation of flight plan data; inability of entry of new flight plan data; inability of cancellation of flight plan data.	Loss of flight data exchanged with CWP	Major
IKS	Data corruption and stealing: generic internal failure, inability to manage hardware resources	Flight data not available at CWP. Loss of automatic advance warning of active flights. Inability of activation of flight plan	Loss of flight data exchanged with CWP	Major

CSCI	Threat	Local Effect	System Effect	Severity
		data; inability of entry of new flight plan data; inability of cancellation of flight plan data. An alert is sent to CSCI of FDPS. Messages generated are lost.		

Table 3 FDPS risk analysis

3.5 Countermeasures identification/risk treatment

The set of countermeasures are the most important output from this security assessment methodology as they can be seen as the recommendations or the security requirements for FDPS.

According to the risk analysis, the following countermeasures can be identified in order to limit the effects of attacks that cause corruption of data:

- Syntactic and semantic check algorithms of AFS and CDB CSCIs.
- Syntactic and semantic check algorithms of CWP.

Regarding the loss of data, no countermeasure can be identified internal to FDPS. Other measures can be identified to protect ATM system from intrusions, as encryption and decryption algorithms passwords.

4 Conclusion

The objective of this paper is the definition of a new methodology for carrying out security risk assessment in the ATM domain. This process is carried out by modelling the system, identifying the assets, threats and vulnerabilities, prioritizing the threats and proposing countermeasures for the weaknesses found.

For demonstrative purposes, we have applied our methodology to a case study on FDPS, which is a component of many ATM systems.

The results are:

- the identification of assets, as services given by the system,
- the analysis of threats, as potential attacks,
- the assessment of vulnerabilities, as CSCI of the system are vulnerable because remotely accessible,
- the analysis of risks, considering the effects of successful attacks,
- and, finally, the identification of countermeasures to limit those effects.

The proposed methodology allows the identification of countermeasures in a systematic way. Countermeasures can be adopted as security system requirements at design level. Nevertheless, not all the countermeasures to protect data can be applied, and identified, at subsystem level, so the security assessment has to be performed also at a higher (system) level.

References

[1] Federal Aviation Administration, *Review of Web applications security and intrusion detection in Air Traffic Control Systems*, U.S. Department of Transportation, 2009.

[2] SESAR Definition Phase Project, *D1 Air Transport framework – The current situation*, v3, SESAR Consortium, 2006.

[3] International Civil Aviation Organization, *Procedures for Air Navigation Services - Rules of the Air and Air Traffic Services*, ICAO, Thirteenth Edition, 1996.

[4] International Civil Aviation Organization, *Security Manual for Safeguarding Civil Aviation Against Acts of Unlawful Interference*, ICAO, Sixth Edition, 2002.

[5] Geer, D., *Risk management is still where the money is*, Computer, *36*(12), 129-131, doi:10.1109/MC.2003.1250894, 2003.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013), in the frame of Marie Curie Industry-Academia Partnerships and Pathways Call (FP7-PEOPLE-2008-IAPP), under Grant Agreement No. 230672 ("CRITICAL-STEP" Project).

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

THE RELATIONSHIP BETWEEN AIR TRAFFIC CONTROL AND COST INDEX

Birsen YORUK ACIKEL and Aydan CAVCAR
Anadolu University, School of Civil Aviation, Eskisehir, Turkey

Ugur TURHAN
Anadolu University, School of Civil Aviation, Eskisehir, Turkey

Abstract

The aviation authorities expect that air traffic volume and complexity will continue to increase and the demand for (ATC) services. The capacity of air transportation system has not kept pace with the growth in demand. Thus costs in terms of lost time and extra fuel consumed continue to grow rapidly for aviation sector. In the face of these developments, air traffic management system (ATM) in pursuit of increasing aircraft operational efficiency, better utilization of the airspace, and decreased environmental impact.

These paper presents use of cost index in air traffic management and its effects on flight time and fuel consumption are given and discussed by real-time radar simulation results.

1 General Introduction

In the EUROCONTROL Statistical Reference Area (ESRA) in 2015, Mid-Term Forecast is that there will be 11.7 million IFR(movements, 16% more than in 2008. The traffic is forecast to decline by nearly 5% in 2009 and to continued growing only slowly in 2010 (growth of 1.5%). In the later years, the growth recovers to more typical rates of 3.5%-4.5% per year [1].

However air traffic in the world has substantially increased and with this development will increase the demand for ATC services. The capacity of air transportation system has not kept pace with the growth in demand. Thus costs in terms of lost time and extra fuel consumed continue to grow rapidly for aviation sector. In the face of these developments, air traffic management system is affected by increasing aircraft operational efficiency, better utilization of the airspace, and decreased environmental impact.

Additionally, in 2008 a Safety Alert due to an increased range of observed speeds in the airspace between identical aircraft types was issued by Eurocontrol. Therefore, The Aircraft Operators and Air Navigation Service Providers were invited to share their experiences regarding the appropriate reasons and consequent effects on the ATM system. Responses clearly pointed out the current lack of information regarding the cost index (CI) based flight planning process is respectively the associated effects on flight profiles. The resulting uncertainties in trajectory prediction affect the provision of the ATC services regarding increased controller workload and reduced capacity as well as potential effects on the safe separation of aircraft. In view of the modernization of the ATM system within the next decade and the aspired service-oriented approach the

appropriate challenges will even rise. Airspace Users' requirements including their wish to fly close to the optimum trajectory will strongly influence the future ATM system [2].

2 Cost Index

The cost index is a parameter set in the cockpit, which determines how the Flight Management System (FMS) will direct the aircraft's speed. It quantifies the choice to fly faster to recover delay, or to fly slower to conserve fuel [3]. In this context, the concept of cost index to be known and taken into account by also from the inside of ATM will benefit the entire air transport sector.

Cost indexing relies on knowledge of aircraft intent and being able to vary aircraft speeds to arrive to the destination at economically optimized times [3]. Defining flight profiles for different cost indices enables efficient flying operations in different conditions. The cost index should be changed to catch slots and avoid delays with consequences like missed connections, dissatisfied customers, airport congestion etc.

The concept of cost index is expressed mathematically as follows:

$$CI = \frac{TC}{FC} \qquad (1)$$

Where *CI* stands for cost index, *TC* is the cost of time (€/min) and *FC* is cost of fuel (€/kg). The unit of cost index is usually given in kg/min [4**]**. The cost index is the key input value for the calculation of the speed and the vertical trajectory based on the most economical in-flight performance. Generally it is given to the pilot within the briefing package provided by the dispatch and entered into the FMS as part of the flight preparation. The flight profile calculation is done by the aircraft's integrated FMS. The calculation process is performed immediately before block-off time and, if required, during the flight in order to adapt the in-flight performance when conditions are changing. Consequently, a trajectory is calculated that balances the costs of time and fuel in order to minimize the sum of all direct operational costs.

2.1 Calculation of the Cost Index

Determination of the cost index value should be based on a careful cost analysis. When designing the flight schedule in an airline's operation, the value placed on time is relatively high, because the faster an aircraft flies the more passenger-miles it produces in a working day. Once operating a fixed schedule, a few minutes gained or lost during flight, usually cannot be converted into productivity. The effect of productivity is therefore excluded from the analysis. The time-related costs are then limited to those maintenance tasks which are a function of flight time, and possibly a portion of the crew's pay.

As already explained, the determination of CT requires careful consideration by the airline. It is not always obvious which cost elements should be included in C_0 and CT.

For instance, crew cost may be based on scheduled block time or actual block time. For the latter condition it is clear that a gain in actual flying time reduces crew cost for a given trip and shall therefore be part of the time-related cost. For the former condition, reducing actual block-time does not affect the crew cost and shall therefore be part of the fixed trip cost. Another example is maintenance cost. The time related maintenance cost per trip can be reduced by decreasing the block time. However, this is usually accomplished by flying faster and thus thrust will have to increase. The higher thrust levels may cause increased wear and tear, which will have an impact on the overhaul cost of the engines, unless "power-by-the-hour" arrangements have been agreed upon with an engine maintenance provider.

Also, as stated before, the effect of a change in productivity (i.e. an additional flight in a working day when sufficient time is gained) is not included. This assumption seems to be reasonable as it is usually impossible to schedule another flight on such short notice. However, one can imagine that a flight has to be

cancelled because the previous flights of the day were flown at low speed to save fuel. Also, passengers missing their connection due to late arrival due to fuel saving policies may cause huge additional losses far higher than the fuel cost gained.

As an illustration, a sample calculation of C_0, CT and CI is presented, based on the following assumptions:

Number of trips: 2000 per year
Fixed cost: 1,000,000 € per year (ownership cost i.e. depreciation, interest, insurance)
Crew cost: 600 € per block hour flown (cabin + cockpit)
Maintenance cost: 800 € per hour
Direct expenses: 1500 € per trip (navigation charges, landing fees, ground handling costs)
Fixed cost per trip:

$$C_0 = \frac{\text{fixed cost per year}}{\text{number of trips per year}} + \text{direct expenses} \qquad (2)$$

$$C_0 = \frac{1{,}000{,}000}{2000} + 1500 \qquad (3)$$

C_0 = 2000 € per trip.
Cost per unit of trip time:
CT = crew cost + maintenance cost per hour
CT = 600 + 800 = 1400 € per hour or 23.3 € per minute

Fuel Cost:
The cost per unit of fuel is assumed to be: CF = 0.73 € per kg.

Cost Index:
The Cost Index CI is then calculated as 23.3 / 0.73 kg/min = 32 kg/min

Note that when the crew cost is defined as a fixed value per scheduled block hour instead of per flight hour, the value of CT will change from 1400 into 800 US$ per hour and as a consequence CI will change from 32 kg/min into 18 kg/min [5].

2.2 Cost Index and It's Use by FMS

Most modern transport aircraft are equipped with a FMS that assists the pilot with flight planning and automatically controlling the aircraft. The FMS blends all navigation data on the aircraft from various sources to maintain an accurate status of aircraft position, velocity, and atmospheric conditions. The pilot enters the flight plan into the FMS, including lateral route and vertical requirements, and the FMS predicts the full aircraft trajectory including altitude, speed, fuel, and time [6].

The FMS performance function ECON (economic speed) allows the computation of a flight profile that yields the lowest possible cost for a given trip by means of the Cost Index. The Cost Index allows the Flight Management System to generate cost optimized flight profiles independent of the individual values of TC and FC. After activation of the FMS performance function ECON, the cost optimization process is started. The optimum cruise speed (Mach) is obtained from a table in the FMS. This optimum Mach is a function of altitude, aircraft weight, wind speed and cost index, and has been pre-calculated during an off-line study [7].

2.3 Econ Speed and Operating Costs

The crucial parameter when balancing time-related and fuel costs is the speed. Depending on the entered CI the Flight Management Computer (FMC) calculates the most economic (ECON) speed for every phase of the flight. For the minimum CI boundary of 0 time costs are neglected and fuel costs are reduced to minimum. In this case the ECON speed will equal Maximum Range Cruise (MRC) speed. For a high CI the ECON speed increases in order to reduce time costs and a speed up to the operational limitation of the aircraft is possible [8]. Fig. 1 presents both fuel and time-related costs depending on the speed.

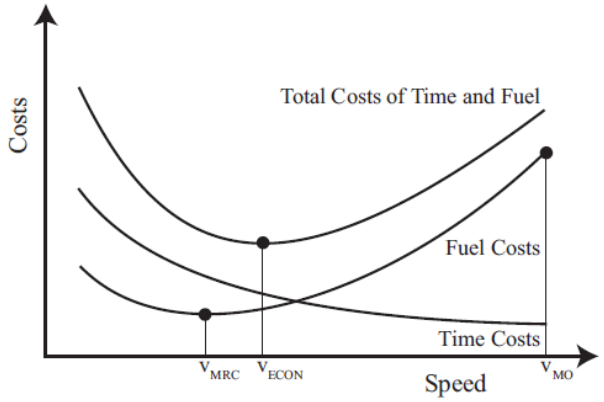


Fig. 1 Cost as a function of speed[8]

According to this figure flying with speed below or above to the ECON speed corresponding to the optimum CI will cause an increase on the total costs. For low speed the fuel savings will not compensate the inevitable higher time costs and vice versa for higher speed. It must be stated that on closer examination the relation between speed and operating costs is more complex because the ECON speed varies significantly in dependence on the flight conditions. Fuel cost curve is dependent on the gross weight of the aircraft, assigned flight level (FL), the air temperature as well as the wind conditions. Time cost curve is dependent on ground speed and with it strongly influenced by the wind. This leads to shifting curves in dependence of the mentioned parameters and with it variations of the ECON speed. However, since during the flight planning process as well as the communication between the pilots and air traffic controllers normally indicated air speed (IAS) or Mach number is used, the upcoming results in Section IV refer to these speeds as well (by default with no wind and no deviation from ISA conditions). For a comprehensive understanding of the cost index concept, it is essential to have a closer look on all operating costs with effects on the in-flight performance of an aircraft [2]. In the following, a short overview on the entire cost structure of a typical airline operator is provided in order to identify all relevant costs related to a certain flight operation and in particular those costs directly related to in-flight performance.

Total airline operating costs consist of direct and indirect costs illustrated in Fig. 2 [9].

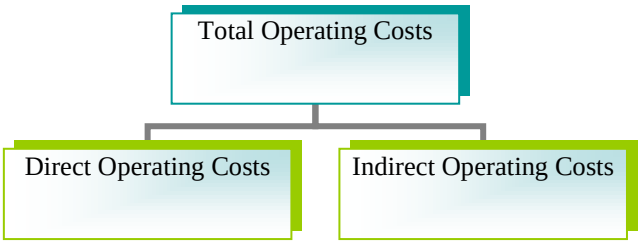


Fig. 2 Airline operating costs

Direct operating costs (DOC) are all those expenses associated with and dependent on the type of aircraft being operated, including all flying expenses, all maintenance and overhaul costs, and all aircraft depreciation expenses [10]. On the contrary indirect operating costs are independent and not connected to the operation of an aircraft mainly including expenditures for administration and distribution. Direct operating costs consist of fixed and a variable part costs. Fixed direct costs are related to the operation of the aircraft but cannot be influenced by the flight operational itself. These are costs for depreciation, insurance and the fixed part of maintenance and crew costs. On the contrary all variable direct operating costs are directly addressable to the flight operational. Therefore, a higher share of variable DOC regarding all operating costs enables an increased cost control in the frame of the flight planning process [2].

Indirect operating costs are all those costs that will remain unaffected by a change of aircraft type because they are not directly dependent on aircraft operations, including expenses that are passenger related rather than aircraft related (such as passenger service costs, costs of ticketing and sales, and station and ground costs) and general and administrative costs [10].

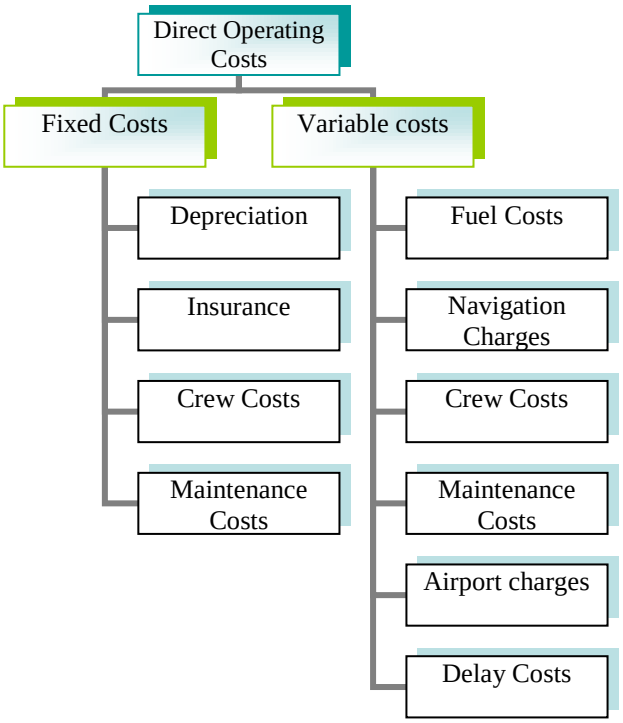


Fig. 3 Direct operating costs [2]

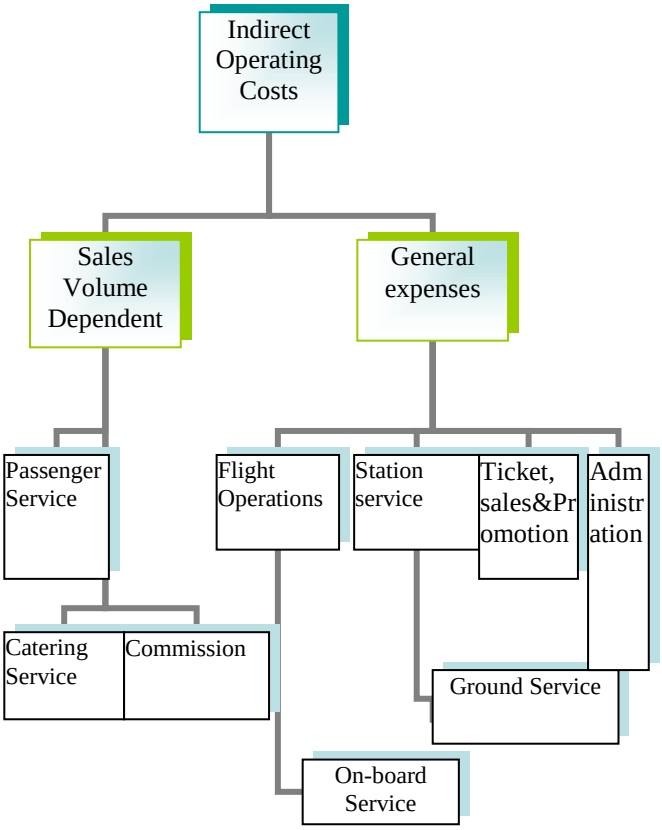


Fig. 4 Indirect operating costs [2]

3 Cost Index on the Effects of ATM

Due to rising fuel prices, airlines are considering the use of the cost index very useful. With the increasing number of aircraft, flight profiles cannot be predicted by the controllers. ICAO flight plan includes information concerning the cruising speed and as such provides particular information, only speed changes of more than 5% have to report to ATC [2]. In addition, depending on the uncertainty of cost index will cause significant changes in the vertical speeds. The Single European Sky ATM Research Programme (SESAR) trying to overcome its with new project and system.

3.1 Simulator Application and Results

In simulation application that carried out A321 aircraft selected as an example aircraft. For selected aircraft three different CI scenarios applied with two different flight distance. In these different cost index scenarios different flight level and cruise speeds chose. Application shows that, increasing CI values decreases flight time, decreasing CI values increases flight times. At the same time consumed fuel quantities changes inversely against flight time. Using same cost index and speed with different flight level resulted different flight times. Thus, causes aircraft stay air for a longer time. Therefore CI should know by air traffic controllers and response to flight requests with respect to CI is very important.

Aircraft	Distance	CI	Flight Level	Cruise speed	Flight time
A321	243 Nm	25	310	0.76	43min
		60	320	0.78	40min
		999	300	0.80	40min
A321	769 Nm	60	370	0.79	1h51
		25	380	0.78	1h55
		60	390	0.79	1h52

Table 1.Analysed flight

4 Conclusion

This paper has identified to importance of cost index. Operational flight planning is becoming more and more of high importance for the ATC and ATM. Therefore, flight profile

optimization based on cost index provides safe and efficient ATM system and it's necessary for modernization during the next year.

Additionally, future work will continue to refine the approaches presented here to better understand the use of cost index for delay environmental performance assessment of ATM.

References

[1] *Eurocontrol Mid-Term Forecast: IFR Flight Movements 2009-2015* (http://www.eurocontrol.int/statfor/public/standard_page/forecast_reports.html).

[2] Rumler W., Günther, T., Fricke,H., and Weishaar, U., "Flight profile Variations due to the Spreading practice of Cost Index Based Flight Planning", *4th International Conference on Research in Air Transportation*, Budapest, 2010.

[3] Cook A., Tanner, G., Williams, and V., Meise, G., "Dynamic cost indexing", *Journal of Air Transport Management*, 2008.

[4] Snare A.S., "Cost optimization of flight profiles and tracks for a regional jet", Master Thesis, Helsinki University of Technology, August, 2006.

[5] Fly Fokker Fuel and Enviromental management, A Fokker Service Suplement, January, 2010.

[6] Jackson M.R.C., Sharma V., Haissig C. M. and Elgersma M., Airborne Technology for Distributed Air Traffic Management," Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on Seville, Spain, December 12-15, 2005.

[7] Boeing, Fuel Conservation Strategies Cruise Flight, AERO Quartly, Seattle, Vol. Quarter 4, 2007.

[8] Airbus Industrie, Getting to Grips with Fuel Economy, 2004.

[9] Wei W. and Hansen M., "Cost economics of aircraft size", *Journal of Transport Economics and Policy*, Vol. 37, Part 2, May 2003, pp. 279-296.

[10] Wensveen J. G., *Air transportation a management perspective*, 6 th ed., Ashgate, 2007, Chaps. 10.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

L-band Compatibility of LDACS1

S. Gligorevic, N. Schneckenburger, N. Franzen, M. Schnell
German Aerospace Center (DLR), Oberpfaffenhofen, Germany

S. Müller, H. Günzel
DFS Deutsche Flugsicherung GmbH, Germany

Keywords: Aeronautical communication, interference, LDACS, DME

Abstract

In order to cope with the increasing demand of communication capacity in the aeronautical sector, the Future Communications Infrastructure has been developed. For air ground communications currently two candidates are considered for the L-band digital aeronautical communication system LDACS. Both L-band systems use frequency bands assigned to both civil and military navigation systems. Hereby of special interest is the distance measurement equipment due to its wide and extensive use in civil aviation. Prior to the deployment of any of the candidates, the compatibility towards those legacy systems has to be confirmed. This paper presents the result obtained during compatibility measurements of LDACS1 carried out at labs of the German Air Navigation Service Provider, DFS Deutsche Flugsicherung GmbH, in March and September 2011. The paper deals with the interference on an airborne DME caused by LDACS1 emissions.

1 Introduction

To enable the modernization of Air-Traffic Management (ATM) as currently pursued by NextGen [1] in the US and SESAR, Single European Sky ATM Research in Europe [2], new and efficient communication, navigation and surveillance technologies are required. For communications, a common understanding within ICAO has been reached that a single data link technology is not capable of covering the communication needs for all phases of flight. Therefore, the Future Communications Infrastructure (FCI) has been developed comprising a set of data link technologies for aeronautical communications [3]. For the airport, AeroMACS (Aeronautical Mobile Airport Communications System) is currently developed within NextGen and SESAR which is strongly based on the WiMAX standard. ESA initiated the development of a future satellite-based communications system for aviation within their ESA Iris program, supplemented by work performed within SESAR. For air/ground communications, currently two L-band Digital Aeronautical Communication System (LDACS) options were identified. LDACS1 option employs a frequency division duplex (FDD) broadband transmission using Orthogonal Frequency-Division Multiplexing (OFDM) [4]. LDACS2 option is a single-carrier system employing time-division duplex (TDD) [5]. In order to provide the final LDACS selection, the radiofrequency compatibility between LDACS and the legacy L-band systems should be proved. The Distance Measuring Equipment (DME) operating as FDD system on a 1 MHz channel grid is the major user of the L-band and its sensitivity to LDACS interference is the prioritized test scenario [6]. Whereas LDACS2 is expected to operate in the DME free frequency band between 960-975 MHz, LDACS1

additionally offers the opportunity to use spectral gaps between existing DME channels, thus increasing the potential number of communication channels. This inlay deployment option for LDACS1 is the most interesting but also the most critical case regarding its interference on DME as well as the interference from DME on LDACS1.

Current work on LDACS, performed under the framework of SESAR within the corresponding SESAR project P15.2.4 "Future Mobile Data Link System Definition", aims to set the evaluation criteria for L-band compatibility testing, to define the measurement set-up for LDACS evaluation, and develop laboratory prototypes to be used in the compatibility tests. Early tasks of the P15.2.4 project produced documents that summarize different interference scenarios and the coexistence criteria. The interference cases where LDACS transmitter (TX) acts as interference source are covered by [7] and the cases where LDACS receiver (RX) is a victim system are reported in [8]. The interfering as well es the victim system can be a ground station (GS) or an airborne station (AS), where different constellations result in different distance ranges between the two systems.

Parallel to the initial SESAR activities, DLR has implemented an LDACS1 physical layer laboratory demonstrator in FPGA technology based on the current LDACS1 specification [4]. The demonstrator enables investigations of both the influence of the LDASC1 waveform on the legacy L-band systems and the interference of the legacy L-band systems on the LDACS1 receiver. Independently from the SESAR activities, DLR in cooperation with the German air navigation service provider Deutsche Flugsicherung GmbH (DFS) started to perform a set of preliminary interference measurements. These test measurements have the goal to assess the sharing conditions between LDACS1 and DME in terms of acceptable levels of the interfering signal as a function of the frequency separation. The first conducted test measurements reported in this paper comprise only an interfering LDACS1 ground or airborne transmitter and a victim airborne DME receiver. The following section describes the interference scenarios relevant for these constellations. The compatibility criteria for DME are explained in Section 3 and Section 4 describes the spectral properties of the LDACS1 signal. Section 5 presents the results obtained in the measurements. A preliminary conclusion on sharing conditions between DME and LDACS1 is given in Section 6.

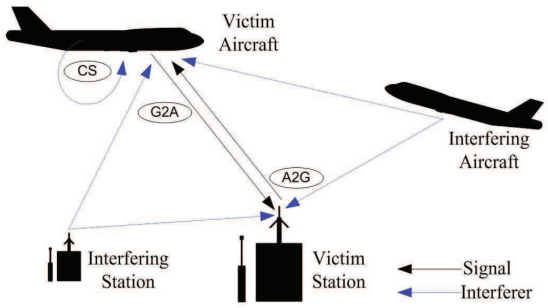


Figure 1: Interference scenarios

2 Interference Scenarios

Fig. 1 shows all theoretically possible interference scenarios between DME and LDACS1 and Fig. 2 the envisaged spectrum allocation of LDACS1 in the inlay deployment option and the DME channel assignment. The relevant cases where LDACS TX acts as interference source are summarized in [7]. However, considering possible deployment scenarios, distances between the interfering and the victim system, as well as the frequency channel usages, the following three scenarios are identified as the most relevant for LDACS1 [6]:

a) LDACS1 GS TX - DME AS RX (G2A)

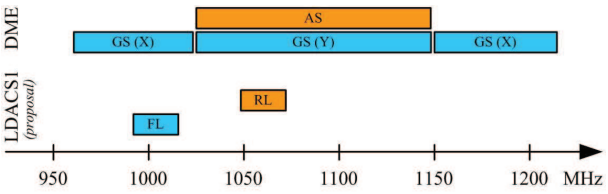


Figure 2: Frequency allocation of DME and LDACS1 in inlay deployment

b) LDACS1 AS TX - DME AS RX (Co-site)

c) LDACS1 AS TX - DME GS RX (A2G)

Case a) represents the situation where signals transmitted by the LDACS1 ground station are being received by an airborne DME station. The closest distance between involved systems is assumed to be as low as 50 m, resulting in a path loss between antennas of 66 dB [7]. The estimated maximum distance in this scenario of 200 nautical miles results in a free space path loss between antennas of approximately 144 dB. LDCAS1 GS transmits continuously with a frequency spacing of $\Delta f = k \cdot 500$ kHz, $k = 1, 2, ...$, from the victim DME system.

Case b) represents the situation where LDACS1 signals transmitted by an aircraft are being received by the DME device installed on the same aircraft. The isolation between two airborne systems on the same aircraft is estimated to 30 dB [7]. DME GS can transmit on any DME channel between 960 and 1213 MHz and hence, any frequency separation between LDACS1 AS TX and victim DME system in multiples of 500 kHz is possible, including co-channel transmission. In general, the victim airborne DME receiver could be influenced by the blocking/desensitisation caused by the strong "in-band" part of the airborne LDACS1 TX signal and its out-of-band products. The performance will then depend upon the selectivity of the DME receiver. Besides, in case of a large frequency spacing between two systems, the radiated LDACS1 TX broadband noise may influence the DME reception. In both cases, the interference impact will also depend on the duty-cycle in the LDACS1 reverse link transmission (see Chapter 4).

Case c) represents the situation where LDACS1 emissions from an airborne aircraft are being received by a DME ground station. Thus, the conditions are similar to case a) with the difference that LDACS1 transmissions are not continuous and therefore, beside frequency spacing, the duty cycle is also a variable test parameter. Due to the current lack of a DME ground station, test case c) could not be measured yet.

3 DME

DME interrogators onboard an aircraft provide the slant range distance between the aircraft and the replying DME transponders on the ground. The slant range distance is measured via determination of the signal traveling time between aircraft, ground and and back to the aircraft. For this purpose the DME interrogator transmits pairs of Gaussian pulses with a defined spacing and the ground transponder replies with another pair of pulses after a determined period of time. During the search mode the maximum allowed repetition rate is 150 pulse pairs per second (ppps) [9]. The aircraft interrogator correlates the replies with the pulses to distinguish, which ones are in response to its own interrogations. This acquisition process should be completed in less than 2 seconds. Once the interrogator obtains lock to the GS, it enters the track mode where the repetition rate is reduced to about 30 ppps.
Typically, the DME devices in use provide much less than the specified maximum number of pulse pairs in the search and the tracking mode. Besides, different DME devices have also different signal processings including filter characteristics and hence show different performance and susceptibility to interference.

The maximum number of pulse pairs transmitted by a DME ground transponder is 2700 ppps in the constant mode. However, some are able to operate with a squitter rate of 3600 ppps under peak traffic conditions. The ground TX operating power is 63 dBm. On the airborne side, the TX power is variable depending of the type of equipment and the maximal possible power setting is 63 dBm. For laboratory test purposes, the ground transponder was replaced by the DME Ground Station simulator JCAir SDX 2000 with the reply frequency set to 2700 ppps.

A general criteria for testing DME ground transponder receiver is the Beacon Reply Efficiency (BRE). BRE represents the ratio of the number of sent pulse pairs to the number of received interrogation pulse pairs. Interference will lower the amount of interrogations which

are successfully detected and therefore reduce the BRE. However, only the JCAir SDX 2000 as ground station simulator was available up to now, which is an inappropriate ground station under test. Therefore, the interference tests have not been performed yet for this scenario.

According to the compatibility criteria proposed in [7] for the DME airborne receiving equipment in the presence of undesired interference, the compatibility tests for DME comprises recording of the following values:

- TTA (Time To Acquire)
- ASOP (Acquire Stable Operating Point)
- BSOP (Break Stable Operating Point)

TTA is defined as the time an interrogator needs to acquire a stable track when the desired signal level is set to the minimum specified value. In [10], a TTA of around 50 s is required for DME devices. When evaluating the impact of UAT on DME in [11], the interference is considered tolerable, when TTA measured with interference present at a reference signal level, $\nu_{TTAwithI}$, fulfills

$$\nu_{TTAwithI} \leq max(\nu_{TTA} + 0.5s, \tilde{\nu}_{TTA})$$

$$\text{with } \tilde{\nu}_{TTA} = \nu_{TTA} + 2 \cdot \sigma_{TTA},$$

where ν_{TTA} is the mean value of TTA in seconds without interference at the reference signal level, and σ_{TTA} is the standard deviation of TTA in seconds without interference at the reference signal level.
Controversially, this value can be less than 2 s which is the upper limit for DME acquisition performance defined by [12]. After obtaining the TTA values for different desired (D) and undesired (U) signal levels, ASOP is defined as the D/U point at which the DME interrogator is able to acquire a stable track (stable within 2 minutes observation time with a tolerable TTA value). BSOP is defined as the D/U point at which the DME interrogator loses the track.

Two different Devices Under Test (DUTs) have been used in the measurements: Rockwell Collins DME900 is an interrogator typically used in air transportation and Bendix/King KDM706A is a device used in general aviation. Whereas the signal processing in DME900 is partly digitalized, KDM706A is an entirely analog device. Table 1 shows the relevant parameters of these two devices.

Table 1: Used airborne DME equipment

DUT	**DME900**	**KDM706A**
search mode [ppps]	**18**	**100**
track mode [ppps]	**2**	**25**
min. D **[dBm]** ([9])	**-82**	**-82**
measured $\tilde{\nu}_{TTA}$ **[s]**	**1.6**	**2.5**

4 LDACS1

In the air-ground operation mode, LDACS1 employs OFDM in the forward link (FL) transmissions from GS to AS and a combined Orthogonal Frequency- / Time- Division Multiple Access (OFDMA/TDMA) in the reverse link (RL) from AS to GS. FL and RL are on separated frequency channels with an envisaged frequency spacing of 63 MHz which should enable the frequency planning to be coupled with the DME frequencies. In the inlay deployment of LDACS1, a 500 kHz LDACS1 channel is located in the middle between two DME channels.

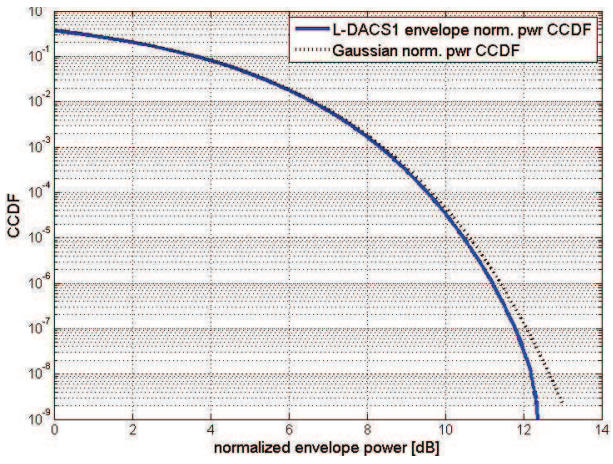


Figure 3: LDACS1 TX power per sample without PA

Due to its broadcast nature, the FL employs a time continuous transmission received by all

AS. As the maximum transmit power of a GS, currently 41 dBm are considered. This power refers to the average OFDM signal power. The peaks of the signal may be theoretically up to 17 dB higher due to the OFDM's Peak to Average Power Ratio (PAPR). In a practical implementation, however, the PAPR is most likely to be limited to a lower value. The hardware realization of the LDACS1 demonstrator used in the measurements limits the PAPR to 12.4 dB. Fig. 3 presents the Complementary Cumulative Density Function (CCDF) of the signal power for each sample, normalized by the average transmission power, thus indicating a low frequency of the occurrence of high PAPR values in the transmitted signal.

RL dynamically allocates a half of or the whole effective bandwidth of 498.05 kHz for a certain time to an AS. Hence, the RL transmission is not continuous and the percentage of time per second in which an aircraft is transmitting is given by duty cycle (DC). The duration of the continuous transmission of an aircraft can vary between $T_{min}^{\mathrm{RL}-active} = 0.6$ μs and $T_{max}^{\mathrm{RL}-active} = 56.88$ ms. In general, the maximum DC can be limited by the resource allocation algorithm. Actually, the highest acceptable DC resulting from the final measurements should be considered by the resource allocation provided by the LDACS1 Media Access Control (MAC). For the average transmit power currently 42 dBm are planned.

The DLR LDACS1 prototype consists of the LDACS1 TX/RX prototype implemented on a Parsec FPGA system and Bögl & Partners Systemtechnik 10.7 MHz to L-Band RF frontend. The baseband demonstrator comprises the TX and the RX physical layer. The TX physical layer including adaptive coding and modulation as well as the complete framing structure for FL and RL is entirely realized in FPGA. The receiver for LDACS1 is not defined in the specification. For that reason and for being able to rapidly implement improved receiver algorithms, the LDACS1 receiver is implemented mainly in software. Only sampling and digital down-conversion from an intermediate frequency to baseband followed by fast data storage of the received baseband samples are implemented in FPGA. The stored signal is processed offline using a software receiver realizing all necessary receiving functions. In order to consider

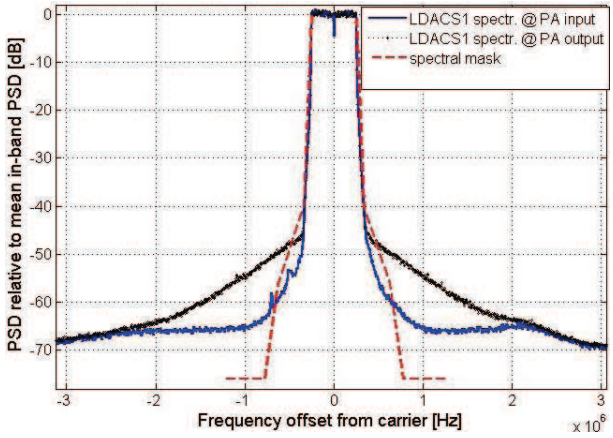


Figure 4: LDACS1 spectrum

the effects of signal amplification, the wideband Power Amplifier (PA) BLMA 0525-35 was used in the measurements to amplify the LDACS1 Tx signal to 35 dBm. The effects of the PA onto the LDACS1 spectrum are presented in Fig. 4.

5 Measurements and Results

To measure the interference caused by an LDACS1 transmitter on the interrogation process of a DME, the interferer is inserted between ground and airborne station as shown in Fig. 5. Two circulators are used to decouple the two paths between the ground and the airborne DME and a coupler to superimpose the victim system with the interfering signal. The power level of the interferer can be adjusted using a step attenuator, while the power level of the desired signal at the victim receiver is kept constant at $D = -82$ dBm. Thus, the attenuator substitutes the propagation loss or the isolation between LDACS1 TX and DME RX.
The undesired signal level U of the LDACS1 interference and desired DME signal level D at DME victim receiver, both measured at the input of the victim DME system, are set into relation in the measurement. Thus, the compatibility criteria are evaluated in terms of desired-to-undesired power ratio D/U.

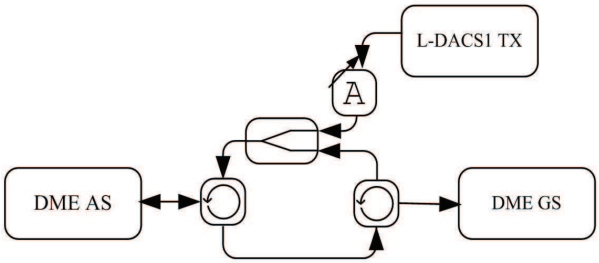


Figure 5: Measurement setup

As a core parameter, TTA has been measured for each undesired signal level U under consideration. Measurements started with a low value of U compared to the desired signal level D. Then U was increased until an D/U point was reached for which no acquisition has been achieved within 20 s. For each of the considered D/U value, mean TTA and standard deviation of TTA have been evaluated from 20 measurements. Whether the interference conditions are acceptable depends on the maximum TTA tolerable. Different tolerance values of TTA are regarded. One possible tolerance margin is the mean value of TTA plus two times the standard deviation of TTA without interference. This rule yields the target TTA values of 1.6 s for DME900 and 2.5 s for KDM706A, according to Table 1. Additionally, maximum TTA of 2 and 5 s are also chosen as possible target values.

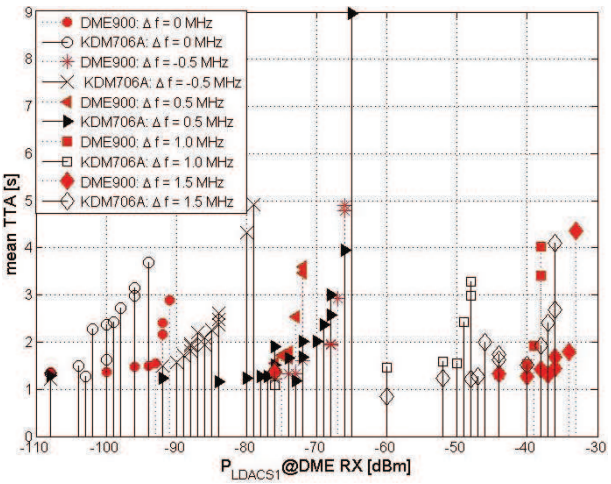


Figure 6: TTA versus LDACS1 power U for $D = -82$ dBm and different frequency offsets.

5.1 G2A Interference Scenario

For the interference scenario a) LDACS GS TX - DME AS RX, the DME frequency was set to 1004 MHz and the LDACS1 frequency was varied from 1002.5 to 1004.5 MHz in steps of 500 kHz. Fig. 6 shows the measured mean TTA values for different U, $D = -82$ dBm and different frequency spacings between the two systems. Noticeable is the different performance of DME900 and KDM706A. The high class Rockwell Collins device is less susceptible to LDACS1 interference. Moreover, due to a non-symmetric spectrum of the devices used, the impact of the undesired signal in not necessary the same on the adjacent channels on both side of the victim DUT. Furthermore, Fig. 7 compares the results of the measurements with DME900 with and without LDACS1 TX power amplifier. Obviously, PA increases the noise floor outside the DME channel and slightly impair the performance at frequency spacings above 1.0 MHz.

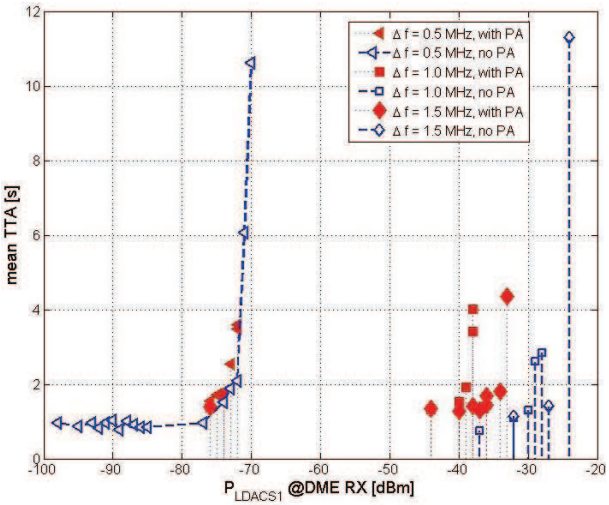


Figure 7: TTA versus LDACS1 power U for $D = -82$ dBm and different frequency offsets measured with DME900 with and without LDACS1 power amplifier.

The lowest D/U for which the mean TTA does not exceed the target value are summarized for different tolerance values for TTA assumed and different frequency spacings Δf in Table 2 for DME900 and in Table 3 for KDM706A.

Table 2: Acceptable D/U [dB] at DME RX for different TTA target values for DME900

TTA [s]	**1.6**	2	2.5	5
$\Delta f = 0$ **MHz**	**11**	11	10	9
$\Delta f = -0.5$ **MHz**	**-9**	-14	-14	-16
$\Delta f = +0.5$ **MHz**	**-6**	-8	-8	-10
$\Delta f = 1.0$ **MHz**	**-42**	-43	-43	-44
$\Delta f = 1.5$ **MHz**	**-45**	-48	-48	-49

Table 3: Acceptable D/U [dB] at DME RX for different TTA target values for KDM706A

TTA [s]	1.6	2	**2.5**	5
$\Delta f = 0$ **MHz**	21	21	**17**	12
$\Delta f = -0.5$ **MHz**	8	6	**3**	-3
$\Delta f = +0.5$ **MHz**	-5	-9	**-13**	-16
$\Delta f = 1.0$ **MHz**	-32	-23	**-33**	-34
$\Delta f = 1.5$ **MHz**	-35	-44	**-45**	-46

5.2 Co-site Interference Scenario

For the interference scenario b) LDACS AS TX - DME AS RX, the DME frequency was set to 1060 MHz. The impact of LDACS1 interference for different duty cycles was measured in co-channel case (LDCAS1 frequency set to 1060 MHz) and on the adjacent channel (LDCAS1 frequency set to 1059.5 MHz) with DME900. Generally, co-site interference is the worst case scenario as the undesired LDACS1 emission from the airborne TX at the airborne DME Rx can be as high as 11 dBm, whereas the desired DME signal as low as -81 dBm. The co-channel case and in general a small frequency separation between two systems in this scenario is theoretically possible, but can be avoided by an appropriate frequency planing. Nevertheless, even in case of large frequency separation, the out-of-band noise could already block the reception of the DME interrogator. However, a low duty cycle of LDACS1 transmission is expected to diminish the impact on the DME performance. In the test set-up, there was no possibility to switch off the PA emission simultaneously with LDACS1 TX transmission. Hence, amplifying LDACS1 signal to 11 dBm would result in a very high noise level of the PA which already blocks the DME unit also when there is no LDACS1 transmission. For this reason, the

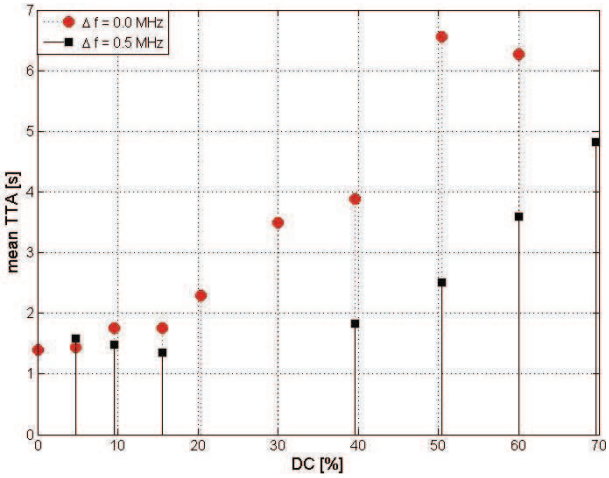


Figure 8: TTA versus DC for different frequency offsets measured with DME900.

Table 4: Acceptable DC[%] of LDACS1 airborne emission for different TTA target values for DME900

TTA [s]	**1.885**	**2.5**	**5**
$\Delta f = 0$ **MHz**	**15.6**	**20.4**	**40**
$\Delta f = 0.5$ **MHz**	**15.6**	**40**	**70**

LDACS1 TX power at PA output is chosen to be $U = -36.5$ dBm. This value is determined in the way to have a mean TTA value if LDACS1 TX is idle similar to the interference free case. The DME power level at the receiver is kept constant at $D = -82$ dBm. With this measurement settings, the no-interference case (DC= 0%) yield a new $\tilde{\nu}_{TTA}$ of 1.8854 s. Fig. 8 shows the measured mean TTA values for different DC settings and Table 4 summarizes the results stating the acceptable DC for different TTA target values.

6 Results Analysis and Conclusions

In general, when interpreting the measurement results, possible deployment scenarios should be taken into account. In ground to air communication, the distance is the major parameter. Considering for example that LDACS1 and DME ground stations are close, i.e. their distances to the airborne DME receiver are approximately the same, the expected D/U is around 22 dB. Besides, the analysis of the results obtained with the two DUTs, should be done regarding their

performance in interference-free case. Results presented in Fig. 6 and Table 2 show that a D/U of -9 dB with LDACS1 interferer in an adjacent channel would not retard the acquisition process for DME900 when considering $\tilde{\nu}_{TTA}$ from Table 1. This TTA target is for DME900 the most stringent TTA requirement. Even 3 dB D/U required for KDM706A in an adjacent channel, regarding a target TTA of 2.5 s, are complied in adequate deployment scenarios. Tolerating TTA of 5 s allows for higher interference than desired power levels. Having an interfering signal with 1.0 MHz separation allows for undesired signal level to be 42 dB for DME900 and of 33 dB for KDM706A above the desired DME power level at the victim receiver.

Regarding the results in the co-site scenario, presented in Fig. 8 and Table 4, the most stringent TTA requirement allows for a DC of at least 15.6%. The next higher DC used in the measurement is 20.4%, hence, it can be expected that the actual DC margin is between these two values.

The results presented show that the LDACS1 co-existence with DME in the most strict inlay deployment option is feasible regarding the presented scenarios. However, the results should be confirmed with additional DUTs and the compatibility should be proved also for other interference scenarios. Furthermore, LDACS1 susceptibility to DME interference should be also tested.

The preliminary work presented in this paper is performed outside the SESAR project. These results will however be available for consideration when performing further tests within SESAR activities.

References

[1] http://www.faa.gov/nextgen/

[2] http://www.sesarju.eu/

[3] Eurocontrol, "Framework for Spectrum Compatibility Analysis in L-Band for FCI technology Candidates, Draft 1.0.

[4] SESAR, P15.2.4 ET, "Updated LDACS1 System Specification", Edition 00.01.01.

[5] Eurocontrol, "L-DACS2 System Definition Proposal", Deliverable D2, Edition 1.0.

[6] SESAR, P15.2.4 ET, "Testing Plan for L-DACS", Deliverable EWA04-2-T2-D1, Edition 00.00.07.

[7] SESAR, P15.2.4 ET, "Updated Compatibility Criteria of existing L band systems", Deliverable EWA04-2-T1-D1.

[8] SESAR, P15.2.4 ET, "Updated Interference Scenarios for Existing L-band Systems", Deliverable EWA04-2-T1-D2.

[9] ICAO, "Annex 10 - Aeronautical Telecommunications, Volume I (Radio Navigation Aids)", 6th edition, Juli 2006.

[10] RTCA, SC-122, "Minimum Performance Standards Airborne Distance Measuring Equipment (DME) Operating Within the Radio-Frequency Range of 960-1215 Megahertz", Document No. DO-151A, November 1978.

[11] ICAO AMCP, Working Group C, UAT Subgroup, "Report on UAT impact on DME - compatibility measurement description and results", June 2003.

[12] EUROCAE, "Minimum Operational Performance Requirements for Distance Measuring Equipment Interrogator (DME/N and DME/P) Operating within the Radio Frequency Range 960 to 1215 MHz (Airborne Equipment)", Document ED-54, January 1987.

CEAS 2011

Composite Materials and Technologies Applications

3rd CEAS Air&Space Conference
21st AIDAA Congress

Size and Shape optimization of Composite Stiffened Panels via MLS free-mesh morphing

Giuseppe Quaranta
Dip. Ingegneria Aerospaziale - Politecnico di Milano, Italy.

Luca Lanzi and Marco Sirna
iChrome Ltd - www.iChrome.eu - United Kingdom

Keywords: *optimization, free-mesh morphing, buckling.*

Abstract

This work combines Moving Least Square mesh morphing methods with classical optimization techniques to optimize Composite Stiffened Structures against buckling constraints. The idea is to use standard optimization methods to handle thicknesses and percentages of the composite laminates and Moving Least Square morphing capabilities to modify the shape of the stringers without generating new meshes from geometry during the optimization process. In this respect, free-mesh morphing methods seem promising methodologies to solve different size-and-shape optimization problems where existing meshes can be modified without specific needs to change neither their connectivity information nor specific model properties such as materials, loading and boundary conditions.

1 Introduction

This paper focuses on so called free-mesh morphing techniques as novel and promising techniques for size and shape optimization of structures.

More and more often, structural optimization problems deal with different geometrical configurations as part of the design process. This is particular true in aerospace and automotive fields: height and flange width of stringer in stiffened structures, beam sections, corrugated elements, reinforcement plates are but few very common examples.

When facing these problems two approaches are commonly used:

- remeshing of the structural assembly so to follow the changes in the geometry;
- mesh-morphing of existing models where positions of nodes are moved so to match the new geometry without changing the connectivity of the mesh.

Both these approaches present strength and weak aspects. The first one is a cad-centric approach where the CAD world, i.e. geometry information, drives the FE one. The main advantage of this approach is that the CAD model is always up-to-date and that FE meshes can be easily generated to respect well defined quality criteria such as maximum size of elements, aspect ratio, angles, etc.

However remeshing techniques often require some supervision from final users who may need to write specific algorithms and macros to propagate changes in geometry from CAD to FE meshes. As a matter of fact and despite the impressive evolution of meshing tools and their increased capabilities to handle parametric

models, re-generating meshes every time a new solution has to be analyzed within an optimization process may turn out to be a complex and expensive process prone to errors when assigning properties to new meshes (boundary and loading conditions, material properties, element definitions, etc).

Accordingly, remeshing techniques may not be always easy to implement or to set-up without efforts. This is particularly true in many industrial applications where stress engineers/analysts are in charge of performing FEM based optimization without specific and continuous support from CAD designers, this leading to a not always easy integration of CAD and meshing tools within structural optimization loops.

In this scenario, mesh morphing techniques seem to be a very attractive alternative to remeshing. In fact mesh morphing techniques:

- do not require close interactions between stress-engineers and designers as stress-engineers will only work on FE models disregarding the CAD world;
- do not require to attribute material properties, loading and boundary conditions to new models;
- are faster as do not require to re-generate and control mesh connectivity.

However, mesh morphing cannot preserve mesh quality above required standards when large changes in geometry are required, often leading to broken or distorted meshes.

The work herein presented moves from these initial considerations and present a novel mesh-morphing technique based on so called Moving Least Square approximation. The technique is first reviewed and then applied to the optimization of composite stiffened panels against buckling loads.

2 Mesh Morphing:

Strictly speaking Free Mesh Morphing is a general purpose technique used to deform an existing mesh - or better the nodes upon which mesh connectivity is defined - without requiring explicit information on the geometry initially used to create the mesh.

More in particular Mesh Morphing allows to modify positions of sets of target nodes in the space by defining a continuous and smooth deformation field to be added to the original position of the nodes. Morphing algorithms can thus be distinguished based on the way they define the deformation field to be applied.

In the structural field two main schemes have been used:

- linear scheme: in this scheme the deformation field is defined by a linear combination of linear deformations spanned wise a set of controlling handles. This technique is known to be used by MSC and Altair in their size and shape tools. Generally speaking this technique is intuitive and easy to use as it produces very predictable deformation fields. It is also computationally effective but as main disadvantage it often leads to broken meshes;
- non-linear scheme: these have been more recently introduced in many research activities [1-4]: Radial Basis Function (RBF) and MLS approaches seem to be the more commons one. Despite being more computational expensive and sometimes less intuitive to use, they outperform the linear scheme with respect of smoothness and final quality of the deformed mesh. Both RBF and MLS can also be adapted to perform linear interpolation by properly down-selecting their inner interpolation functions.

This paper focuses on the use of MLS as available in Shaper [6] – commercial software for general purpose size and shape optimization.

From a general standpoint, MLS techniques belong to the field of data approximation form a scattered set of points. Applications of MLS can be found in many different fields: computer graphics and visualization, image processing, regression models, supervised learning, mesh-free finite element methods are but few examples [1-5].

The following section briefly outlines the fundamentals of MLS approximation methods. More detailed information can be found in [4].

3 Moving Least Square:

The Moving Least Square technique can be used to solve any function approximation problem. The problem can be formulated as follows: given N points located at positions $\mathbf{x}_i \in \Re^3$ where $i \in [1,\ldots,N]$, find a globally defined function $p(\mathbf{x})$ that approximates a set of given values $f(\mathbf{x}_i)$.

In order to have the maximum versatility, no specific structure is supposed in the data point distribution $\mathbf{x}_i$, which can be considered scattered in the space $\Re^3$.

This very abstract formulation can be easily applied to mesh morphing. In fact, by solving function approximation it is possible to find the function that represent the mesh nodes displacement field given the displacement of few points, the controlling handles.

A classical approach to solve this problem is represented by the Least Square that looks for the solution of the following minimization:

$$\min_{p\in\Pi_m} \sum_i \left(p(\mathbf{x}_i) - f(\mathbf{x}_i)\right)^2 \qquad (1)$$

The real problem here is to find the appropriate functional space for the function $p(\mathbf{x})$ able to be a good approximation of the real function $f(\mathbf{x})$ globally. Usually the space of polynomial of maximum degree m is taken.

To avoid the necessity to find a functional approximation space that is valid globally, Lancaster and Salkauskas proposed the MLS technique [5]. The idea is to perform individually a Weighted Least Square for each point $\mathbf{x}$ in the domain. So, the global approximation $p(\mathbf{x})$ is obtained from the union of a series of local functions i.e. $p(\mathbf{x}) = p_{\mathbf{x}}(\mathbf{x})$, where the functions $p_{\mathbf{x}}(\mathbf{x})$ are obtained as solution of the following minimization problem:

$$\min_{p_{\mathbf{x}}\in\Pi_m} \sum_i \left(p_{\mathbf{x}}(\mathbf{x}_i) - f(\mathbf{x}_i)\right)^2 \theta\left(\|\mathbf{x} - \mathbf{x}_i\|\right) \qquad (2)$$

The function θ is a non-negative weight function that depends only from the Euclidean distance between points in the solution space. The approximation is localized as much as possible if $\theta(r)$ is rapidly decreasing as $r \to \infty$. Extremely useful as weight functions are the compact support RBF functions suggested by Wendland [7].

Then, the approximation problem with MLS is localized in two ways:

- first, a local polynomial fit is evaluated continuously over the entire domain; each point $\mathbf{x}$ has its own local approximation.
- second, the number of points $\mathbf{x}_i$ that effectively have an influence on the local polynomial representation at $\mathbf{x}$ is kept limited by using RBF weight functions (better if with compact support).

By looking at Eq. (2) it is easy to see that if $\theta(0) \to \infty$ the MLS fit will be forced to interpolate the prescribed values at known points $f(\mathbf{x}_i)$.

Of course, the MLS approach is more expensive than other approach because for every point $\mathbf{x}$ where the solution must be computed it is required to solve a Linear System of Equations to find the minimum of Eq. (2). However, the size of these LSEs an be kept small by using small support dimension for the weight functions.

In order to solve the MLS problem the user should consider the following aspects:

- the functional space for the local approximation functions $p_{\mathbf{x}}(\mathbf{x})$. Usually the functions are polynomial, so the user must decide the maximum polynomial order.
- the type of weight functions. This choice may have an influence on the smoothness of the obtained approximation (see [4]).
- the size of the local support. Usually a good strategy is to fix the number of points that must belong to each support, and then adapt the local size accordingly. In this way a natural adaptation of the

algorithm is realized, since where there are few spread points the support is large, while for more dense areas the support becomes smaller.

4 Morphing Examples:

This section presents a first example on how MLS morphing techniques can be used to modify the shape of a very simple FE mesh: a cube made of bricks (8-nodes solid elements).

The deformation field is defined by using 8 control points (so called free handles in Shaper) located on the 8 vertices of the cube.

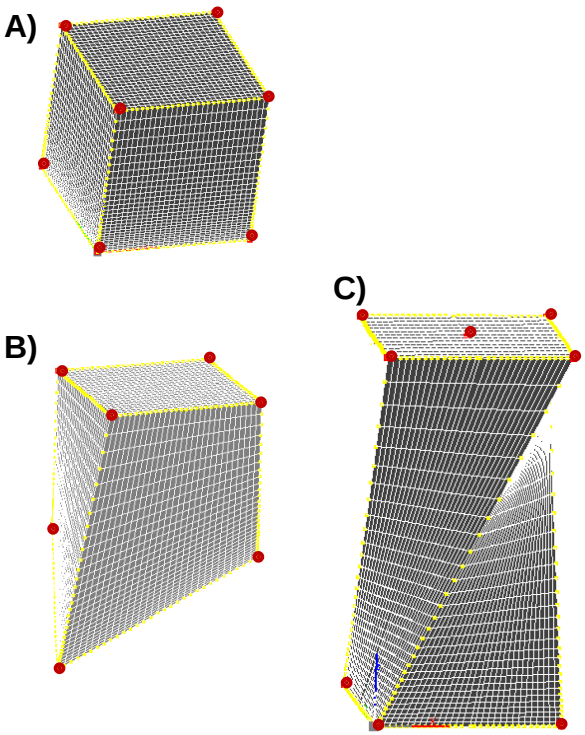


Fig. 1: a first example of MLS morphing: a) original mesh; b) single handle movement; c) combined handle movements.

Figure 1.A) presents the original FE mesh of the cube together with 8 controlling handles located at the vertices. In Fig.1.B) a handle on the bottom-face is translated and the cube deformed accordingly; finally in Fig. 1.C) a new handle is added on the upper face of the cube and it is linked (by means of a rigid constraint) to the other 4 top-face handles. The handle is then translated vertically and rotated of 90 degrees, the mesh of cube following the deformation imposed by the handles. The same configuration of handles would as an example allow to inflate and to contract the cube as well as to deform it to become a more general parallelogram.

It is worth noticing that:

- a relative small number of handles has been required to apply a deformation field overall the model;
- the deformation field works by deforming elements the more they are closer to a moving handles. This is exactly what we expect from a morphing algorithm assuring the mesh to remain almost untouched in those regions where no handles are moved.

4.1 Dealing With Constraints:

A challenging aspect when performing mesh-morphing on structures is related to the need to preserve particular features of the model such as holes, corners, fixed positions for connections, etc.

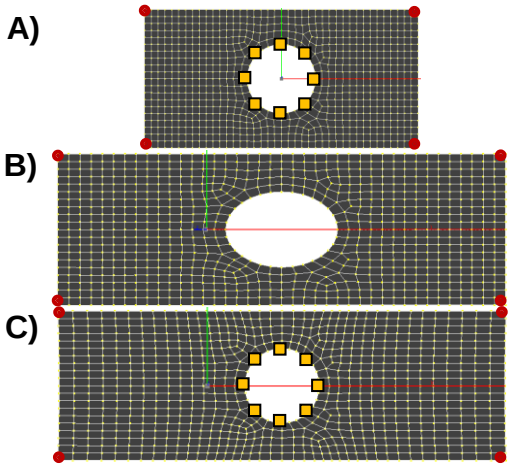


Fig. 2: MLS morphing with additional constraints. Controlling handles in red circles, constraint handles in orange squares. a) undeformed model, b) deformed without constraint handles, c) deformed with constraint handles active

When using MLS interpolation, this can be achieved by hierarchically combining handles to assure rigid transformations. As an example, a simple square panel with a hole in the middle is considered in Fig. 2.

4.2 Linear and Non-Linear Morphing:

An advantage of MLS against classical linear morphing techniques is the wide choice of interpolations and number of support points than can be actually used to define the mesh deformation field. Clearly this freedom turns out to be extremely powerful but at the same time may be counterintuitive requiring more expert users than a pure linear morphing approach would probably do.

This section attempts to prove the flexibility of the MLS by using two different interpolation schemes to morph the cross-section of typical stringers. Shape optimizations on stringer sections are very common problems in the structural engineering field, particularly when dealing with lightweight structures. In the following, a typical stringer is considered and different morphed solutions are presented.

Figure 3 reports a first simple example on how MLS with linear interpolation scheme can be used to control the cross section of general structural components – in this case back-to-back C-shaped stringers. Results of Fig. 3 have been achieved with a limited number of 7 controlling handles at each end of the stringer.

Similarly, Fig. 4 illustrates possible shapes to be achieved with non-linear MLS approximations. In the first case an overall number of 74 handles has been used to morph the foot flange of the stringer; in the second case an overall number of 125 handles has been used. It is worth noticing that despite the relatively high number of controlling handles, parent handles can be defined and chained to reduce the final number of parameters to be finally controlled toward an optimization process.

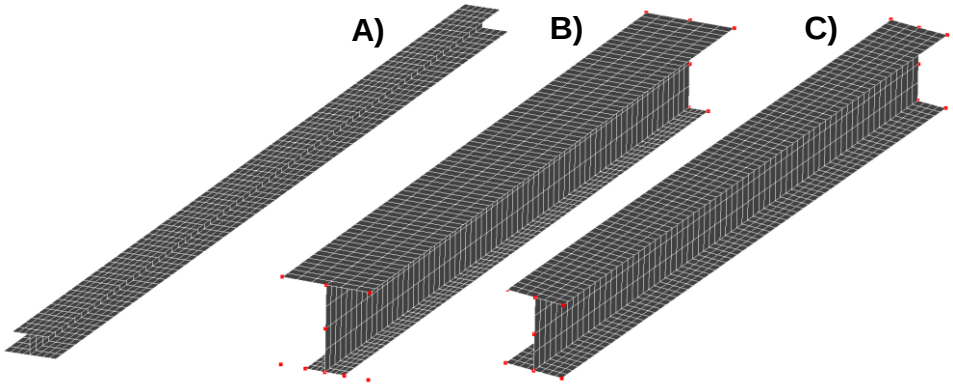


Fig. 3: MLS linear morphing: a) undeformed stringer geometry, b) and c) morphed stringer geometries

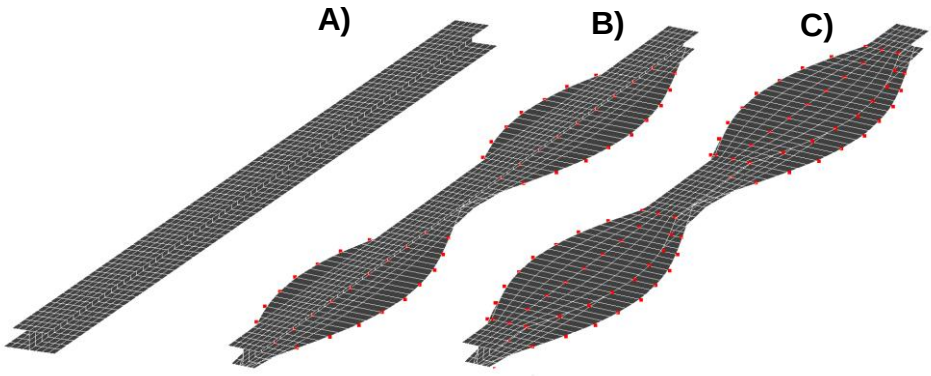


Fig. 4: MLS non-linear morphing: a) undeformed stringer, d) morphed foot stringer flange, c) morphed foot and top stringer flanges

5 Real World Application: Stiffened Panel Optimization.

The proposed morphing technique is thus applied to minimize the weight of a flat composite stiffened panel against bucking loads.

Composite stiffened panels are widely used in primary lightweight structures as they represent weight-effective structural configurations: airplane wing covers, fuselage panels, helicopter tails and tailplanes are but few common applications.

Furthermore, the recent trend of using composite laminates in aerospace and automotive fields strongly increases the number of design variables on which engineers can leverage to achieve required performances.

In this respect, the optimization of stiffened composite panels subjected to buckling and strength constraints has been dealt with several studies in the last decades. Among the first studies on this field a minimum weight design was performed by Butler and Williams [9] using VICONOPT; other minimum weight optimizations including buckling load constraints were proposed by Wiggenraad et al. [10]. Often, the presence of local optima and integer variables, like the number and the orientation of the layers, make the use of the genetic algorithms (GA) appealing in optimization involving composite structures [10-13]. As an example, Kaletta and Wolf [14] applied a parallel computing GA, considering buckling and maximum strength constraints, to stiffened composite plate panels. The fitness evaluation was performed using directly eigenvalue finite element analyses.

This paper moves from the aforementioned works and applied MLS morphing to consider changes in the cross section of the stringers of a composite stiffened panel assembly.

It is highlighted that this work in focused mainly on the application of mesh morphing techniques more than on the engineering design of composite stiffened panels. Mainly for this reason, the choice of the author was to keep the optimization process as simple as possible from a computational standpoint and to underline the main advantages of using MLS. In this respect, only buckling constraints have been considered and the optimization problem has been formulated as a multi-objective search where panels with minimum weight and maximum buckling load are searched for.

5.1 Stiffened Panel configuration:

The considered structural assembly is reported in Fig. 4. More in particular, a flat panel with five stringers and 2 rib bays has been considered as an exemplificative aeronautical structure. The stiffened panel is 1750 mm wide and 2300 mm long, with equally spaced stringer with a pitch of approximately 250 mm.

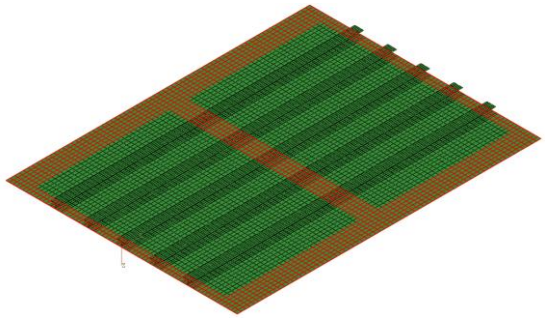


Fig. 4: considered structure assembly: a three-rib composite stiffened flat panel.

The panel is made of a unidirectional CRFP material whose nominal mechanical properties are reported in Tab. 1.

Description	Value
Elastic modulus E11	13500[N/mm2]
Elastic modulus E22	8500 [N/mm2]
Poisson coefficient ν12	0.35
Shear modulus G12=G13	4200 [N/mm2]
Shear modulus G23	3150[N/mm2]
Compression strength X	1280[N/mm2]
Compression strength Y	250[N/mm2]
Tension strength X	2210[N/mm2]
Tension strength Y	75[N/mm2]
In plane shear strength	95[N/mm2]
Nominal ply thickness	0.30[mm]
Density	1600[Kg/m3]

Tab. 1: mechanical properties of the stiffened panel's material.

The rationale is to optimize not only the thicknesses of panels and stringers but also the cross section of the latter via MLS.

5.2 Finite Element Model:

Finite Element Analyses have been performed to predict the behavior of the structure assembly using Abaqus [15].

All parts of the structure have been modeled using four nodes laminated shell elements with reduced integration (S4R) with six degrees of freedom at each node and a single integration points throughout the thickness for each generalized composite ply. The elements used to model the skin have dimension of about 25mm and aspect ratio very close to 1.0, resulting in a total number of 4780 elements. Each stinger is modeled by using 1260 shell elements: 4 elements on each side of the foot flanges 3 elements on each side of the upper ones and 4 elements along the web. Linear material models are used for all parts.

The connection between the skin and the stringer feet has been modeled using tie contacts so to allow the nodes of the foot flange to change their position, as a result of the morphing, without needs to modify the mesh of the skin.

Boundary conditions have been defined to account for spars and ribs: vertical translations of nodes corresponding to spars and ribs landings have been constrained not to translate vertically.

One end in the span-wise direction has been constrained against longitudinal translation; the other is loaded by applying a uniform displacement.

Eigenvalue FE analyses are performed to obtain the first bucking load of the structure, requiring an average CPU time of about 3 minutes. Figure 5 a), b) and c) shows local and global buckling patterns for different analyzed configurations.

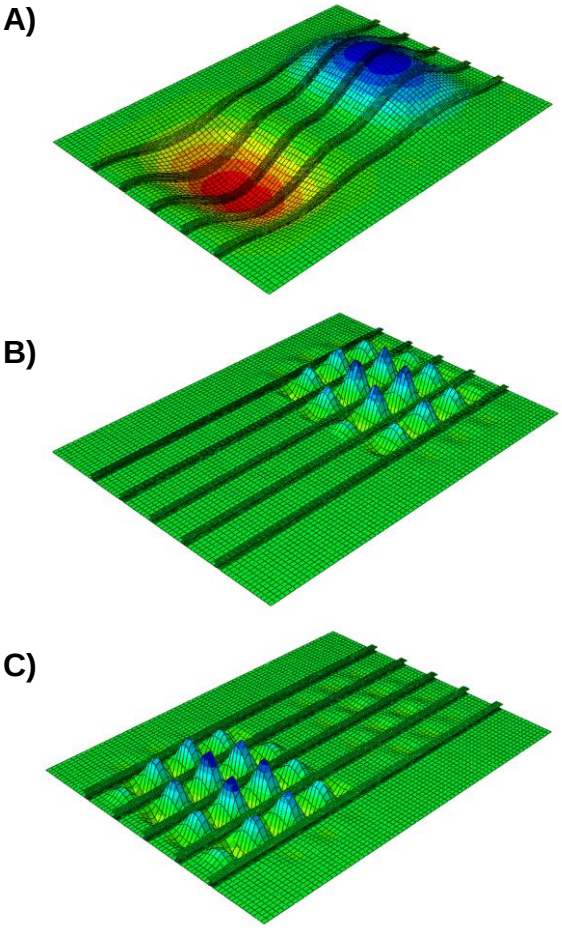


Fig. 5: buckling pattern of different analysed configurations: a) global instability; b) and c) local instability.

5.3 Optimization Problem:

The goal of the proposed optimization is to reduce the overall panel weight while minimizing the buckling load in the attempt to find the best layup (thicknesses and percentages) for skin and stringers, as well as the optimal shape for the stringers via MLS morphing as shown in Eq. (3).

$$\begin{cases} \min\limits_{\vec{x}\in D}\left(W(\vec{x})\right) \\ \max\limits_{\vec{x}\in D}\left(F_{CR}(\vec{x})\right) \\ F_{CR}(\vec{x}) \geq F_{Limit} \end{cases} \qquad (3)$$

A constrained multi-objective genetic algorithm (MOGA) has been applied to identify the solutions representing the best trade-off

between weight, by changing laminate thicknesses and percentages as well as geometry of the stringers, and first buckling load under pure assail compression.

Even if it is recognized that the proposed approach is a simplified one, it is also believed to represent typical aerospace design conditions.

5.3.1 Design variables:

Design variables can be grouped in two main sets: variables controlling the shape of the stringers, i.e. the position of the handle in the morphing process, and variables controlling the thicknesses and the lay-ups of panel skin and stringers.

As far as the stringer shape is concerned, Tab. 2 reports the range of variation of the stringer flanges and webs being the shape controlled via 3 displacements of the controlling handles: one to stretch/shrink the upper flanges, one to stretch/shrink the foot flanges and the last one to stretch/shrink the height of the web.

Description	Min	Max
Stringer height *[mm]*	23.5	82.5
Stringer half-foot width *[mm]*	12.5	62.5
Stringer half-head width	12.5	62.5

Tab. 2: optimisation variables: stringer cross-section domain of interest.

For the panel skin the following pseudo-homogenized symmetric stacking sequence has been assumed: [+α/-α/0/90/90/0/-α/+α], where α represents a general ply orientation. A value of α=45 will lead to +/- 45 degree layers so commonly used in real world applications, herein equally adopted in the stringers.

Description	Min	Max
Stringer 0 ply thickness *[mm]*	0.75	5.00
Stringer 90 ply thickness *[mm]*	0.75	5.00
Stringer +/-45 ply thickness*[mm]*	0.75	5.00
Skin 0 ply thickness *[mm]*	0.50	5.00
Skin 90 ply thickness *[mm]*	0.50	5.00
Skin +/-α ply thickness*[mm]*	0.50	5.00
Skin α ply orientation *[degrees]*	20.0	80.0

Tab. 3: optimisation variables: composite laminates.

In both cases the optimization has been performed considering overall equivalent thicknesses orientation-by-orientation anyway, assuming the limits reported in Tab. 3.

5.3.2 Constraints:

Six constraints have been defined to upper- and lower-limit the total thicknesses of panel and stringers, as reported in Tab. 4. These also guarantee to obtain a feasible shape for the stringers, assuring the width of the foot flange to be greater than the top flange width.

Dimension	Min	Max
Panel thickness *[mm]*	3.0	7.0
Stringer Flanges thickness *[mm]*	2.5	6.5
Stringer Web thickness *[mm]*	5.0	13.0

Tab. 4: limits of the imposed geometrical constraints

5.3.3 Optimisation Algorithm:

The proposed multi-objective Genetic Algorithm is a modified version of the NSGA-II algorithm [16] with a double-swap crossover operator and a non-uniform mutation operator (namely the magnitude of mutation decreases with the number of iterations progressively confining the search in the most promising area of the domain). The algorithm available in Nexus [17] has an enhanced rank-based turnover strategy for the searching of the Pareto's set on the base of which feasible solutions are preferred to the unfeasible ones and, in the subset of feasible solutions, the ones with better objective functions are selected.

Limiting to the present work, the following settings have been used for the Genetic Algorithm:

- maximum number of iterations = 75
- population size = 300
- kind of crossover = full weight
- probability of crossover = 0.6
- kind if mutation = uniform
- probability of mutation = 0.3

5.3.4 Optimization results:

The optimization has been started with an initial population of 300 members randomly selected in the design space.

However, it is worth noting that the total skin and stringer thicknesses as well as the width of the flanges can be calculated directly once the values of the design variables have been assigned without need of any FEM analysis.

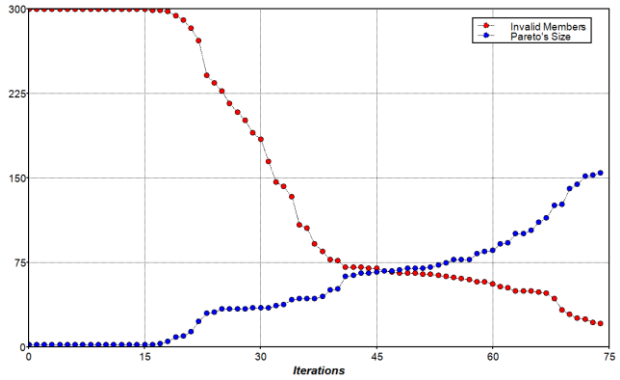


Fig. 6: convergence history: number of invalid members (red) and Pareto's Size (blue) per iterations.

An effective optimization strategy would thus be to evaluate FE analyses (required to estimate the buckling load) only on those configurations that satisfy the imposed geometrical constraints.

This constraint ranking permits to reduce the number of FE evaluations in the first phase of the optimization procedure, when most of the population members may not satisfy the imposed constraints (invalid members), as showed in the Fig. 6.

The procedure has been stopped when the maximum number of iterations (75) has been reached instead. The time required to complete the optimization procedure on a common Pentium Intel Core i7-940 Processor is about 24 hours, performing eight analyses at a time.

The main results of the procedure are summarized in Fig. 7, showing the total evaluations performed and the Pareto's set of the problem.

The Pareto's front, which represents the minimum-weight panel configuration at a given buckling load, envelopes the performed FE analyses, identifying solutions which range from 30.25 Kg up to 67.37 Kg.

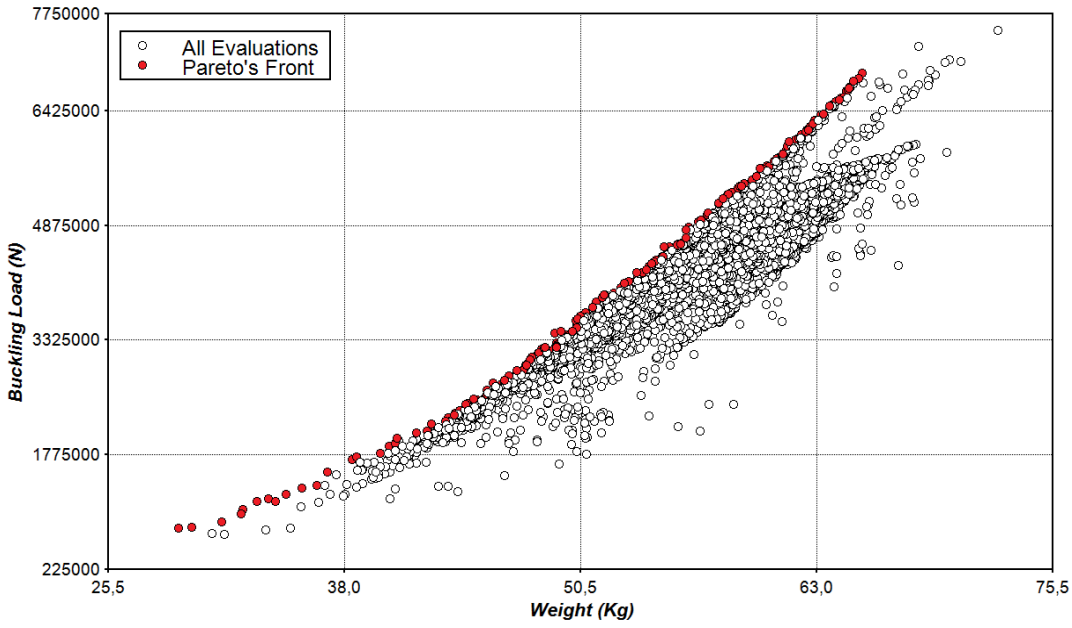


Fig. 7: identified Pareto's Front and all performed evaluations after 75 iterations

6 Conclusion

The paper focuses on the use of mesh morphing applied to structural optimization problems. A practical example considering a composite stiffened panel against buckling constraints has been presented.

The proposed morphing procedure seems an appealing alternative to more classical (linear) approaches and allows- in this specific application - to modify size and shape of the stringers.

The resulting optimization process uses 10 design variables to control both shape variables (that define the geometry of the stringers) and lay-up variables controlling the thicknesses of each orientation of the composite laminates.

A modified version of the NSGA-II multi-objective Genetic Algorithm has been used to outline the trade-off between weight and first buckling load.

References

[1] Alexa, M., Behr, J., Cohen-Or, D., Fleishman, S., Levin, D., and T. Silva, C. *Computing and rendering point set surfaces*. IEEE Transactions on Visualization and Computer Graphics, Vol 9, 2003, p. 3-15

[2] Belytschko, T., Krongauz, Y., Organ, D., Fleming, M., And Krysl, P. *Meshless methods: An overview and recent developments*. Computer Methods in Applied Mechanics and Engineering, Vol 139, 1996, p. 3–47

[3] G. Quaranta, P. Masarati, P. Mantegazza: "A Conservative Mesh-Free Approach for Fluid-Structure Interface Problems", *International Conference for Coupled Problems in Science and Engineering*, Greece 2005.

[4] Levin, D. *The approximation power of moving least-squares*. Math. Comp. 1998, 67(224): 1517–1531.

[5] Lancaster, P., And Salkauskas, K. *Surfaces generated by moving least squares methods*. Mathematics of Computation, Vol. 87, 1981, p.141-158

[6] iChrome Ltd. *Shaper: User Manuals - Keywords*, Bristol, 2011 - www.ichrome.eu

[7] Wendland, H. *Piecewise polynomial, positive definite and compactly supported radial basis functions of minimal degree*. Advances in Computational Mathematics, 1995,Vol. 4, p. 389-396

[8] Butler R, Williams FW. *Optimum design using VICONOPT, a buckling and strength constraint program for prismatic assemblies of anisotropic plates*. Computers & Structures 1992; 43(4):699-708.

[9] Wiggenraad JFM, Arendsen P, da Silva Pereira JM. *Design optimization of stiffened composite panels with buckling and damage tolerance constraints*. American Institute of Aeronautics and Astronautics, AIAA-98-1750, 1998. p. 420-430.

[10] Muc A, Gurba W. *Genetic algorithms and finite element analysis in optimization of composite structures*. Composite Structures 2001; 54:275-281.

[11] Tabakov PY. *Multi-dimensional design optimisation of laminated structures using an improved genetic algorithm*. Composite Structures 2001; 54:349-354.

[12] Nagendra S, Jestin D, Gürdal Z, Haftka RT, Watson LT. *Improved genetic algorithm for the design of stiffened composite panels*. Computers & Structures Vol. 58, 1996, p. 543-555.

[13] Lillico M., Butler R., Hunt GW, Watson A., Kennedy D, Williams FW. *Optimum design and testing of a post-buckled stiffened panel*. American Institute of Aeronautics and Astronautics, AIAA-2000-1659, 2000. p. 1-10.

[14] Kaletta P, Wolf K. *Optimisation of composite aircraft panels using evolutionary computation methods*. Proceedings of ICAS 2000 Congress, Harrogate, UK, 27 August - 1 September 2000. p. 411.1-411.10.

[15] Abaqus. *FE user's manual*, version 6.7. Habitt, Karlsson& Sorensen Inc., 1080, Main street, Pawtucket, Rhode Island, 02860-4867, USA. 2002.

[16] Deb K., Agrawal S., Pratap A., Meyarivan T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation - TEC, 6(2) 182-197 (2002)

[17] iChrome Ltd. *Nexus: User Manuals - Keywords*, Bristol, 2011 - www.ichrome.eu

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Design optimization of a non-crimp fabric Pi-shaped multi-functional profile for composite aero-structures

Konstantinos I. Tserpes and Spiros G. Pantelakis
Laboratory of Technology & Strength of Materials, Department of Mechanical Engineering & Aeronautics, University of Patras, Patras, 26500, Greece

Keywords: Non-crimp fabric composites, Adhesively bonded joints, Finite element analysis, Optimization

Abstract

A numerical design optimization of a non-crimp fabric Pi (⏄)-shaped multi-functional profile is presented. Initial design has been done such that the profile to replace a riveted junction between composite sub-components in a helicopter structure. The geometry of the profile has been optimized for the load-cases of T-pull and T-bending. Optimization is accommodated using the sub-problem approximation method implemented on the basis of a 3D finite element model. First, the critical stresses that cause damage initiation and dominate damage progression are identified using progressive damage modeling. Minimization of these stresses serve as the objective function for the optimization. Then, a decision on the design variables (joint dimensions) and state variables is taken and the optimization analyses are performed in order to obtain the optimal configuration of the Pi-profile.

1 Introduction

Today, a rapid growth in composite applications for fuselage, wing and other structures (28% in A380 and 50% in Boeing 787) has been realized due to the superior 'specific' mechanical properties of composites allowing weight, associated fuel consumption and pollution reductions. However, the substitution of metals with composites has presented formidable challenges to realize lighter, stiffer, stronger and damage tolerant structures that are cost-efficient, regarding development, manufacturing, inspection, or even operating costs.

A major design problem of composites is the efficient joining systems for high performance aero-structures; in particular, transfer of distributed loading along rivets lines in the assembly of flat or curved thin-walled stiffened panels, and introduction of concentrated loads for connecting lugs with thicker laminates are two indicative problems. Due to the large number of such interfaces in aircrafts, a solution to these problems is highly promising in terms of weight and cost reductions. Besides, such interfaces and joints have not reached their best possible mechanical performance. Indeed, high in-plane stiffness, out-of-plane strength, bearing strength, resistance to delamination/damage tolerance of fastened areas are concurrent mechanical objectives that are difficult to achieve simultaneously within the current design, analysis and manufacturing practice. Moreover, it can be stated that current design practices give non-optimal jointing systems that are poor in terms of weight penalties and cost efficiency. Consequently, composite designs are often rejected in favor of traditional metallic concepts.

During assembly of fuselage shells and barrels in civil aircrafts, splices are generally used (adding significant weight), as well as a large amount of rivets (that increase the costs).

Three alternative assembly options would be [1,2]:

- To lower the number of rivets used in butt joints by introducing the concept of geometrical interlocking and bonding of the panels,
- To lower (or even to eliminate) the number of rivets used to join fuselage panels by using 3D woven and/or non-crimp fabric (NCF) joining 'H profiles', replacing the classical splices,
- To join those large components together though specific multi-functional profiles combining the function of stiffening (frames, stringers) and joining (splices), and displaying sufficient 3rd direction properties (through the optimal combination of 3D weaving and braiding technologies).

The present work aims to contribute in the effort for developing efficient adhesive-based joining technologies between composite parts.

2 Statement of problem

The scope of this work is to perform a numerical design optimization of a NCF Pi(⏚)-shaped multi-functional profile (Fig.1a) destined to replace a riveted junction (Fig.1b) between a frame and a beam in a composite generic helicopter structure (Fig.2).

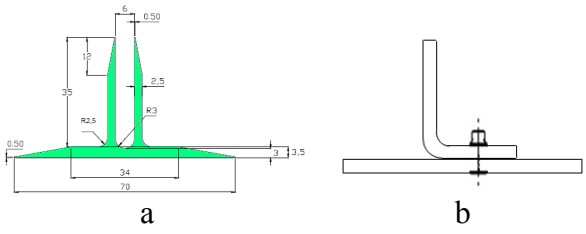


Fig. 1 Schematic of a. the Pi-shaped profile, b. the riveted junction.

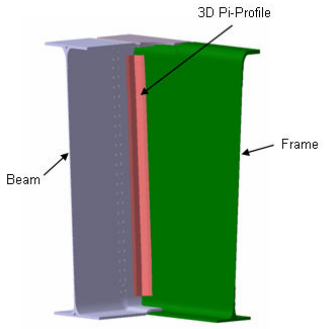


Fig. 2 Schematic of the Pi-shaped based assembly at the helicopter structure.

Optimization will be done on the basis of two loading conditions, which are representative of the service conditions of the composite joint in the helicopter structure, namely T-pull (Fig.3a) and T-bending (Fig.3b). Geometry of the joint will be optimized with regard to its strength under both loading conditions.

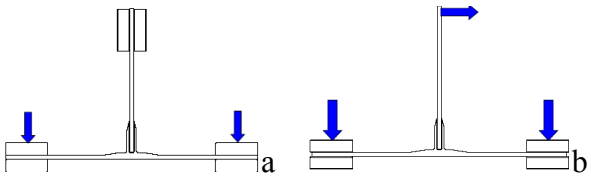


Fig. 3 a. T-pull and b. T-bending loading conditions.

The Pi joining element, the beam and the frame are made from quasi-isotropic HTS/RTM6 NCF quad-layers. The adhesive used is a mix of the paste adhesives EA9395 and EA9396. The mean thickness of the bondline is 0.5 mm.

3 Optimization module

To perform optimization, the optimization module integrated in the ANSYS commercial FE code was employed [3]. The module employs three types of variables to characterize the design process: the *design variables*, the *state variables* and the *objective function*. These variables are represented by scalar parameters in the ANSYS Parametric Design Language.

The independent variables in an optimization analysis are the design variables. The vector of design variables is indicated by:

$$\mathbf{x} = (x_1 x_2 x_3 \ldots x_n) \qquad (1)$$

Design variables are subject to n constrains with upper and lower limits, that is,

$$\underline{x}_i \le x_i \le \overline{x}_i \qquad (i = 1,2,3,...n) \tag{2}$$

where n is the number of design variables. The design variable constraints are often referred to as side constraints and define what is commonly called feasible design space. Now, minimize

$$f = f(x) \tag{3}$$

subject to

$$\begin{aligned} g_i(x) &\le \overline{g}_i \qquad (i = 1,2,3,...m_1) \\ \underline{h}_i &\le h_i(x) \qquad (i = 1,2,3,...m_2) \\ \underline{w}_i \le w_i(x) &\le \overline{w}_i \qquad (i = 1,2,3,...m_3) \end{aligned} \tag{4}$$

where f is the objective function, g_i, h_i, w_i are state variables containing the design with underbars and overbars representing lower and upper bounds, respectively, and the sum $m_1 + m_2 + m_3$ is the number of state variables constraints with various upper and lower limit values.

4 Methodology

The methodology adopted to achieve the goal of the present study is the following:

1. Progressive damage modeling of the joint with the initial geometry under the specific loading conditions aiming to find the critical stresses that cause initiation of main damage mechanisms and govern damage propagation. Minimization of the maximum values of these stresses at the sites of damage initiation present the objective functions.
2. Decision on which dimensions of the joint will be the design variables.
3. Determine the restrictions in which design variables must be subjected based on information from the manufacturing sector and the overall assembly of the helicopter structure. These restrictions will be the state variables of the process.
4. Optimization loops are executed in order to obtain the (optimum) configuration of the Pi profile at minimum is capable to transfer the desired loads (loads efficiently transferred by the initial riveted junction) and at maximum is capable to transfer the desired loads with a maximum safety factor.
5. Validation of optimization by ensuring that the optimum geometry possesses higher strength than the reference one.

4.1 Progressive damage modeling of the joint ─ Objective functions

Strength of the joint was predicted using a mesomechanical model based on progressive damage modeling. The model was developed in [4] and it is able to simulate the mechanical performance and predict strength of textile structural parts. To date, it has been successfully applied to H-shaped adhesively bonded joints [5] and double-lap shear joints [6]. The model implements progressive damage modeling at two different scales: at the micro-scale in order to numerically characterize the homogenized mechanical behavior of the NCF material and at the macro-scale in order to predict strength of the joint. The first model component concerning the NCF HTS/RTM6 material has been described in detail in [4]. FE mesh of the NCF joint is shown in Fig.4. All parts have been represented using the ANSYS SOLID185 element [3].The model is able to simulate the initiation and progression of different failure modes in the NCF material (Pi and skin) as well as debonding due to fracture or shearing of the adhesive.

The model predictions have been verified by mechanical tests in [7].

Analyses revealed that under T-pull loading the joint fails due to interlaminar fracture (delamination) initiated at the Pi-base/skin interface and propagated parallel to skin. This is also verified by the experiments as shown in Fig.4. Interlaminar fracture was due to large normal stresses maximized at the edge of the Pi. Fig.5 shows the distribution of the normal stress computed by the model. Fracture initiated at Pi edge and propagated inwards as revealed by the

transfer of the stress maximization point. Final failure of the specimens was due to either a combination of delamination and bending of the skin or due to a complete separation of the Pi from the skin without any significant damage in the skin.

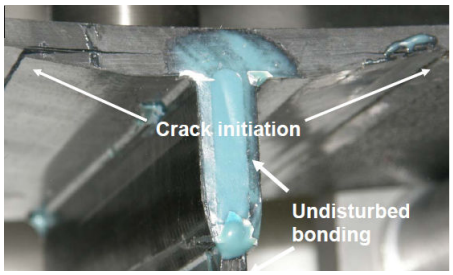


Fig. 4 Initiation of interlaminar fracture in the Pi-shaped joint due to T-pull loading.

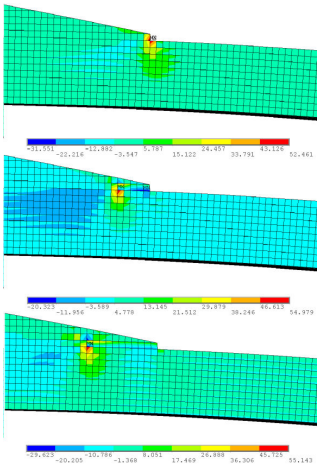


Fig. 5 Computed contour of normal stress σ_y (in MPa) at the joint subjected to T-pull as a function of the applied load: focus on the edge of the Pi.

For T-bending, the failure mechanism was debonding of the insert from the Pi occurred at the stretched part of the joint, as can be seen in Fig.6, due to large tensile normal stresses. The computed distribution of normal stress at the adhesive is illustrated in Fig.8 for different load-steps. The normal stress is maximized at the top of the joint. Debonding initiated at this point and propagated downwards as revealed in Fig.7 by the movement of the point of maximum stress. Actually, the area above the line of stress maximization has been debonded.

Objective functions of the present optimization problem are:

$$\sigma_y = \sigma_y^{\min} \tag{5}$$

for T-pull load-case and

$$\sigma_x = \sigma_x^{\min} \tag{6}$$

for T-bending load-case.

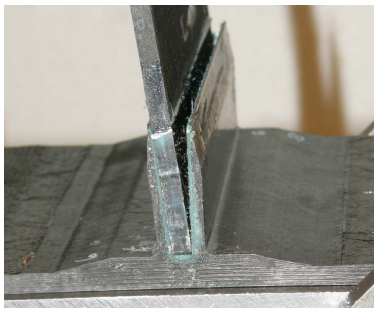


Fig. 6 Failure due to debonding of the specimen loaded in T-bending.

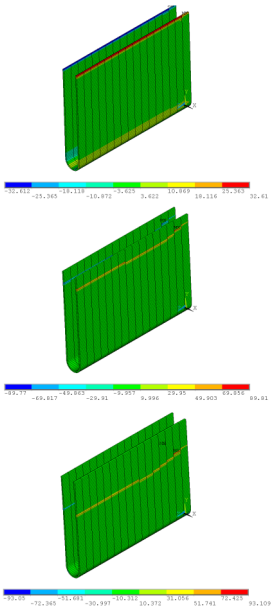


Fig. 7 Computed distribution of normal stress σ_x at the adhesive (in MPa) being responsible for debonding (only the adhesive is displayed).

4.2 Design variables

Design variables of the optimization are the dimensions of the Pi profile and the adhesive that influence the stresses that are to be minimized (Eqs. (5) and (6)). To find these dimensions, a preliminary parametric study was conducted. Assuming that stresses are mainly influenced by the dimensions of the cross-

sectional, the study was done using a 2D FE model of the joint representing only its cross-sectional area. The 2D FE mesh of the joint with the adhesive is shown in Fig.8.

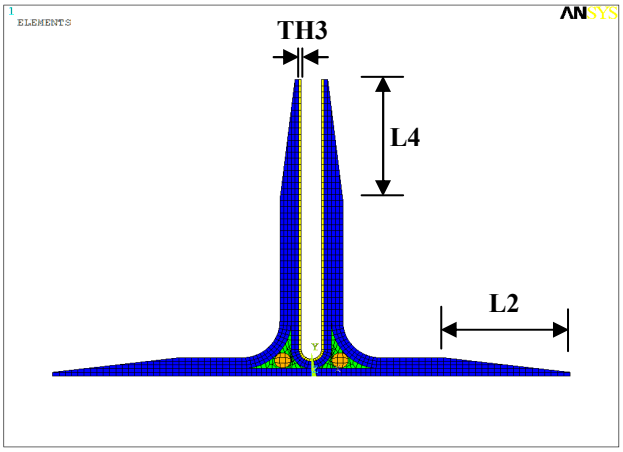


Fig. 8 FE mesh of the Pi-based bonded joint and indication of dimensions served as design variables.

Two separate preliminary studies have been performed; one for T-pull and one for T-bending load-case. In the case of T-pull loading, it was found that σ_y at the Pi base is mainly influenced by tapered length of Pi base (L2), while in the case of T-bending σ_x at the adhesive is mainly influenced by adhesive thickness (TH3) and tapered length of Pi arms (L4). The three dimensions, representing design variables of optimization, are shown in Fig.8. At this point, it is mentioned that stress field at the adhesive and Pi joint, it is also influenced by other dimensions as well as from combinations of dimensions, however, as design variables only dimensions with major effect can be selected.

4.3 State variables

Dimensions of the Pi and the adhesive are subjected to restrictions originated from the manufacturing sector and the assembly requirements. The restrictions, representing state variables in optimization, refer mainly to the shape of the joint and the capability of manufacturing processes. They concern mainly the outer dimensions of the joint as well as the thickness of the arms and base related to the thickness and number of composite layers.

4.4 Optimization analyses

Optimization module is not capable to consider two objective functions simultaneously. A way to overcome this constraint is to perform two sequential optimizations under the condition that design variables between the two optimizations are different and independent. This condition stands for the current application. Therefore, the optimum geometry of the Pi-based joint was first obtained for the T-pull load-case and then, this geometry was further optimized for the case of T-bending load-case.

Figure 9 depicts the log file with the optimization sets for the T-pull load-case. The results are for axial pull-out displacement of 1 mm applied to the insert. L2 varied from 1 to 10 mm. σ_y minimized at maximum value of L2 (10 mm). Optimum set 1 is indicated by stars. Similarly, Fig.10 shows the optimization sets for the T-bending load-case, which have been performed using L2=10 mm. The results are for a transverse displacement of 1 mm applied to the insert.TH3 varied from 0.1 to 0.5 mm and L4 from 1 to 10 mm. σ_x minimized for maximum TH3 (0.5 mm) and maximum L4 (10 mm). The values of design variables from the overall optimization of the Pi joint are:

L2=10 mm, TH3=0.5 mm and L4=10 mm.

The FE mesh of the optimum geometry is shown in Fig.11.

LIST OPTIMIZATION SETS FROM SET 1 TO SET 4 AND SHOW
ONLY OPTIMIZATION PARAMETERS. (A "*" SYMBOL IS USED TO
INDICATE THE BEST LISTED SET)

		SET 1 (FEASIBLE)	SET 2 (FEASIBLE)	SET 3 (FEASIBLE)	SET 4 (FEASIBLE)
L2	(DV)	10.000	9.2359	7.2859	9.9924
SSY1	(OBJ)	33.484	49.687	37.021	33.498

Fig. 9 Optimization output listing for T-pull load-case.

LIST OPTIMIZATION SETS FROM SET 1 TO SET 6 AND SHOW
ONLY OPTIMIZATION PARAMETERS. (A "*" SYMBOL IS USED TO
INDICATE THE BEST LISTED SET)

		SET 1 (FEASIBLE)	SET 2 (FEASIBLE)	SET 3 (FEASIBLE)	SET 4 (FEASIBLE)
TH3	(DV)	0.50000	0.43887	0.25520	0.31502
L4	(DV)	10.000	6.7148	9.6546	8.7240
SSX	(OBJ)	3.9120	3.1208	4.5396	4.1196

		SET 5 (FEASIBLE)	*SET 6* (FEASIBLE)
TH3	(DV)	0.48201	0.49495
L4	(DV)	2.1650	1.2898
SSX	(OBJ)	2.2346	1.7379

Fig. 10 Optimization output listing for T-bending load-case.

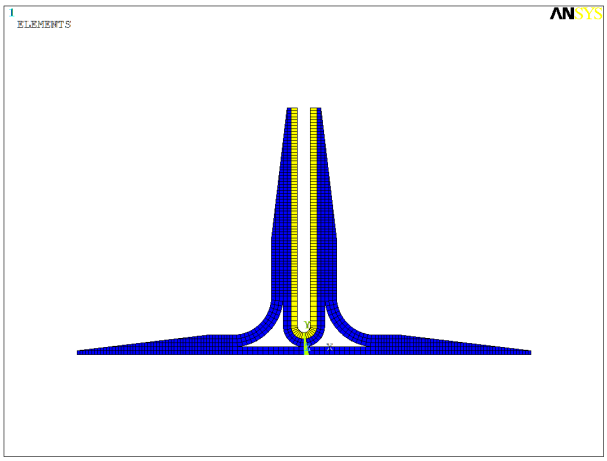


Fig. 11 FE mesh of the optimum geometry of the joint and adhesive.

5 Validation of optimization

For the effected optimization to be effective, the optimum geometry must lead to the maximum feasible strength of the joint. In order to validate optimization, progressive damage analyses have been carried out using the optimum geometry under both load-cases. The predicted failure loads for the optimum geometry are compared with those of the reference geometry in Fig.12. The predicted failure loads of the optimum geometry show an increase of 11% for T-tension and 6.9% for the T-bending load-case, thus validating the conducted optimization.

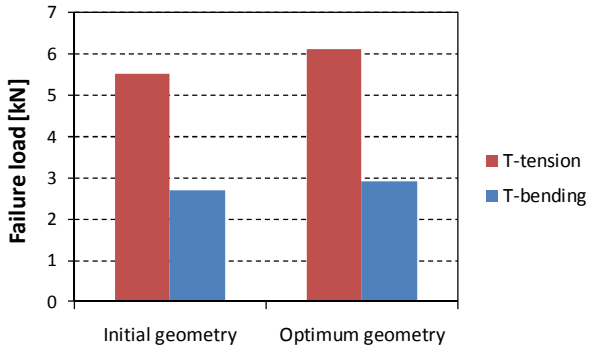


Fig. 12 Comparison between predicted failure loads for the reference and optimum joint geometries.

6 Conclusions

In this paper, a numerical design optimization of an NCF Pi-shaped multi-functional profile was presented. The joint is destined to replace a riveted junction between composite sub-components in a helicopter structure. Optimization of joint geometry was accommodated using the sub-problem approximation method implemented by means a 3D FE model. Optimization was done on the basis of two different load-cases, namely T-pull and T-bending. Optimum geometry possesses higher strength than the initial geometry by 11% for T-pull and 6.9 for T-bending load-case.

The proposed numerical design optimization concept is based on the combination of a optimization algorithm with the progressive damage modeling technique. By performing analyses with this technique, the design variables of the optimization algorithm are derived, while by analyzing the optimum geometry the optimization is validated. The concept can be implemented for the design optimization of any composite aircraft structural part.

References

[1] Modular Joints for Composite Aircraft Components (MOJO), EU FP7 Project, 2006-2009.

[2] Cost Effective Reinforcement of Fastener Areas in Composites (CERFAC) EU FP7 Project, 2010.

[3] ANSYS Version 12.0, *Documentation.*

[4] Tserpes, K.I. and Labeas, G.N., "Mesomechanical analysis of non- crimp fabric composite structural parts" *Composite Structures*, Vol. 87, No. 4, 2009, pp. 358-369.

[5] Tserpes, K.I., Cinquin, J., Pantelakis, Sp. "On the mechanical performance of noncrimp fabric H-shaped adhesively bonded joints", *Journal of Composite Materials*, Vol. 45, No. 15, 2011, 1607-1618.

[6] Llopart P, Ll., Tserpes K.I. and Labeas, G.N. "Experimental and numerical investigation of the influence of imperfect bonding on the strength of NCF double-lap shear joints", *Composite Structures*, Vol. 92, No. 7, 2010, pp. 1673-1682.

[7] Tserpes, K.I. Ruzek, R. and Pantelakis, Sp., "Strength of Pi shaped non-crimp fabric adhesively bonded joints", *Plastics, Rubber and Composites*, in print.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Performance evaluation of a full composite Ultra-Light Aircraft and guidelines for their improvement

G. Scarselli, A. Maci
University of Salento, Department of Engineering for Innovation, Lecce, Italy

R. Zecca
Pro. Mecc. srl, Corigliano d'Otranto (Lecce), Italy

Keywords: *Ultra-Light Aircraft, aircraft performance, composite*

Abstract

The present job is focused on a full composite ultra-light aircraft, and has the main goal of evaluating the existing margins for the improvement of the vehicle performance through a dedicated optimization of the aerodynamic and structural properties. For what concerns the aerodynamic properties of the entire aircraft, at first the attention has been focused on the airfoil chosen for the wings. Then the analysis of the complete wing and aircraft has been carried out. The preliminary investigation on the structural properties of the aircraft has been concerning an accurate evaluation of the operational external loads in order to obtain a confident estimation of the structural solicitation parameters and the identification of the most stressed aircraft structural parts.

1 Introduction

In the present job the tools for the performance evaluations of an Ultra-Light Aircraft have been focused and the key parameters for the subsequent optimization phase have been selected. For what concerns the aerodynamic properties of the entire aircraft, at first the attention has been focused on the airfoil chosen for the wings: numerical evaluations have been performed using the commercial code Fluent. Then the analysis of the complete wing and aircraft has been carried out through simple codes in order to obtain a rough preliminary estimation of the aerodynamic properties. The preliminary investigation on the structural properties of the aircraft has been concerning an accurate evaluation of the operational external loads in order to obtain a confident estimation of the structural solicitation parameters and the identification of the most stressed aircraft structural parts. The International Standards concerning the engine powered ultra-light aircrafts have been taken in account and the National and International Market perspectives for such category have been depicted starting from the vehicle performances. The position in this scenario of the Sparviero has been carefully evaluated in order to quantify possible improvements through the comparison with similar aircrafts. The general description of the Sparviero is reported and a brief explanation of its single parts production process is outlined: the hand lay-up technology is analyzed in detail for what concerns its effect on the structural properties of the aircraft components. The technological progress and the recent advances in the design tools has allowed the building of ultra-light aircrafts able to satisfy the different market requirements. Together with the traditional high-wing configuration, the current

tendencies to increase the performance also in this aeronautical market segment, have encouraged the development of vehicles able to perform medium range missions keeping on average speeds around 180 km/h or able to execute moderate acrobatic maneuvers up to 6 g load factors. This is possible only if the aerodynamic and structural design of the vehicle is carefully carried out keeping in mind the technical restrictions of this aircraft category.

2 The ULM regulations

The technological progress in the fields of the materials adopted for the aeronautical applications, of the engines and of the advanced CAE tools allowing an improved structural and aerodynamic design, led to the continuous development of Ultra-Light Aircrafts able to meet a wide range of different market needs. The architectural configurations can vary from the conventional high-wing to the advanced low-wing with outstanding performance able to cover medium range missions with speeds of about 180 km/h and able to perform acrobatic maneuvers. In Italy the Ultra-Light flight, like all the aeronautical products, is ruled by a specific regulation[1,2]: the law n. 106 issued on 25th March 1985 (after the investigations presented in this paper a new regulation has been issued, DPR n. 133 of 9th July 2010[3]).

ULM (law n. 106/1985)	• Double-seat with $W_{to} \leq 450$ kg • $V_{s0} \leq 65$ km/h
RAI V.EL	• $W_{to} \leq 600$ kg • Installed power ≤ 90 kW
CS-VLA (EASA)	• $W_{to} \leq 750$ kg • $V_{s0} \leq 83$ km/h • Take-off distance $\leq$ 500 m • RoC > 2 m/s
CS/FAR 23	• $W_{to} \leq 5670$ kg

Table 1: light aircraft regulations

In the Table 1 a synthetic sketch of the regulations, concerning the Ultra-Light and Light Aircraft, with the related fields of application is reported:

Without going in the technical details concerning the design specifications fixed by each of these regulations, in the following there is a list of some rules related to the Ultra-Light flight:

1. Aircraft specifications

The aircraft can be:

- single-seat, engine powered with a maximum take-off weight (W_{to}) of 300 kg and stall speed not exceeding 65 km/h;
- double-seat, engine powered with a maximum take-off weight (W_{to}) of 450 kg and stall speed not exceeding 65 km/h.

2. Helmet

The employment of a proper rigid helmet is mandatory during the flight on any aircraft.

3. Double-seat engine powered

The employment of this category of aircraft is allowed for: flight trainers; pilots with private license in validity course; pilots with at least 30 flight hours as commands responsible and achievement of an examination.

4. Employment of areas for take-off and landing

It is allowed the employment of proper areas for the take-off and landing operations with the license of the area owners. For operations in proximity or on the civil airports it is required a specific license.

5. Operational restrictions

It is allowed the flight on the entire national territory but not on the towns.

6. Aircraft identification

The aircrafts can fly only if they are equipped with a metallic plate with the identification number released by the AeCI (Italian aeroclub). The identification number must be impressed on the wing bottom surface with letters of minimum dimension of 30x15 cm in contrasting

color. All the documents concerning the aircraft identification must be kept on board. Modifications at the structure or at the colors, and eventual property transfers must be communicated to AeCI.

7. Traffic restriction

The aircraft can fly only from the sunrise to the sunset. The maximum flight altitude is 500 feet (about 150 m); on Saturday and Sunday, is 1000 feet (about 300 m).

3 Sparviero: description and operational information

Sparviero, produced by the Pro.Mecc. located in Corigliano d'Otranto (Lecce), is an Ultra-Light Aircraft fully realized in composite material (epoxy resins reinforced by glass and carbon fibers). This specific segment of the aeronautical market has seen in the last ten years a global increase with special attention focused on those vehicles characterized by advanced technical performance. In the last twenty years the aeronautical industries operating in this sector have spent many efforts to find improved solutions in the aerodynamic and structural fields. The employment of CAE tools and the validation of the numerical predictions in the wind-tunnels have allowed the design of high-performance profiles for the lifting surfaces exhibiting high levels of efficiency; in the same way the structural aspects have been improved through a massive employment of composite materials or solutions aimed to the weight saving.



Fig. 1: Sparviero

Sparviero is a low-wing monoplane aircraft, double seated, with pulling propeller, exhibits traditional aerodynamic solutions and conventional horizontal and vertical tails. The wing is composed by two different rectangular shapes parts with rounded tip, connected to the fuselage through two assembling nodes. The wing is equipped with plain flaps, rotating on a simple hinge. The landing gear is fixed, tricycle with fore wheel able to steer and its aerodynamic properties are improved by proper covers on the wheels aimed to minimize the drag. The aircraft fuselage is fabricated by a specific number of layers with the addition of local reinforcements (glass and carbon fibers) wherever structural reasons require it. The cabin is covered by a canopy with vertical opening; the pilot and passenger seats are fabricated in composite materials properly covered. A little vane is present for luggage. The horizontal tail is rectangular fully fabricated in composite material equipped with a trim. The adopted engine is the widely well-known Rotax 912 100 hp with four opposed cylinders, four stroke, with combined cooling system, electronic injection and electrical starter engine. In the following table the main performances of the Sparviero are listed.

Sparviero technical data	
W_{to} [kg]	450
S [m^2]	11.5
b [m]	8.2
AR	5.8
V_{S0} [km/h]	69
V_{SFF} [km/h]	60
V_c [km/h]	200
V_{max} [km/h]	256
TO/L distances [m]	80/90
RC [m/s]	6

In the first part of this investigation the performances of typical Ultra-Light Aircrafts have been analyzed and compared with Sparviero and it has been realized that this vehicle exhibits interesting features essentially due to the use of light materials. In the following Figures the main performance parameters of different Ultra-Light Aircrafts are

reported on the same graph to allow comparisons. From the Figure 2 it is clear that Sparviero is characterized by an interesting value of the maximum speed (256 km/h) that locates it in the high range of the market.

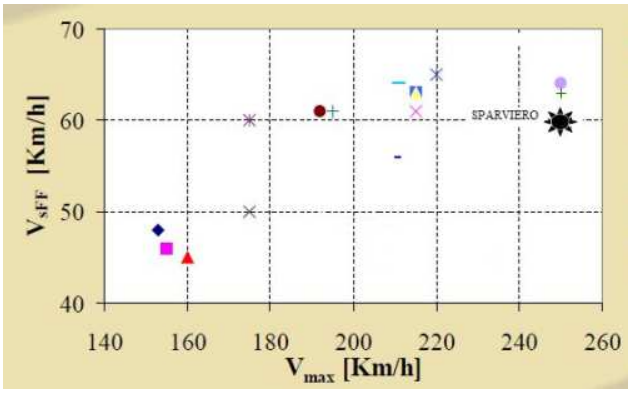


Fig. 2: V_{sFF} versus V_{max} for different ULMs

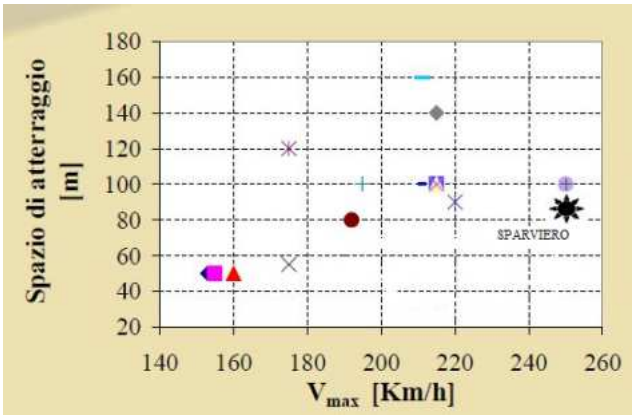


Fig. 3: comparison on the landing distance

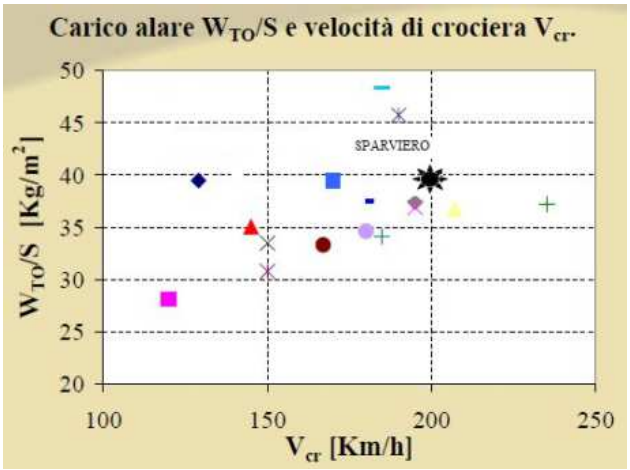


Fig. 4: comparison on the wing load

4 The production process

The different parts of the Sparviero aircraft are fabricated through the hand lay-up technology and subsequently assembled. The procedure adopted consists in the following steps: layers of reinforcing material are disposed inside the mould representing the part; the reinforcement is impregnated with liquid state resin; these operations are repeated until the desired thickness is reached; the system is consolidated through application of the vacuum bag and of a thermal flux for the curing process (it occurs in an oven at about 80°C).

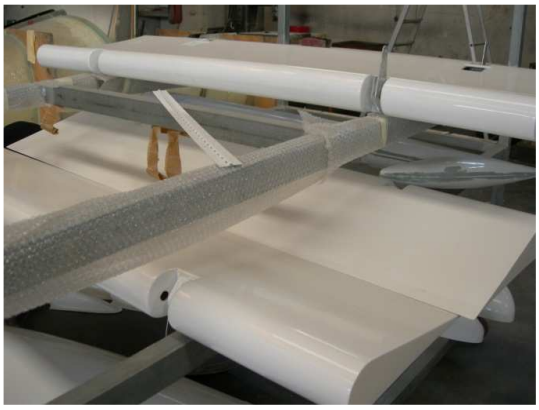


Fig. 5: some Sparviero components made in composite



Fig. 6: carbon and glass fiber rolls

The adopted fibers are provided in big rolls and the operator uses proper tools for cutting the different layers that compose the desired lamination sequence according to the geometry of the part under fabrication.

The adopted resin is epoxy type with addition of two chemical components serving as catalyst and accelerator. These chemical components have a strong influence on the viscosity of the resin and its pot-life, i.e. the overall time that resin, after the addition of the before mentioned elements, spends in the pot before starting to solidify. Finally, specific gel coats are disposed on the mold for improving the surface finish of the single Sparviero components and, at the same time, for ensuring a better environmental resistance. The impregnation of resin and fibers is performed through the employment of special metallic or plastic rullers that are carefully managed by qualified operators: they, before each production phase, control and certify, through a precision balance, the proper weight ratio between resin and fibers in order to obtain optimal structural properties for the component under fabrication. After the polymerization phase, each component receives an accurate and detailed quality control for what concerns the polymerization process and, if this control has had positive result, it is employed in the subsequent aircraft assembly phase.



Fig. 7: Sparviero assembly phase

5 Aerodynamic and structural aspects

At first, the aircraft aerodynamic properties have been evaluated[4]: the codes adopted for the numerical simulations have been Fluent and Xflr5[5] based on Xfoil. Xflr5 is an analysis tool for airfoils, wings and planes operating at low Reynolds Numbers. It includes: XFoil's direct and inverse analysis capabilities; wing design and analysis capabilities based on the lifting line theory, on the vortex lattice method, and on a 3D panel method. The first simulations have been carried out on the bidimensional aerodynamic field around the profile employed for the Sparviero wing and have been aimed to the evaluation of the C_l at the different angles of attack for all the operating conditions characterizing the aircraft flight envelope. In the Figure 8 the typical aerodynamic analyses are reported for the profile in an investigation case.

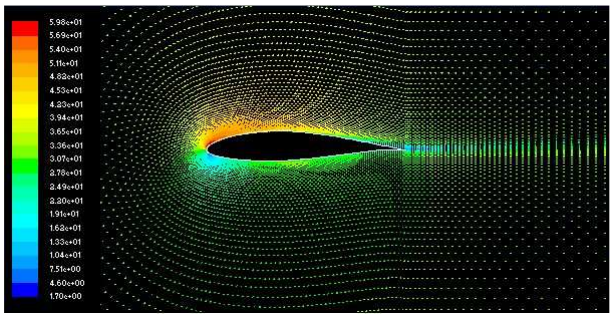


Fig. 8: aerodynamic field for α=8°

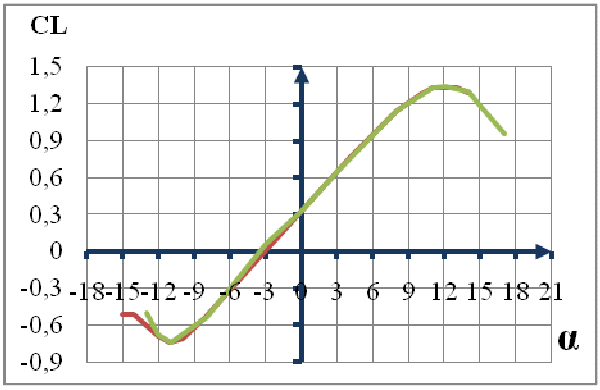


Fig. 9: C_l versus α

In Figure 9 the lifting coefficient at different angles of attack are reported obtained through the Fluent code and through the Xflr5 code for comparison purposes. The analyses have shown that Xflr5 provides reliable numerical results and this code has been used for the following computations keeping well in mind that this activity is just preliminary for a further and deeper investigation. The attention has been then focused on the finite wing. As mentioned before, Xflr5 allows the calculations through three different methods: in this context the lifting line model has been adopted that, in its classical formulation, does not take in account the viscous effects. Just for sake of completeness, it is right to mention that in Xflr5 a formulation taking in account these effects has

been implemented based on the NACA Technical Note 1269. A tridimensional model representing the Sparviero wing has been developed (Figure 10).

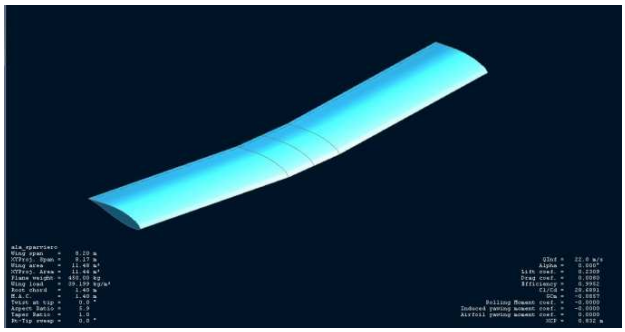


Fig. 10: 3D model of the Sparviero wing

Once developed the geometric model, aerodynamic computations have been carried out in order to obtain: the finite wing polar; the pitching moment coefficient; the wing efficiency at different Reynolds numbers. In the following Figure 11 the complete aircraft polar, result of the summation of the different contributions, is reported.

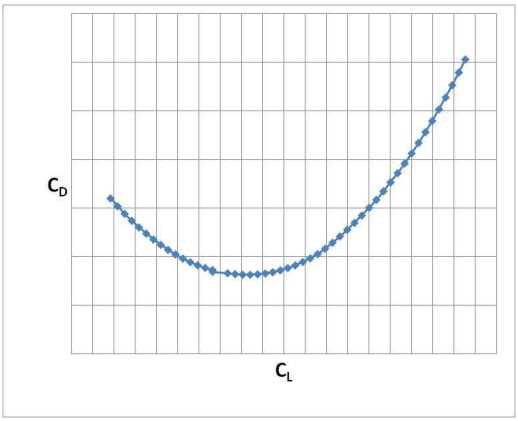


Fig. 11: complete aircraft polar

The simulations have been repeated for all the lifting surfaces of the aircraft. A sensitivity study has been carried out in order to evaluate the influence of the different geometric parameters on the aircraft aerodynamic properties and the effect of these properties on the main aircraft performances that are: take-off and landing distance; range; rate of climbing; the glide performance; the turning performance; the maximum flight altitude[6,7]. The aircraft parameters that play a major role on these performances are: the maximum lift coefficient of the complete aircraft (C_{Lmax}); the wing load, W/S (Nm^{-2}); the ratio between weight W and engine power P (NW^{-1}). In the following part of this paragraph several design equations adopted for the performance evaluation are reported, from which the influence of the before mentioned parameters can be derived:

Take-off distance

$$S_{TO} = K' \frac{1}{\varepsilon C_{Lmax}} \frac{W/S}{75m \frac{\eta \psi_z P}{W V_d} - \mu} \qquad (1)$$

where K', m and η are constants, ε is the relative density, ψ_z is a coefficient depending on the altitude, μ is the friction coefficient and V_d is the lift-off speed.

Landing distance

$$S_L = \frac{V_a^2}{2g.(\frac{C_{Da}}{C_{La}} - \mu)} ln \frac{C_{Da}}{\mu C_{La}} \qquad (2)$$

where g is the gravity, V_a is the landing speed, C_{Da} and C_{La} are the aircraft drag and lift coefficients at landing.

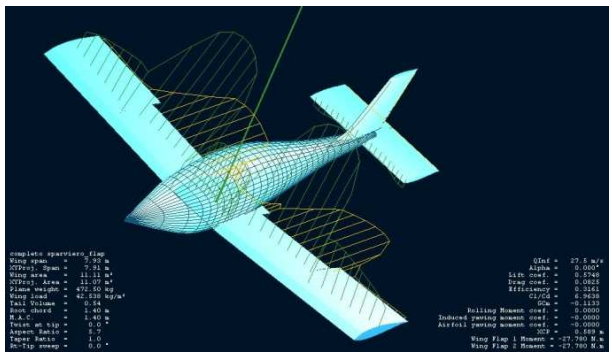


Fig. 12: complete aircraft aerodynamic analysis

A proper optimization of the design parameters allows the improvement of the performance. The following statistical formulations can be furthermore adopted for the evaluation of S_{TO} and S_L providing results very close to analytical formulations and to the actual performance.

$$S_{TO} = 0.0649TOP^2 + 5.0024TOP \ \ [m] \qquad (3)$$

where TOP (Take-Off Parameter) is:

$$TOP = \frac{(W/S)(W/P)}{\varepsilon C_{Lmax}}$$

and:

$$S_L = 0.038V_S^2 - 0.641\ V_S\ [m] \qquad (4)$$

where V_S is the stall speed.
Once defined the aerodynamic properties of the Sparviero, the load calculations in the different representative points of the flight envelope have been carried out.
These computations have been performed according to the RAI V.El. regulation and in the following Figures several results, obtained in one point of the flight envelope, are reported. The wing loads have been evaluated through the Schrenk method and, from the knowledge of these, the solicitations diagram on the lifting surfaces has been drawn. The subsequent step would be an accurate (CAE) evaluation of stress induced in the wing sections by the calculated solicitations[8].

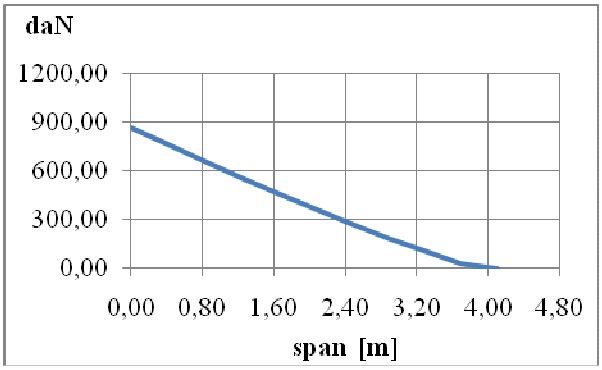


Figure 13: shear at the different sections of the wing

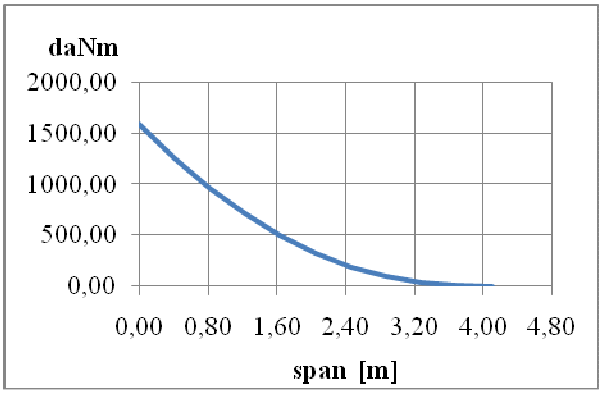


Figure 14: bending moment on the wing

RAI V.El. states several methods for the evaluation of the loads on the rear tails from which the following computations are made.

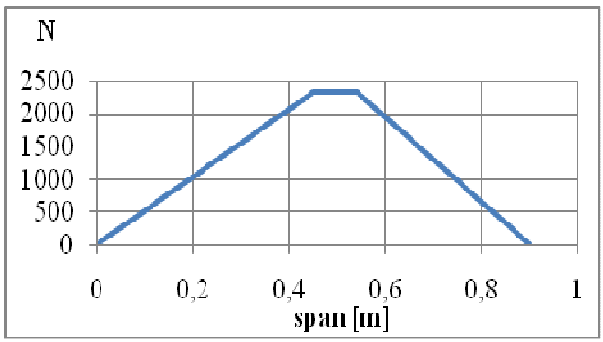


Figure 15: symmetric load on the horizontal tail

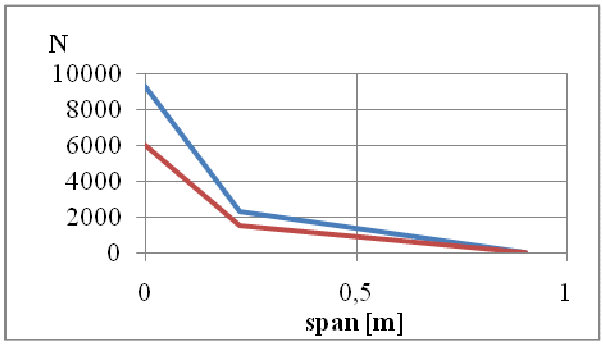


Figure 16: non symmetric load on the horizontal tail

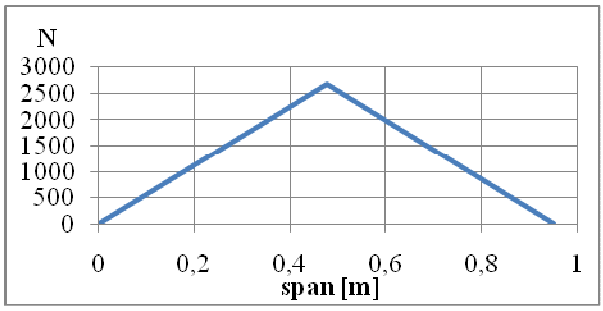


Figure 17: symmetric load on the vertical tail

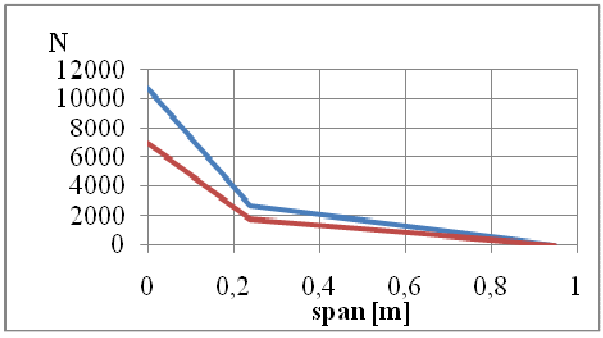


Figure 18: non symmetric load on the vertical tail

6 Guidelines for the performance improvement

The analyses carried out and briefly discussed in this paper allow a knowledge of the aircraft technical properties according to the limitations of the predictions provided by the numerical tools adopted. More refined results could be obtained through more complex and reliable approaches. On the other hand, the activities herein reported have the main goal to fix a procedure for a preliminary evaluation of the aircraft performance and the subsequent optimization of these through the improvement of the key design parameters on which they essentially depend. In the specific case under investigation the attention is focused on the adopted structural solutions and on the production technology.

For what concerns the structural aspects, an improved efficiency can be obtained through the employment of advanced numerical tools aimed to the structural computations: these tools, starting from a deep and complete characterization of the materials coming from a highly reproducible process, allow to establish with a high degree of accuracy the stress and strain distributions in the most stressed parts of the aircraft. In this way a decrease of structural weight of the different parts can be obtained with a strong effect on the aircraft performance as before highlighted.

The adopted production process is hand lay-up type that is widely adopted for the boat applications where large coefficients of safety are used. For such kind of process, a control can be performed on final voids and resin content. Obviously compression, shear and interlaminar shear properties are adversely affected by high void contents. The presence of voids can be responsible of reduced fatigue life and of increased water absorption with related possible reduction of stiffness of the composite. This is the reason for which in the last years aeronautic constructions started to consider processes alternative to autoclave lamination looking at resin film infusion (through thickness flow) and resin transfer molding[9].

It is right to remark that the activities herein presented and discussed refer to a version of the Sparviero aircraft active in 2009. In the meanwhile Pro.Mecc. has continuously improved this vehicle in order to increase the main performance and, at the same time, to guarantee a highest possible level of safety.

7 Conclusions

The present job has been performed inside a cooperation between the aerospace manufacturer Pro.Mecc located in Corigliano d'Otranto (Lecce) and the Department of Engineering for Innovation of University of Salento. Pro.Mecc. entirely builds and sells the Sparviero, a full composite ultra-light aircraft, and has assigned to University the task of evaluating the existing margins for the improvement of the vehicle performance through a dedicated optimization of the aerodynamic and structural properties: the tools for the performance evaluations have been focused and the key aircraft parameters for the subsequent optimization phase have been selected.

References

[1] Italian Law n. 106 issued on 25th March 1985, *Disciplina del volo da diporto e sportivo*, issued on Gazzetta Ufficiale n. 78 of 1st April 1985.

[2] ENAC, *RAI V.El Standard, Standard di aeronavigabilità per velivoli elementari*, Edition 1 – October 1999.

[3] Italian Regulation, DPR n. 133 issued on 9th July 2010, *Nuovo regolamento di attuazione della Legge n. 106 del 25 Marzo 1985*, issued on Gazzetta Ufficiale n. 193 of 19th August 2010.

[4] Ira H. Abbott, Albert E. Von Doenhoff, *Theory of wing sections*, Dover publications, Inc. New York, 1958.

[5] Xflr5 User's manual, *Analysis of foils and wings operating at low Reynolds numbers*, 2008.

[6] V. Losito, *Fondamenti di aeronautica generale*, Edizioni dell'Accademia Aeronautica (Pozzuoli).

[7] J. Roskam, Chuan-Tau Edward Lan, *Airplane Aerodynamics and performance*, published by DAR corporation, Edition 2003.

[8] T.H.G. Megson, *Aircraft structures for engineering students*, Buttherworth & Heinemann, 3rd Edition, 1999.

[9] C. Soutis, *Fibre reinforced composites in aircraft construction*, Progress in Aerospace Sciences 41 (2005) 143–151.

CEAS 2011

Aircraft Preliminary Design

Analysis of Pareto-Frontier for Multidisciplinary Design Optimization of Aircraft

F. Mastroddi, S.Gemma
Dept. of Mechanical and Aerospace Engineering - University of Rome "La Sapienza"
Via Eudossiana 18, 00184 Roma, Italy

Keywords: Multidisciplinary-Design Optimization, Aircraft Preliminary Design, Pareto frontiers

Abstract

This work aims to develop a methodology for an integral analysis and a computational environment in order to perform the preliminary design of a complete aircraft. For this purpose an integrated multidisciplinary design optimization (MDO) has been used to minimize a composite multi-objective function. The use of a MDO techniques can improve the solution obtained compared to a classical sequential design, in that it is able to utilize all the interactions and influences that exist between the disciplines under investigation. The multi-objective problem will be solved by generating a set of Pareto optimal solutions and analyzing the consequent Pareto frontiers. A Weighed Global Criterion (WGC) was used in three different formulations. It has also been analyzed in detail the selected formulations and it has been formulated the concept of shape function. The algorithm was developed in Fortran environment and involves the use of a commercial optimization code and a commercial finite element code, that is able to perform the analysis required by several disciplines.

1 Introduction

The Multidisciplinary Design Optimization (MDO) is a design approach that has been widely developed for the engineering design with a special emphasis in the aerospace engineering field. This methodology is able to reduce design time and costs by integrating all the disciplines involved in the design process which, following different objectives, may typically provide different and contrasting design solutions. Indeed, MDO solutions are so more optimal with respect to standard optimal sequential solutions in that they are able to utilize all the coupling interactions that are present in all the involved disciplines.

An essential issues that characterize the MDO analysis are the natural conflicts that may arise between different objectives within the same design process. Sometimes it is not possible to find a design solution that simultaneously optimize all the objectives but often it is possible to search and identify a set of solutions that perform the best compromise. Originally this concept was introduced by Pareto (Ref. [1]) that defined a solution as Pareto-optimal if it is not possible to find another one that can improve one objective at least without making worse the other ones. The set of Pareto-solutions in the space of the objectives is known as Pareto frontier.

The purpose of the present paper is to determine and study Pareto frontiers for a multi-objective MDO problem of a complete aircraft structure and shape in a preliminary design phase by including as disciplines in the initial design the structural analysis (statics and dynamics), the

steady and unsteady aerodynamics, the flight mechanics and the aircraft performance evaluation.
Typically, the use of Finite Element techniques for structures, CFD, or dynamic analysis involves a high computational burden. Therefore, a compromise between quality and complexity of the studies has necessarily required. Two approaches are possible. First, a preliminary optimization that uses low-fidelity models: this approach is able to study all state space and to overcome the problems of entrapment of local minimum points. Following, second approach consists of optimizing the suboptimal solution and using high fidelity models, with the possibility of employing algorithms based on gradient method. In the present paper a gradient-based approach has been used in order to determine Pareto frontier. Pareto frontier allows to enhance the MDO technique capabilities, as the Pareto-frontier analysis provides a complete problem overview (in terms of coupling degree, influence of design variables and constraints) and it allows a more aware final choice.
As an actual application, a middle range transport aircraft has been considered in this paper. This case study has also proposed as Garteur G70 in the European community research activity (Ref. [5], [6]). The multi-objective MDO problem has involved three different objectives: the aircraft weight, the aerodynamic efficiency and the aircraft range capability. The constraints has been considered by several disciplines (structural stresses due to static and dynamic loads, maximum trim angles for flight mechanics, aeroelastic constraints for stability and response).

2 Study case

Three different objectives have been considered: minimization of structural weight, maximization of aerodynamic efficiency and maximization of range capability; six constraints have also been imposed: stress (traction, compression, Von Mises), trim angles (incidence and balancing), maximum wing displacement. The design variables, 36 overall, are both structural and shape variables, which allowed to explore all possible geometric configurations, Table 1.
The minimization of structural mass W is the first objective considered in the present work. It is estimated during the optimization process by the weight estimator of MD.NASTRAN commercial software (Ref. [8]). The second objective is the aerodynamic efficiency of the wing in flight at constant speed, namely in the cruising condition in which the aircraft operates for most flight time.

$$E = \frac{C_L}{C_D} \tag{1}$$

Range, the third objective, is obtained by Breguet formulas for jet aircraft, and it is expressed as:

$$R = \frac{2}{c_t \cdot C_D}\sqrt{\frac{2 \cdot C_L}{\rho \cdot S}}\left(\sqrt{W + W_F} - \sqrt{W}\right) \tag{2}$$

where c_t is the specific fuel consumption. The weight of fuel W_f on board in the wing box is obtained through knowledge of the fuel density and the evalueted volume of the wing box.

WING and TAIL		
Structural	Skin panels	root thickness thickness ratio
	Spars	root web thickness thickness ratio root cap area cap area ratio
	Stringers	root area area ratio
Shape	Angles	sweep dihedral
	Length	root chord taper ratio span
FUSOLAGE		
Structural	Skin panel	nose thickness wing connection thickness tail thickness
	Stringers	nose area wing connection area tail area
	Floor Skin panel Floor stringers	thickness area
Shape	Length	wing position tail plane height

Table 1: *Design Variables*

3 Multi-objective optimization strategy

The optimization problem aims to minimize the objective function by searching the set of optimal design variables x_n and within the constraints.
Basic knowledges for the problem under review are respectively the state and objectives spaces, where are contained the designs and their images, mapped on the objective function Φ. The constraints define a feasible decision set (X) of design variables vectors, each of which defines a unique project in the objectives space (Y). The Pareto optimal points are on the frontier of the feasible objectives set.
There are several methods to generate Pareto frontier (Ref. [2]). For this paper, based on the context work and depending on the purpose and available resources, the gradient-based Weighted Global Criterion (WGC) has been chosen.
Following this method, objectives are combined into a single functional module (Φ), transforming the multi-objective problem into a single-objective problem.

3.1 Composite objective function

Three main objectives of a aircraft has been considered. Obviously, it seeks to minimize weight and maximize efficiency and range. Moreover, objectives have different orders of magnitude and units so they have to be normalized in order to have a homogeneous combination of objectives of the same order.
Each objective has been normalized using a reference value, which may be its value, for instance, in the initial configuration. In order to build a unique objective function to be minimize, a linear combination of the weight together with the reciprocal of aerodynamic efficiency and the reciprocal of the range is used:

$$\Phi_p = \sum_{i=1}^{3} w_i \, {f_i}^p \qquad (3)$$

with:

$$\begin{aligned} f_1 &= \frac{W}{W_r} \\ f_2 &= \frac{E_r}{E} \\ f_3 &= \frac{R_r}{R} \end{aligned} \qquad (4)$$

where p represents the order of the assumed global objective function, whereas w_i are the weight coefficients and W_r, E_r and R_r are the reference values.

3.2 Shape functions

Analyzing in detail WGC properties, this method allows to compute a Pareto points only if in that points the *shape function* associated to the functional Φ is tangent to the Pareto frontier.
For a N_{obj}-objective problem, the chosen shape functions, in the objective space, once the N_{obj} objective has been esplicited in Eq. 5 and Φ_p value is fixed, one has:

$$f_n = \frac{\left(\Phi_p - \sum_{i=1, i\neq n}^{N_{obj}} w_i f_i\right)^{(1/p)}}{w_n} \qquad (5)$$

Figure 1 shows the shape functions for different value of p, in the simple bi-objective case, with weight coefficients and final value of the functional be respectively $w_1 = w_2 = 0.5$ and $\Phi_p = 0.5$. The higher is the degree of functional the smaller is the radius of curvature. Because of the tangency condition, the ability to capture the curvature is an important property for shape functions: this ensures the tangency with the frontier. It is clear that higher is the value of this parameter, the greater is the shape function capability to capture the points of the Pareto frontier in areas with high non convexity of the surface.
Therefore, the simplest formulation, $p = 1$ (the weighted sum), does not allow to calculate the non-convex areas of the front, because the shape function associated is a hyperplane in N_{obj}-dimensional space of objectives. The set of hyperplanes is the envelope of the front. Furthermore, through the vector of weights we can

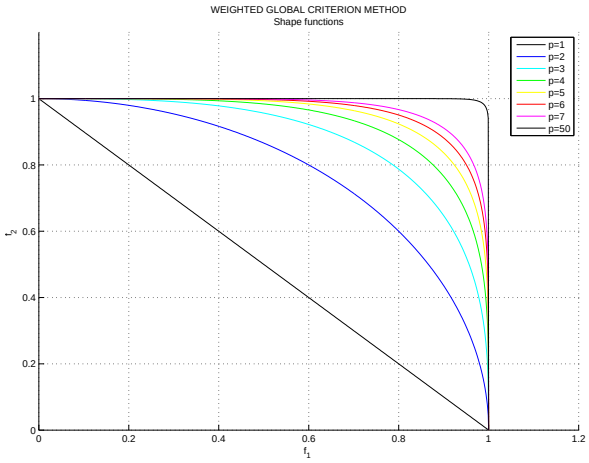


Figure 1: *Shape functions for bi-objective problem with fixed* $w_1 = w_2 = 0.5$, $\Phi_p = 0.5$

calculate the normal vector to the Pareto frontier, which provides interesting information on the geometry of the front, allowing to hypothesize, but not to distinguish, any areas of non-connection of low curvature and non-convexity (Refs. [4], [3]).
Because of the previous comments, the weighted sum can be applied for the first phase of exploration of the Pareto front. In order to overcome the limitations of the first order, higher-order functional were analyzed. By studying the associated shape functions, it was observed that the curvature is proportional to the degree p of the functional. Therefore, their capability to calculate points in non-convex areas increases with the p degree. therefore, 2nd and 5th order formulations will be used for detailed analysis of interest areas.

4 MDO Algorithm

In order to perform the MDO analysis for a complete aircraft, a monolithic gradient-based algorithm with a high-fidelity approach for the disciplines analysis is proposed. Specifically, the MDO procedure is able to take into account the change of aircraft geometry during the optimization process.
This MDO procedure is implemented in FORTRAN 90 language and is based on the use of: i) SNOPT 6.6 code (Ref. [7]) a commercial suite of subroutines for Nonlinear Programming; ii) a home-made geometry and mesh generator, for wing and fuselage structures useful to generate a Finite Element Model (Ref. [9]); iii) MD.Nastran code (Ref. [8]) able to perform the structural and aeroelastic analyses of the designed configuration. The Pareto frontier has been calculated by varying the weight coefficient components vector in the domain [0,1] with a step of 0.05.

$$w_1, w_2, w_3 = (0, 1) \qquad with \quad \Delta w = 0.05 \quad (6)$$

233 designs have been obtained for each method formulation. In the following only projects matching the condition of Pareto optimality are selected through a filter and consequently the frontier has been built.

5 Numerical results

The results concerning Pareto analyses and final design project for a regional-aircraft configuration will be presented in the following subsections.

5.1 Pareto frontier evaluation

To build the Pareto frontier three basic steps are necessary.
Because of gradient-based method, the initial condition choice is critical and requires preliminary analysis. A configuration that already has a acceptable low value of structural weight was chosen.
Secondly, it was necessary to calculate the designs that optimize each objective functions. They are the vertices of the frontier. They also allow to calculate the *utopian point*, which is defined as the vector whose components are the objective function minima. It represents the vector that minimizes all objectives simultaneously, but as unreachable, is utopian, and it really is a reference ideal point in the multi-objective prob-

lem solution.

$$\left.\begin{aligned}
\mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T \xrightarrow{min\,\Phi_1} \mathbf{x}^{*1} \longrightarrow \begin{bmatrix} f_1(\mathbf{x}^{*1}) \\ f_2(\mathbf{x}^{*1}) \\ f_3(\mathbf{x}^{*1}) \end{bmatrix} \\
\mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \xrightarrow{min\,\Phi_1} \mathbf{x}^{*2} \longrightarrow \begin{bmatrix} f_1(\mathbf{x}^{*2}) \\ f_2(\mathbf{x}^{*2}) \\ f_3(\mathbf{x}^{*2}) \end{bmatrix} \\
\mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^T \xrightarrow{min\,\Phi_1} \mathbf{x}^{*3} \longrightarrow \begin{bmatrix} f_1(\mathbf{x}^{*3}) \\ f_2(\mathbf{x}^{*3}) \\ f_3(\mathbf{x}^{*3}) \end{bmatrix}
\end{aligned}\right\}$$

$$\longrightarrow \mathbf{f_{ut}} = \begin{bmatrix} f_1(\mathbf{x}^{*1}) \\ f_2(\mathbf{x}^{*2}) \\ f_3(\mathbf{x}^{*3}) \end{bmatrix} \tag{7}$$

Finally, the Pareto frontier is calculated on the basis of first, second and fifth order formulations of the weights method, or using:

$$\begin{aligned}
\Phi_1 &= w_1 \frac{W}{W_{ut}} + w_2 \frac{E_{ut}}{E} + w_3 \frac{R_{ut}}{R} \\
\Phi_2 &= w_1 \left(\frac{W}{W_{ut}}\right)^2 + w_2 \left(\frac{E_{ut}}{E}\right)^2 + w_3 \left(\frac{R_{ut}}{R}\right)^2 \\
\Phi_5 &= w_1 \left(\frac{W}{W_{ut}}\right)^5 + w_2 \left(\frac{E_{ut}}{E}\right)^5 + w_3 \left(\frac{R_{ut}}{R}\right)^5
\end{aligned} \tag{8}$$

An elementary case of construction of Pareto

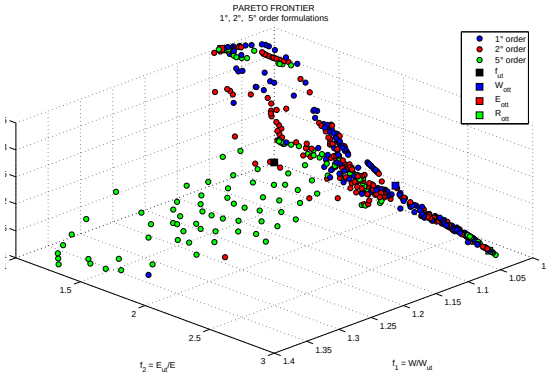
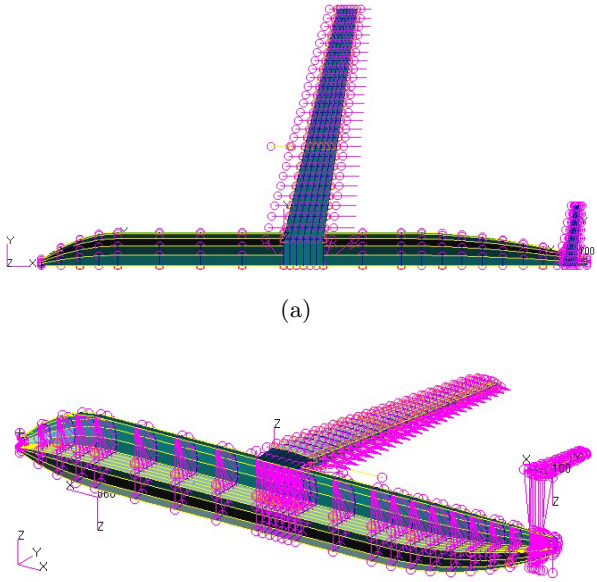


Figure 2: *Pareto frontier calculated on the basis of first, second and fifth order WGC formulations*

frontier in a 2D objective space using the proposed methodology is presented in A.
The Pareto frontier is built in the 3D objectives space, whose axes represent, respectively, weight, efficiency and range (Figure 2). We can distinguish the vertices (squared marks) which are the extreme designs. The minimum weight design W_{opt} is presented in Figure 3(a), 3(b). The maximum range design R_{opt} implies a maximum wing volume for the fuel loading (Figure 5(a), 5(b)). On the other hand, the maximum efficiency design E_{opt} yields a high wing aspect ratio, minimum chord dimension therefore, in order to comply trim conditions, it has been greatly increased the tail size (Figure 4(a), 4(b)). In the present case the utopian point f_{ut}

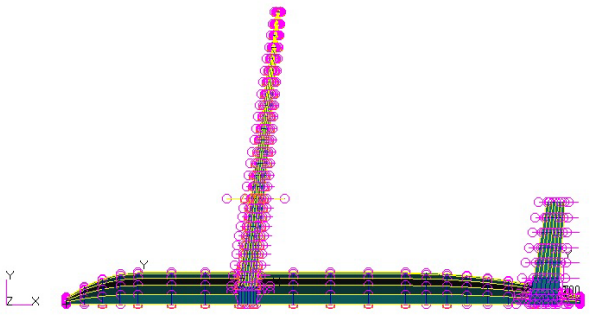


(a)

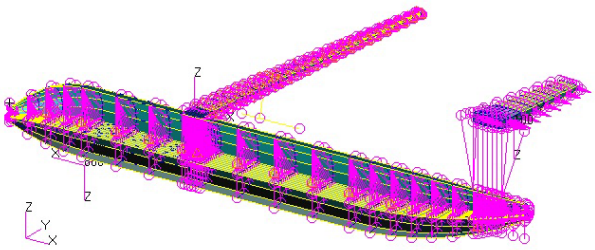
(b)

Figure 3: *Minimum weight design.*

coincides with the axes origin.
Increasing the order of the functional, more central areas are populated by new surface points. The three formulations have provided complementary results, because they have the ability to reconstruct different parts of the Pareto frontier.
Therefore, the strategy to integrate the results has been adopted and the final frontier has been obtained by selecting only the Pareto optimal designs.
In addition, it was noticed that the computation time increases with the order of functional. With the order also increases the probabilities to calculate designs points in areas with strong convexity and more distant from the point utopian (see Sez. 5.2). Therefore, it is necessary to eval-

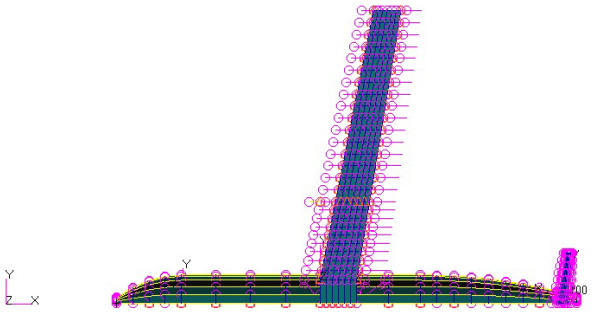


(a)

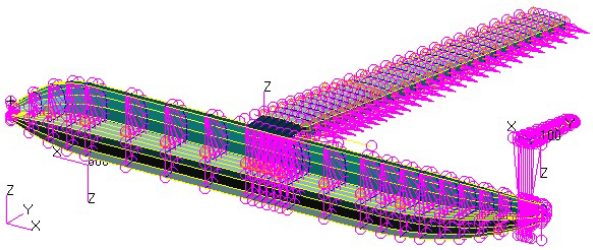


(b)

Figure 4: *Maximum efficiency design.*



(a)



(b)

Figure 5: *Maximum range design.*

uate whether it makes sense to use higher-order functional.

5.2 Choice of final designs on Pareto surface

In order to choose a final optimum design, evaluation criteria have been defined.
The first adopted parameter to do that is typical and is the distance from utopian point:

$$d_{ut}=\sqrt{\left(\frac{W-W_{ut}}{W_{ut}}\right)^2+\left(\frac{E_{ut}-E}{E_{ut}}\right)^2+\left(\frac{R_{ut}-R}{R_{ut}}\right)^2}$$

Mapping the Pareto frontier in a color scale proportional to that parameter, the most satisfactory area has been found in the central part of the frontier (Figure 6). The second parameter d_{Gcr} represents the exploitation of constraints, which indicates how much the constraints value G_i is close to the limit value G_i^{cr}.

$$d_{Gcr}=\sqrt{\left(\sum_i \frac{G_i-G_i^{cr}}{G_i^{cr}}\right)^2} \qquad (9)$$

It also is an indication of the design oversizing. In this case the best area is nearby, but not identical, to that found previously (Figure 7). Figure 8 shows the design obtained minimizing the quantity given by Eq .9: note the this result is very close to that obtained by minimizing the weight, and this is due to the fact that the constraints driving the design are those on stresses in both cases. Finally, preferences has been introduced based on the specifical application on

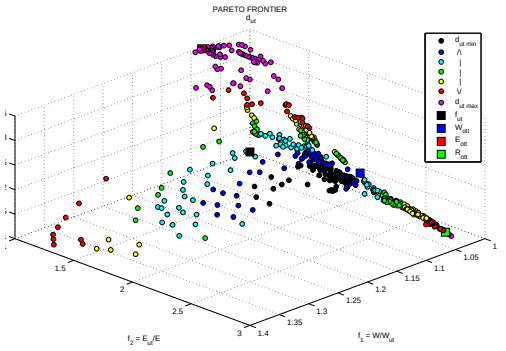


Figure 6: *Pareto frontier mapping on distance from utopian point.*

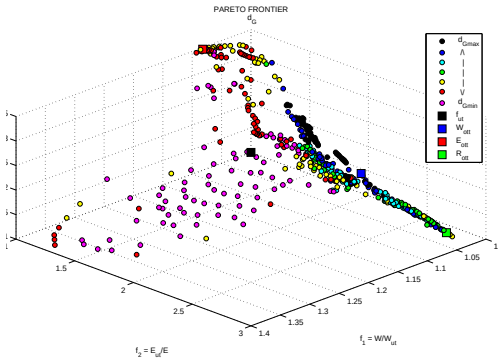


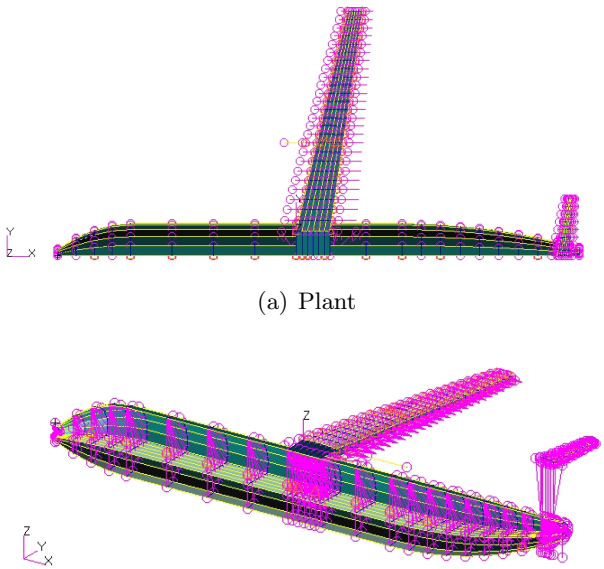
Figure 7: *Pareto frontier mapped by d_{Gcr}.*

regional aircraft. It has been given priority to the structural weight, because if it increase the other goals worse and operating costs increase. Moreover, analyzing the Pareto frontier, it has seen that it is unprofitable to pursue the maximum efficiency and range because they involve unconventional geometries.
A final selected design with structural weight close to the minimum value and constraints close to the limit values has been chosen (see Figure 9 for the related Pareto frontier and Fig. 10 for the final design): note that the final design has been selected taking into account all the criteria discussed above. The configuration has a balanced geometry with high aspect and taper ratio which ensure good efficiency but at the same time sufficient wing volume for a good range. The reached optimal design is a quite conventional geometry as the variables were chosen as to optimize an existing design. However, enlarging the state space, it is possible to obtain non-conventional geometries too.

6 Goal methods application

In order to verify the nature of Pareto optimality of the points numerically obtained in the previous section, a test on some points has been performed as in the following. Indeed, the Pareto problem can be rewritten using the so-called ε-constraints method, or:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_l(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \leq \varepsilon_i \quad i = 1, 2, \ldots, N_{obj} \quad i \neq l \\ & g_j(\mathbf{x}) \leq 0 \quad j = 1, 2, \ldots, M \end{aligned} \tag{10}$$



(a) Plant

(b) Lateral view

Figure 8: *Design closer to constraints violation.*

where $l = 1, 2, \ldots, N_{obj}$. This method applies literally the definition of Pareto optimality: it searches for any objective best value by requiring that the other do not get worse. It then selects one of the objective functions and transformed all other into constraints, by imposing the upper bound. Applying this method randomly for several designs of the frontier, the property of Pareto optimality was confirmed for the multi-objective MDO analysis of an aircraft.
Another application of goal methods has been made. Individually optimizing the distance from the utopian point d_{ut} and the distance constraints d_G from the same initial condition, were obtained designs less satisfactory than the one selected through same criteria from the frontier. However, the application of goals methods does not guarantee that the only result is global optimum, nor that it is an excellent compromise. The goal methods are favored because the cost calculation are apparently minor. But when an application is limited to the determination an initial project (not Pareto optimum) from which to start the process of classical sequential design, design time and costs are therefore higher and the final design is not optimal.

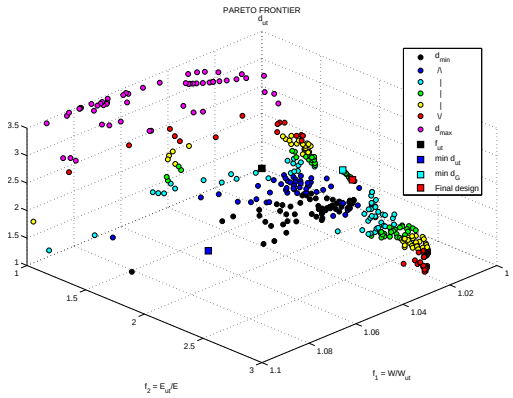
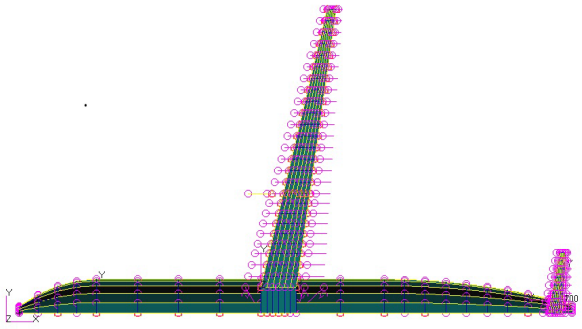
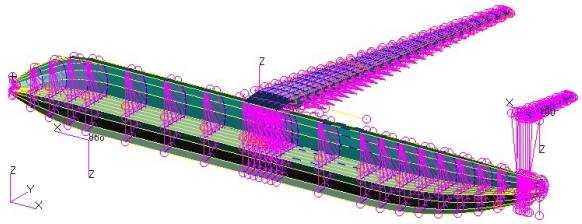


Figure 9: *Pareto frontier mapped by d_{ut} with final chosen design.*



(a) Plant



(b) Lateral view

Figure 10: *Final design.*

7 Concluding remarks

In addition, it was noticed that the computation time increases with the order of functional. With the order also increases the probabilities to calculate designs points in areas with strong convexity and more distant from the point utopian (see Sez. 5.2). Therefore, it is necessary to evaluate whether it makes sense to use higher-order functional. The method adopted, Weighted Global Criterion, has proven to be a good alternative to the global methods for the multi-objective MDO analysis of an aircraft. The formulations of higher order overcomes the limitation of the weighted sum method. Integrating the results for several composite objective function (Eq. 8) we obtain a complete Pareto frontier. Unlike the evolutionary algorithm, it allows the control and analysis by the designer and give the opportunity to focus research in a well-defined area of the frontier through the choice of the weight coefficient w_i. It is also much less expensive because it requires fewer calculations of objectives and constraints.

In addition, it was noticed that the computation time (an Intel Core 2 Duo 6600 processor at 2.40Ghz has been employed for the calculation) increases with the order of functional (see Fig. 11) and with the order the probabilities to calculate designs points in areas with strong convexity also increases: this implies that a compromise between functional accuracy and computational efficiency has to be reached.

The implemented tool for the Pareto frontier construction has provided several design solutions that satisfy all the constraints. This gives the designer a complete vision of multidisciplinary problem because it allows to quantify the coupling between the disciplines, in terms of constraints and design variables influences, values range of the objectives, etc. One can also make the best choice of the design because one knows how much an objective can be improved by the expense of others.
Finally, a better design with respect to the traditional sequential design and goal methods can be achieved. If the design process from preliminary to the advanced stages were based on a multi-objective approach by analyzing the Pareto frontier, the time would be reduced and the design would result in a better performance.

A Test case

The validation of the algorithm and metodology used in the MDO process has been conducted by analytically solving an optimization problem based on Bredt theory of the torsion

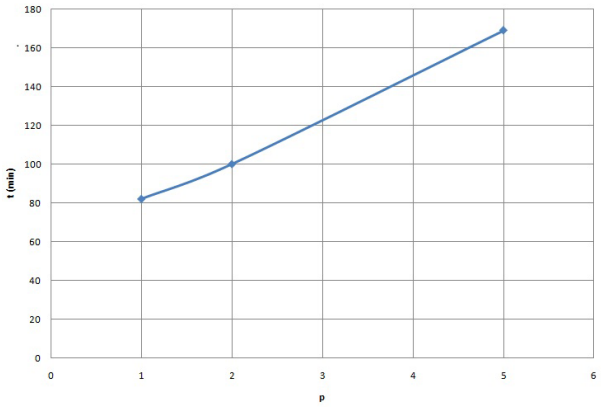


Figure 11: *CPU time for evaluating a design point on the Pareto surface as function of p.*

of beams having simply-connected squared cross section. This problem consists of minimizing the structural mass W and maximizing internal volume V of a hollow thin-walled cantilever beam, with squared section, subjected to a assigned tip torque (Figure 12). The design variables are

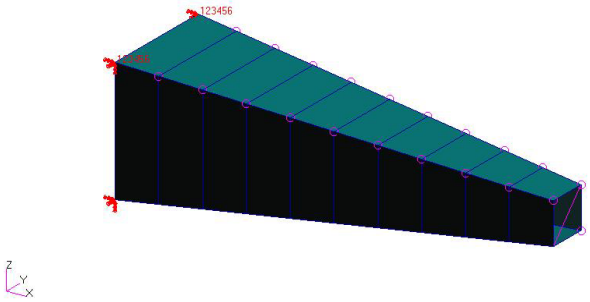


Figure 12: *Test case cantilever beam.*

the square edge of the tip section side C_{tip}, the length l and the thickness t. A value for the initial design of $C_{tip} = C_{root} = 3m$, length $l = 10m$ thickness $t = 1mm$ has been assumed, and a linear trend from root to tip value of the section geometry is assumed too. A constraint on the maximum value of the rotation of the extremity free section ϑ_{tip} has been also applied. So the optimization problem consists in 2 contrasting objectives, 3 design variables (1 structural and 2 shape variables) with 1 constraint.
Solving the optimization problem with three different WGC formulation:

$$\Phi_1 = w_1 \frac{W}{W_0} + w_2 \frac{V_0}{V} \quad (11)$$

$$\Phi_2 = w_1 \left(\frac{W}{W_0}\right)^2 + w_2 \left(\frac{V_0}{V}\right)^2 \quad (12)$$

$$\Phi_5 = w_1 \left(\frac{W}{W_0}\right)^5 + w_2 \left(\frac{V_0}{V}\right)^5 \quad (13)$$

for:

$$w_1, w_2 = [0, 1] \qquad with \quad \Delta w = 0.1 \quad (14)$$

with $\sum_{i=1}^{n} w_i = 1$, Pareto frontier in Figure 13 is obtained. Note that the projects obtained with

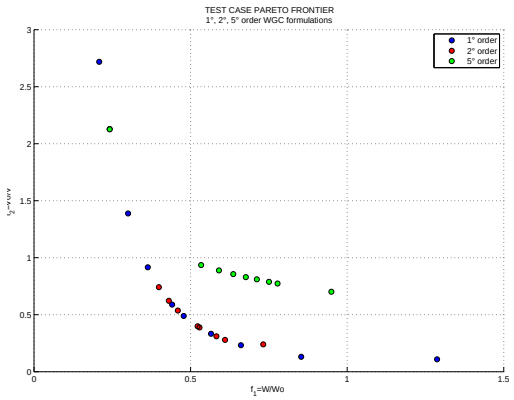


Figure 13: *Pareto frontier for the test case problem.*

the formulation of the fifth order does not comply with the condition of Pareto optimality and thus not part of the Pareto frontier.
Figure 14(a), 14(b), 14(c) shown the shape functions tangent to the frontier, whose designs are part.

References

[1] V. Pareto,*Manuale di economia politica con una introduzione alla scienza sociale*, Piccola biblioteca scientifica, Milano, 1906

[2] M. Ehrgott, *Multicriteria Optimization*, 2005, Springer editor.

[3] R.T. Marler, J.S. Arora, "Survey of multi-objective optimization methods for engineering", *Structural Multidisciplinary Optimization*, 26, 2004, pp.369–395.

[4] T.W.Athan, P.Papalambros, "A note on wheighted criteria methods for compromise solution in multiobjective optimization", *Engeneering Optimization*, Vol.27, 1996, pp.155–176.

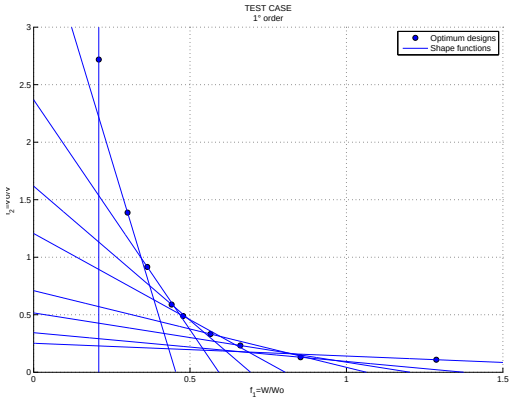
[5] Fransen S.H.J.A., "Conceptual Design of a 70 Passenger Airliner Propelled by Fuel-Efficient Turbofan Engines", *Memorandum M725*, Technical University of Delft, March, 1996.

[6] P. Arendsen, "Final Report of the Garteur Action Group SM AG-21 on Multi Disciplinary Wing Optimisation", NLR-TR-2001-557, 2001.

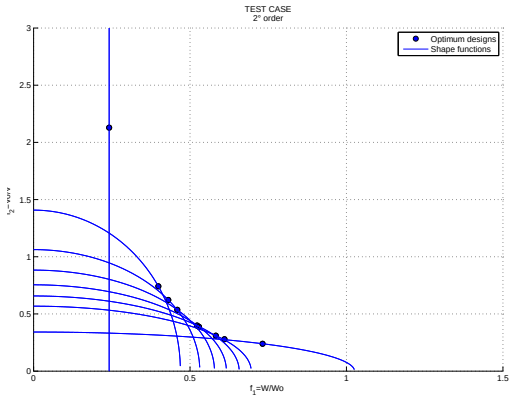
[7] Philip E. Gill, Walter Murray, Michael A. Saunders, *SNOPT: An SQP Algorithm for Large-Scale Constrained Optimization*, SIAM Review, vol.47, n.1, 2005, pp.99-131.

[8] W. P. Rodden, P. F. Taylor, S. C. McIntosh Jr., *Aeroelastic Analysis User's Guide MSC.NASTRAN*

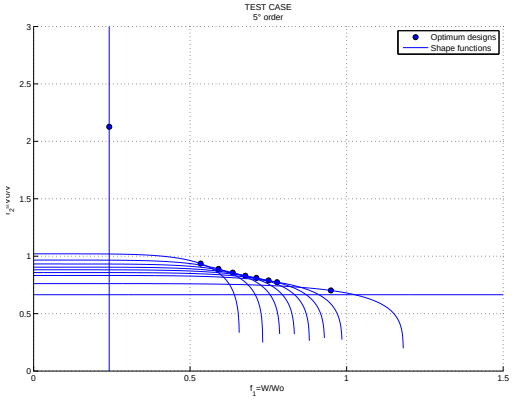
[9] Mastroddi, F., Tozzi, M., Capannolo, V., "On the use of geometry design variables in the MDO analysis of wing structures with aeroelastic constraints on stability and respons", *Aerospace Science and Technology, Vol. 15, No. 3, 2011, pp. 196-206.*



(a) $p = 1$ case



(b) $p = 2$ case



(c) $p = 5$ case

Figure 14: *Shape function and optimum designs for test case problem.*

HANDLING OPTIMIZATION PROBLEMS ON AN EXAMPLE OF MICRO UAV

Jacek Mieloszyk
Warsaw University of Technology
st. Nowowiejska 24/330
00-665 Warszawa
jmieloszyk@meil.pw.edu.pl

Abstract

Optimization techniques became very common in airspace science in recent days. Designers reached high level of efficiency in their constructions so it became difficult to introduce much butter designs. To stay competitive on the market, designers have to use more sophisticated solutions, than they use to, to gain small improvement, or take some risk entering new technologies which are not yet fully understood. This paper will show an example of optimization methodology applied to a real world problem of designing a MAV.

Keywords: MAV, Design, Optimization

integrated autopilot may help making the MAV resistant on wind gusts and atmospheric turbulence, by actively dumping disturbances [6].

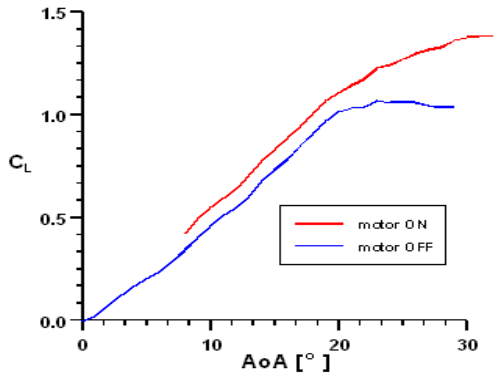


Fig. 1 Lift coefficient characteristics

Unfortunately configuration, with propeller working in a slot, produces additional drag. The main objective of this research will be reduction of drag for cruise conditions, without degradation of the advantageous features of the MAV. Obtaining useful MAV solution, applicable in true field conditions, was also very important. Contrary to theoretical examples of optimization tasks, real world objects generate far more design problems. Some constraints meeting requirements of manufacturing and field operations has to be applied. Defining such constrains is connected with choosing appropriate design variables and setting limits for them basing on engineering assumptions.

Equipment components distribution

Initial placement of components provided appropriate plane balance, but did not include connections between the components. Before starting the optimization payload components needed to be arranged Fig. 2. Limited volume and slot in the wing were additional issues, which had to be taken into account. Internal volume of the plane could be enlarged, but this is opposite the condition for minimal size of the aircraft for easy transport and storage. Size of payload components is easily determinable, but space for wires they occupy isn't. Sockets of the equipment make desired volume even bigger. Wires which lie too close to servomechanisms can make them stop working. The problem arises even more during thirst experimental flights, when additional equipment and sensors are attached to the aircraft. Place for this additional equipment isn't normally provided Fig. 3.

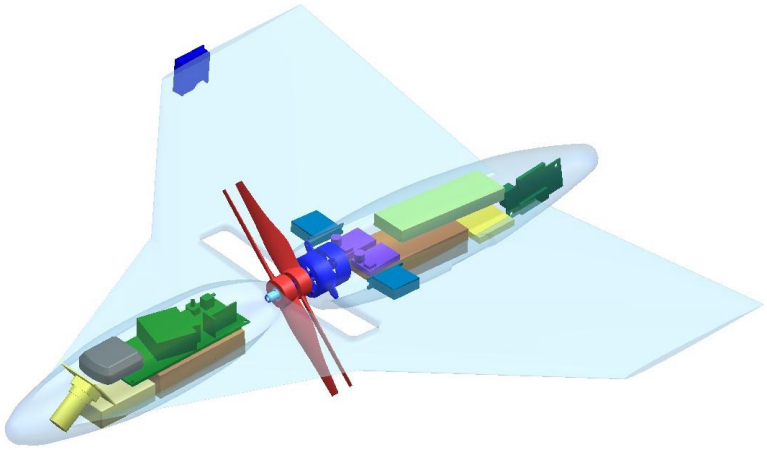


Fig. 2 Payload distribution in the MAV

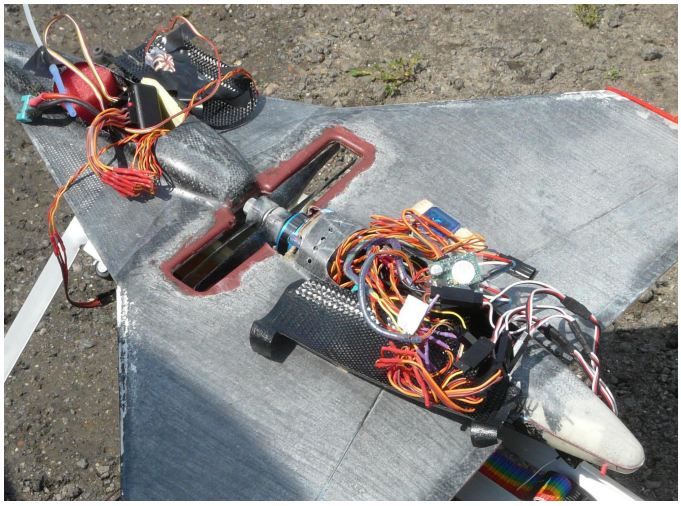


Fig. 3 Wires to the equipment components and additional sensors.

The most critical place in the designed aircraft is place between the slot and the leading edge where all needed connections between equipment components in front and rear part, must go through. There are two such places on the left wing and the right wing. It has to be pointed out that wires from main batteries, through which big currents go through, can excite currents in neighboring signal cables and become source of electric noise. To avoid this situation, signal wires and wires from power supply should lie separately on the opposite sites of the slot. Proposed solution of wiring on the MAV shows Fig 4.

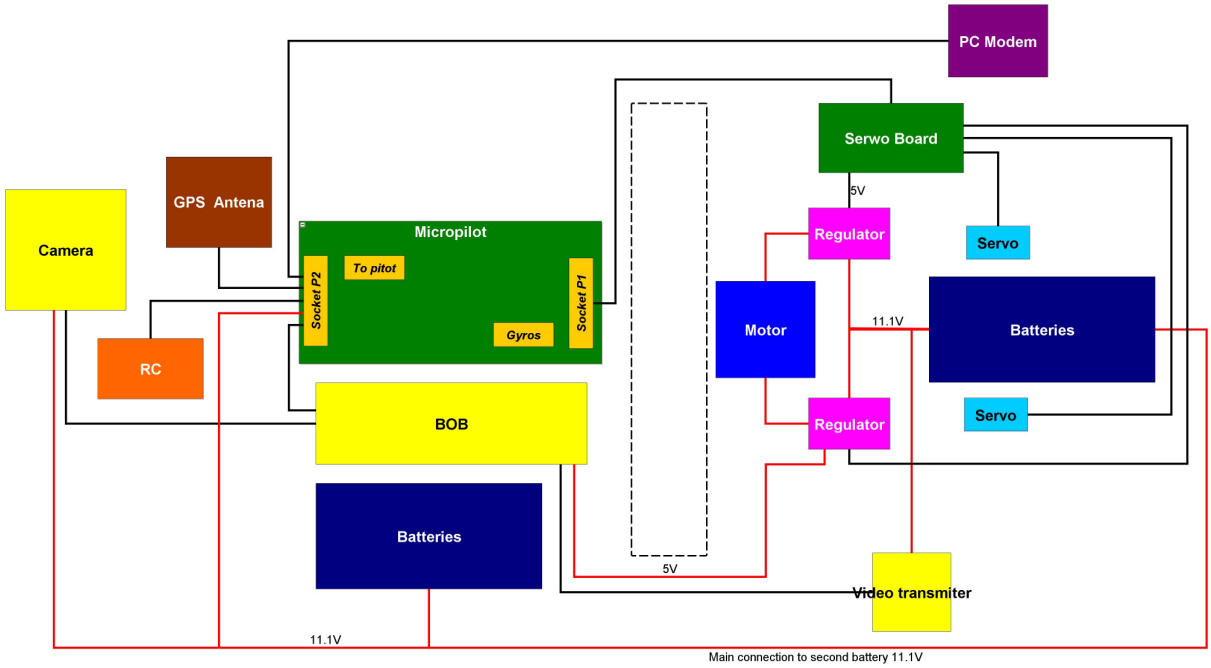


Fig. 4 Wiring on the designed MAV.

Aircraft – propeller interference

Aerodynamic flow is very complicated in the area of the slot, where strong interaction is present between the counter rotating propeller and the wing. This paragraph treats about quantification of generated drag and it's initial reduction, by testing different types of slot edges. 2D numerical simulation with application of "Fluent" software is done and compared with clean wing configuration, which shows how big is the drag penalty because of the slot and what is the most desirable edges geometry. Example of the investigated geometry shows Fig 5.

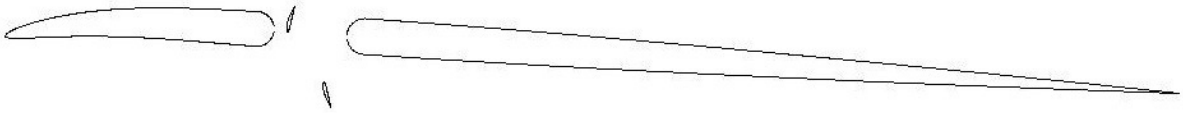


Fig. 5 Example of investigated 2D geometry.

Obtained results are shown on Fig.6. Symbols on the graph refer to the types of edges of the slot. For example: sharp edge in the middle of the wing, in the upper part of the wing, in the lower part of the wing, rounded edges. Two horizontal lines refer to the clean wing without propeller and propeller without presence of the wing. The last configuration was calculated for comparison reasons. For every geometry configuration, set of calculated and normalized parameters is shown. That is: wing's lift coefficient, drag coefficient, aerodynamic efficiency, longitudinal moment coefficient and propeller's thrust and drag coefficient. Values of the parameters were normalized, every value was divided by arithmetic mean for easy comparison issues, to show them on one chart.

Comparing configurations with each other indicates, that thrust coefficient and drag coefficient of stand alone propeller are two times lower than in the rest of the configurations. What is interesting ratio of the thrust coefficient to drag coefficient of the propeller doesn't change. It means that the propeller's aerodynamic efficiency is insensitive for presence of the slot in the wing. In almost every configuration, with slot present, lift coefficient of the wing is higher than for configuration without the slot, but in consequence drag and pitching moment also increase. High positive pitching moment, which in the used coordinate system rises nose of the aircraft, has negative effect, because it increases trim drag of the MAV. It has to be compensated by elevons to

attain appropriate balance. For the investigated cruise conditions, on small angles of attack, it is important to reach MAV's high aerodynamic efficiency, which has direct impact on the flight endurance. From Fig. 6 one can see that definitely the highest aerodynamic efficiency has wing without the slot and it is approximately 2.5 times higher than in other configurations. It is price paid for above average maneuvering capabilities. Without propeller working in the slot the MAV couldn't have such good lift coefficient characteristics as presented in Fig 1. Comparing only configurations containing the slot the best aerodynamic efficiency has configuration with both sharp edges on the lower side of the wing. It occurred that in that case propeller blades pass each other directly in the area of the sharp edges of the slot. It has to be remembered that in the simulation propeller's blades speed and position were controlled. In reality electric motors used don't have any synchronizing mechanism and the blades can get through the slot in different arrangement. It is even possible that the front and rear propeller will pass through the slot separately. It can be also seen that configuration with the rounded slot edges is almost as efficient as configuration with the sharp edges. In the case of rounded edges any position by which blades pass can be considered as the critical point. It means that without big loss on efficiency more general case can be chosen, that will meet condition for any position arrangement of the propeller's blades.

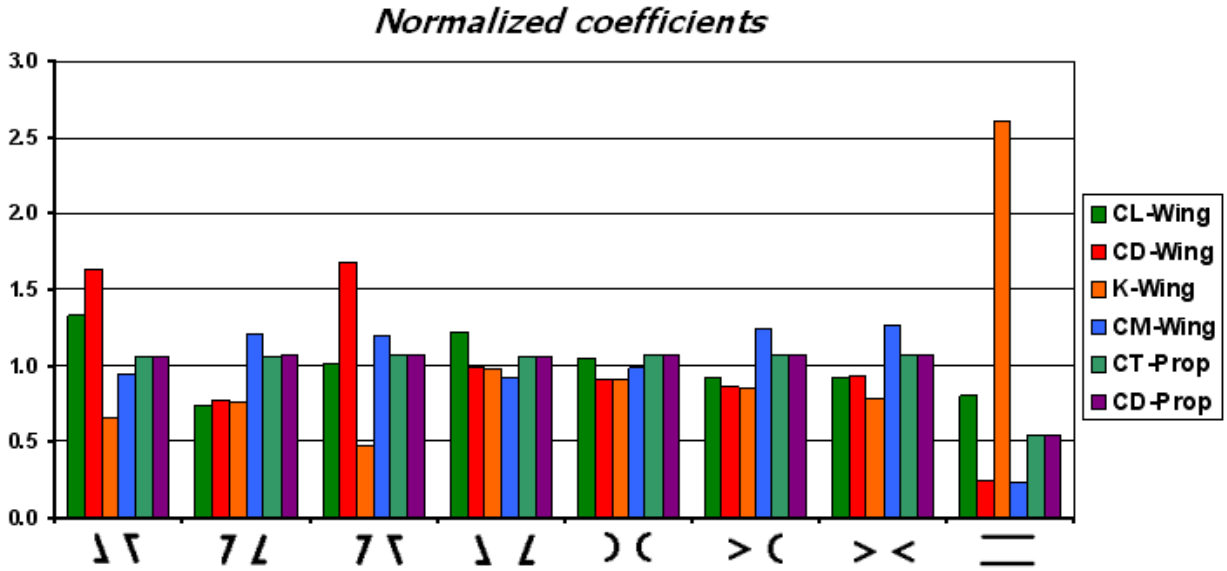
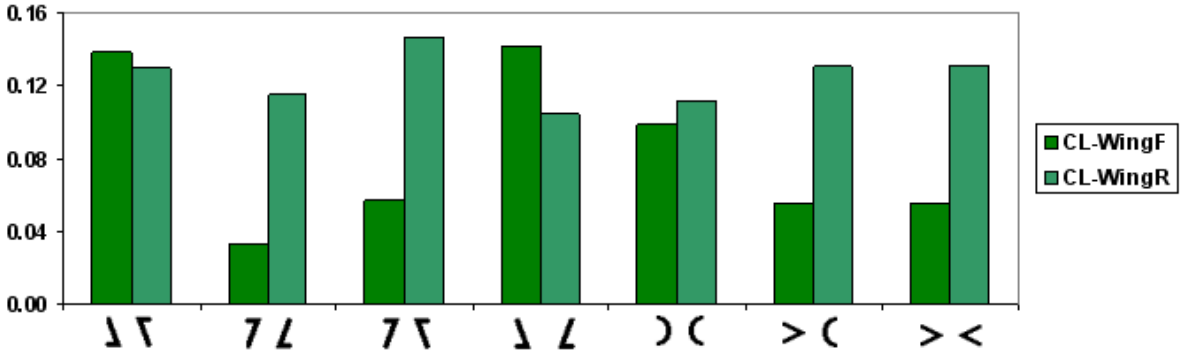


Fig. 6 Results for 2D geometry.

Influence of the propeller blades on the front and rear part of the wing was checked and is presented on Fig 7. Lift coefficient on the front and rear part of the wing can differ much for individual configurations, but the drag coefficient is always higher on the rear part of the wing. Rear edge of the slot has crucial meaning for drag reduction.



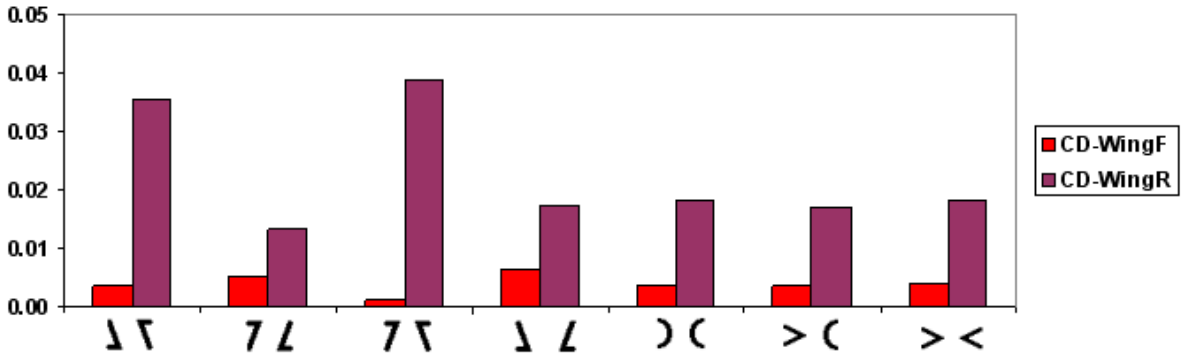


Fig. 7 Lift and drag coefficient on the front and rear part of the wing for different slot configurations.

Below values of the calculated parameters for clean wing and stand alone propeller are compared with the most practical configuration chosen with rounded slot edges.

	Clean wing, propeller	Rounded slot
CL	0.1625	0.21
CD	0.0058	0.0218
CM	0.0037	0.0152
K	27.8	9.7
CTprop	0.193	0.39
CDprop	0.0513	0.1025
Kprop	3.77	3.8

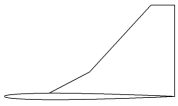
Observations from test flights

Having unoptimized version of the MAV, set of vertical stabilizers was investigated during an outdoor test flights. MAV which was used for tests can be seen on Fig 8. Basing on the output from the data acquisition system and experienced pilot opinion first general conclusions were drown. Number of test flights were conducted, with varying stabilizers geometry. Type of vertical stabilizers used had significant impact on the lateral characteristics of the aircraft. Types of vertical stabilizers used and their influence on the plane are presented below.

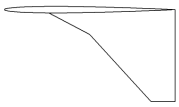



Fig. 8 MAV during test flights.

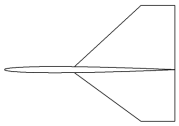
Type of vertical stabilizer used:



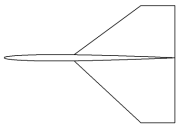
small upper stabilizer, strong undumped Dutch roll oscillations



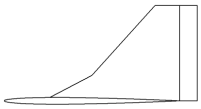
small lower stabilizer, no oscillations, small spiral divergence, higher drag than for upper stabilizer



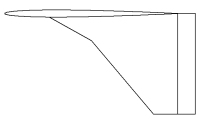
bilateral stabilizer with bigger upper surface, fine flying characteristics



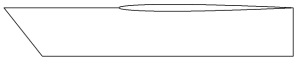
bilateral stabilizer with bigger lower surface, fine flying characteristics



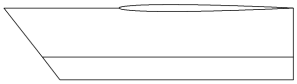
enlarged upper stabilizer, butter dumped Dutch roll oscillations



enlarged lower stabilizer, no Dutch roll oscillations, small spiral divergence, higher drag than for upper stabilizer



single stabilizer on the centerline of the plane, below fuselage, strong coupling of directional and lateral instability



enlarged single stabilizer on the centerline of the plane, below fuselage, lateral instability and trimming shift

Pilot was able to control the aircraft in almost every configuration. Upper vertical stabilizers at the ends of the wing favor lateral oscillations with simultaneous tendency to bringing out from roll during the turn. Lover vertical stabilizers don't excite lateral oscillations, but cause small spiral divergence, which however can be easily corrected by the pilot. As the final solution bilateral stabilizers were chosen, which provided sufficient flight stability characteristics and limited drag. Choosing the type of stabilizers based on flight experiment reduced design variables and simplified optimization task.

Posing optimization task

Uncertain points of the design were set and optimization task of reasonable complexity could be defined. Author was aware that conducting successful optimization solution is not easy. The first optimization was conducted on an inviscid solver with some simplifications imposed. Dropping viscous solution for the moment greatly speeded up computations and enabled to make tests with different solvers and solver settings. The additional simplifications were: coarse grid, no vertical stabilizers, no propeller and slot in the wing, no vortex lift – small angle of attack assumption.

The objective was defined as minimization of total drag for cruise conditions. Design variables were: angle of attack, length of the tip chord, wing sweep and parameters, which controlled nonlinear wing twist distribution defined as a fourth order polynomial Fig. 9. As a result of the optimization tests variable defining position of center of gravity was later added.

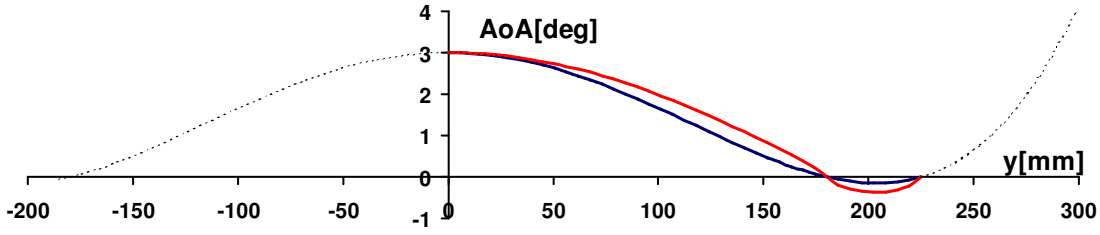


Fig. 9 Example of the wing twist distribution.

To enforce obtaining real solutions constrains on longitudinal static stability were set using penalty function method. This method of making constrains is sufficient for gradient [7] as well as for genetic algorithms [8], which were used in this work. Objective function returned to the optimization algorithm was defined by equation (1) and included constrains on balance of vertical forces (2) and constant static stability coefficient (3). μ is predefined scalar by the user.

$$F_{objective} = C_x + P_1 + P_2 \quad (1)$$

$$P_1 = 0.5 \cdot C_1^2/\mu; \quad C_1 = (m \cdot g) - (0.5 \cdot \delta \cdot V^2 \cdot S \cdot C_z) \quad (2)$$

$$P_2 = 0.5 \cdot C_2^2/\mu; \quad C_2 = -0.1 - dC_m/dC_z \quad (3)$$

The first optimization results showed that the optimization task was over constrained. Solution satisfied stability constrains by varying angle of attack, but the geometry parameters didn't change significantly maintaining the old geometry. After this experience variable of position of center of gravity was added. This way geometry could vary still satisfying the constrains.

For optimization genetic and gradient algorithms were used, both giving corresponding results. Genetic algorithm was very robust and gave always solutions. Contrary to that experience gradient algorithm needed quite some time to set be set to start solution to converge, but after it was done it converged much faster than genetic algorith. The best type of gradient solver was second order Newton method. Although this algorithm

needed second order derivatives it's estimation of search direction was so fine that it's efficiency overtaken Steepest Descent, Conjugate Gradient and Quasi Newton gradient methods. Interestingly genetic algorithm, which can theoretically lead to random solutions, showed that two competitive solutions are possible, with the current optimization task definition, with completely different geometry Fig 10.

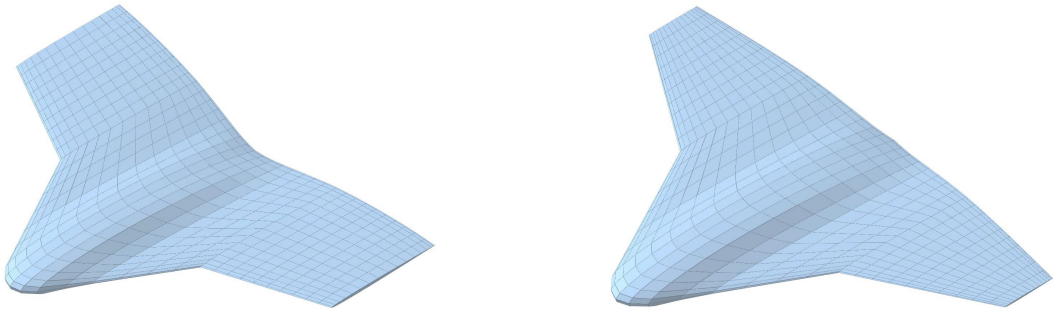


Fig. 10 Competitive solutions from the optimization.

Dynamic stability analysis

The obtained solution, from the first optimization tests, revealed question are the configurations dynamically stable in all modes? To have confidence about it simplified dynamic stability analysis were made [9]. Aerodynamic analysis were made, on viscous solver VSAero, to obtain aerodynamic derivatives from angle of attack and side slip. All remaining dynamic derivatives were calculated using DATCOM reports [10]. Actually obtained MAV solutions from inviscid optimization were investigated. Vertical stabilizers were added to the plain, grid was refined, propeller was added in a form of flow field disturbance. There was no vorticity and low angle of attack assumption stayed in force. It became clear that the investigated configuration is stable longitudinally and has unstable Dutch roll oscillations.

This problem needed a closer look and further computations were done. Using pure DATCOM methods wide analysis of dynamic lateral stability were made. Estimations of aerodynamic derivatives in DATCOM reports depended on wing sweep, aspect ratio and taper ratio, so the direct impact would have optimization variables of wing sweep and tip chord. Root chord of the wing and the wing span were constant. Utilizing DATCOM methods carpet plots of Dutch roll damping coefficient and lambda parameter of spiral stability from wing sweep and tip chord were made Fig. 11.

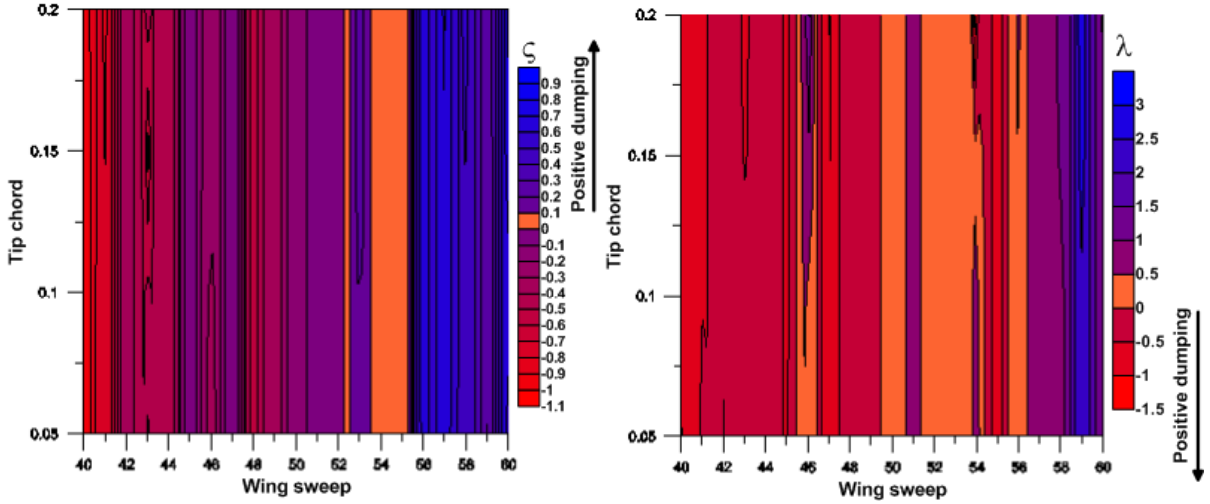


Fig. 11 Lateral dynamic characteristics from the wing sweep and the wing's tip chord.

It can be seen that the tip chord has less impact on lateral stability compared to the wing sweep. Orange color in a color scale shows place were stability is neutral and divides stable region from instable. Wing sweep of about

54deg indicates border between the regions. It would mean that the real plane isn't capable of flying, nevertheless the test flights showed it flies quite well. Analysis made by DATCOM methods are only an approximation. They can introduce large errors, specially for such a small object like the MAV, for which they weren't particularly designed. Moreover LEX part of the wing wasn't included in the DATCOM computations and it means that the real plane has much bigger average wing sweep. This methods can't directly answer if the MAV will be dynamically stable, but still provide some useful information. Wing sweep has crucial impact on the lateral stability and is equally important as the influence of the vertical stabilizers. Crossing between Dutch roll instability and spiral divergence is very rapid. This assumption is additionally confirmed by observations from the test flights, were even small changes in vertical stabilizers geometry changed lateral stability characteristics.

This considerations can be widened on a wider group of MAVs. Demand on miniaturization drives MAV designs into configurations, which have low aspect ratios below two and have small moments of inertia. Very often they are built in a flying wing configuration, were vertical stabilizers aren't far from the center of gravity. If the vertical stabilizers are mounted at the wing tips, the efficiency of the vertical stabilizers depends also partially on the wing sweep. The conclusion is that the MAVs in a configuration of flying wing should have significant wing sweep to maintain favorable lateral stability characteristics.

Utilizing the derived conclusions and knowing about good flying characteristics of the current MAV, the wing sweep design parameter was frozen and it's value was taken from the flying MAV. As the wing tip chord had modest influence on the lateral dynamic characteristics of the MAV this design parameter remained free to change.

Numerical optimization

Gradient optimization was faster than genetic one and since results from the previous inviscid optimization were available, it was used at first. It occurred that gradients obtained by numerical differentiation computed with the viscous solver, had unrealistic values and optimization algorithm couldn't determine appropriate search direction. All the following computations were done using genetic based algorithms.

After obtaining the first results from the optimization, which used viscous solver VSAero, new conclusions were drown. Analyses made during optimization showed areas of unfavorable pressure gradients, which couldn't be controlled by any defined optimization variable. Four additional variables were added, which modified tips of the wing and added filet in the center part of the wing Fig. 12. The newly introduced variables defined quarter areas of ellipse, which was subtracted or added to the wing area.

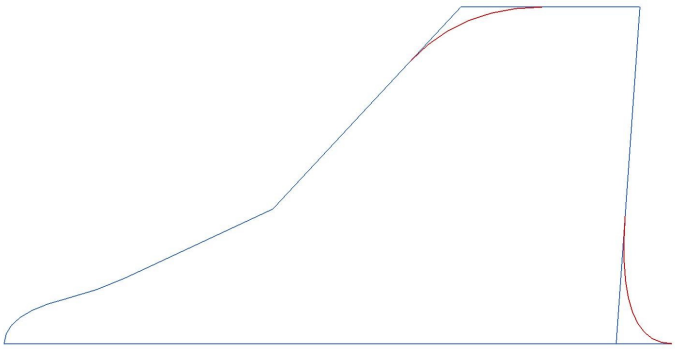


Fig. 12. Geometry modification with additional variables.

Soon it became clear that stability constrains dominate the solution and that the drag couldn't be minimized sufficiently. Weight coefficient of one hundred was added to level magnitude of total drag coefficient and values of penalty constrains (4). Equation (4) is a modification of equation (1).

$$F_{objective} = 100 \cdot C_x + P_1 + P_2 \qquad (4)$$

The optimized MAV meets all optimization constrains and fulfills requirements for bigger internal volume for equipment components. Shape of the MAV became more smooth and is more practical for maintenance and producing composite structures from which the MAV is made. The total drag coefficient, for the cruise speed stayed approximately constant and is equal to 0.0595. Fig. 13 show the final optimized geometry. Optimization with changed airfoils, to reduce drag more, is still on going.

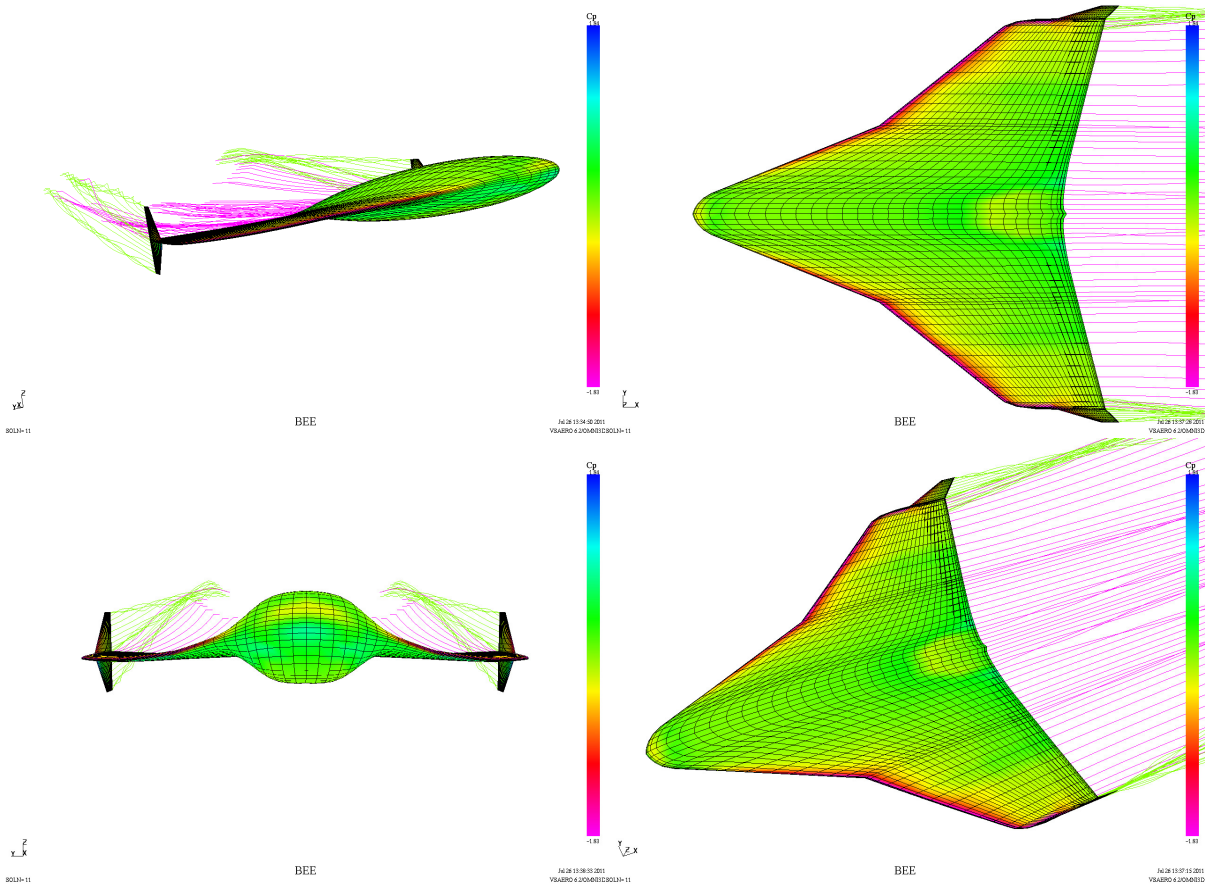


Fig. 13. The final optimized geometry.

Conclusions

It was shown that optimization task needs careful planning. Designer should include all needed constrains, with appropriate amount of flexibility needed for optimization algorithm to converge and get realistic solution. Number of optimization variables can be often significantly reduced, by using engineering assumptions.

In case of MAVs connections between equipment components play important role. They can determine need for additional internal volume.

Different types of slot edges, where wing interferes aerodynamically with the propeller, were studied and the best one for the current design was chosen.

General conclusions about lateral stability for MAVs in configuration of flying wing were drawn, showing that such aircrafts need significant wing sweep to maintain favorable lateral dynamic characteristics.

Final results from genetic optimization satisfy all constrains, new structure will improve maintenance and the total drag remained on the constant level.

Acknowledgments

This work was supported by Polish Ministry of Science and Higher Education through the grant O N509 025836.

References

[1] Morris, S. J., "Design and Flight Test Results for Micronized Fixed-Wing and VTOL Aircraft," *Proceedings of the FirstInternational Conference on EmergingTechnologies for Micro Air Vehicles*, Georgia Institute of Technology, Atlanta, GA, February 1997.
[2] Bovais, C., Mackrell, J., Foch, R., Carruthers, S., "Dragon Eye UAV: Concept to Production," *Proceedings of UAVs XVIII International Conference*, Bristol, UK, 31 March - 2 April 2003, pp. 3.1-3.12.
[3] Grasmeyer J. M., Keennon M. T.,"Development of the Black Widow MicroAir Vehicle," AIAA Paper 2001-0127, January 2001
[4] S. Watkins, W. Melbourne "Atmospheric Winds: Implications for MAVs", proceedings of the XVIII International UAV Conference, 31 March – 2 April 2003, Bristol, UK
[5] S. Watkins, G. Vino "The Turbulent Wind Environment of Birds, Insects and MAVs", *15th Australasian Fluid Mechanics Conference*, Sydney, Australia 13-17 December 2004
[6] C. Galiński, J. Mieloszyk, J. Piechna „Prograss in the gust resistant MAV programme", *27th International Congress of the Aeronautical Sciences,* Nicea 19-24.09.2010
[7] J. Nocedal, S. J. Wright „Numerical Optimization" Springer, 1999
[8] D. P. Raymer "Enhancing Aircraft Conceptual Design Using Multidisciplinary Optimization" Doctoral Thesis, 2002
[9] R.C. Nelson "Flight Stability and Automatic Control" McGraw Hill, 1998
[10] D. E. Hoak, D. E. Ellison, L. V. Malthan "USAF Stability and Control DATCOM" Douglas Aircraft Company, 1965

3rd CEAS Air&Space Conference
21st AIDAA Congress

Mesh morphing applied to transonic wing optimization

E. Fabiano, G. Quaranta†
Politecnico di Milano - Dipartimento di Ingegneria Aerospaziale, Italy
A. Colbertaldo‡ **and L. Lanzi**‡
iChrome Ltd - Bristol (www.ichrome.eu)
e–mail: †quaranta@aero.polimi.it
‡angelo.colbertaldo,luca.lanzi@ichrome.eu

Keywords: aircraft design, shape optimization, mesh morphing, computational fluid dynamics.

Abstract

The paper present a mesh morphing technique based on Moving Least Square (MLS) used for aircraft aerodynamic shape optimization. The objective in this case is the drag minimization of the main wing of the three lifting surface configuration. Since the crucial advantage of a three surface configuration stays in the theoretical ability to trim the aircraft with the minimum induced drag possible the trim problem is not tackled independently but is part of the drag minimization problem. An initial sample problem is presented here, where only the movable surface position and the wing twist distribution are considered has parameters for the optimization. However, it will be shown how the MLS mesh morphing possess a great potential to solve shape optimization problems applied to complex geometries.

1 Introduction

Aircraft design is a long iterative process involving different fidelity models at different design phases. At the beginning of the design process designers just have a basic idea of the overall aircraft so that simple analytical model as those described in are needed. As the design keeps going aircraft models get more and more refined and this is true for the wing too. Focusing on the aerodynamic design of the wing, in the past designers chose main wing parameters based on experience, nowadays its common to rely once more upon optimization procedure, as thoroughly described by Jameson in [1]. In the transonic regime shock waves will most likely be present. Indeed the goal of a pure aerodynamic optimization is a shock-free wing, and since these features can only be captured by advanced CFD methods, these represent the basic tool for any aerodynamic design, even thought they are still expensive in terms of computational time required. The classic approach to shape optimization requires the definition of a geometrical parametrization of the Computer-Aided Design (CAD) model. In this case, whenever a parameter is changed, a new mesh must be generated for the new configuration. For large three-dimensional models this operation can have a very high computational cost.

With the mesh morphing approach proposed in the following, the parametrization is defined directly on the numerical model used for calculations, which is generated only once at the beginning of the procedure and is then updated. Thus, there is no need to mesh the CAD model for each new configuration required by the opti-

mizer, significantly simplifying and speeding-up the optimization process. In addition, this is often the only feasible approach in applications where the requirements on the quality of the final mesh make it difficult to fully automate the grid generation process.

The morphing methodology is based on Moving Least Square (MLS) technique used in surface reconstructions from scattered data [2]. The technique has already been successfully applied to fluid-structure interface problems [3]. The MLS scheme is based on locally supported reconstruction operators that allow to efficiently deform the original mesh with a limited number of control points, without changing the grid topology and leaving the user the possibility to easily control the smoothness of the deformation. The main advantages of the proposed technique are its independence from the numerical formulation of the Computational Fluid Dynamics (CFD) solver and its computational efficiency, which makes this algorithm significantly faster than other methods proposed in literature, including the state-of-art Radial Basis Function (RBF) morphing [4]. Compared to other strategies, the MLS morphing also preserves in a more accurate way the initial quality of the mesh. A good quality of the domain discretization is an essential requirement in CFD when an accurate estimation of aerodynamics forces, especially of drag, is desired [5].

As a proof-of-concept, the potential of this new approach is exploited for the drag minimization of the main wing of the innovative three lifting surface configuration thoroughly investigated at Politecnico di Milano [6] by means of an aeroelastically scaled wind tunnel model denominated XDIA. The aircraft currently optimized is the Target Aircraft, an upscaled variant of the XDIA model shown in Figure 1. This configuration features an anterior all movable canard, a 15 degs negative sweep mid-wing and a T-tail. Scientific debate about the three surface aircraft is endless. Many authors claim for three surfaces superiority, as Kendall [7, 8, 9], because of the theoretical ability to trim the aircraft with the minimum induced drag possible, no matter where the center of gravity is, provided both the canard and the horizontal tail

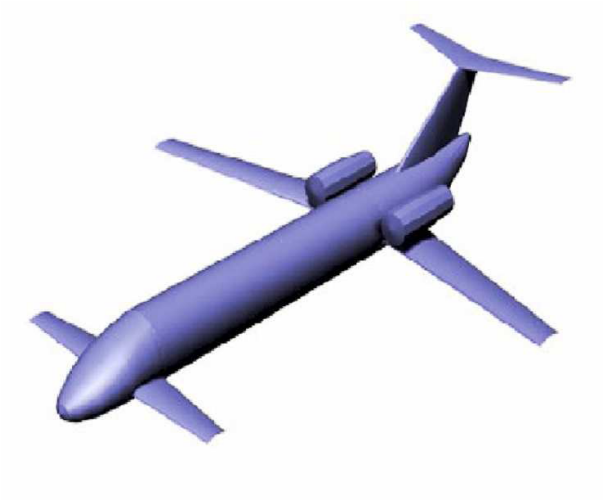


Figure 1: CAD model of the Target Aircraft

are independently operated. The situation however is not so clear-cut, drawbacks being additional weight, complexity and interference drag [10]; but it is believed that the TA configuration could results in a synergic cooperation between structure and aerodynamic toward weight saving and drag reduction [6]. The optimization looks for the wing twist of minimal drag evaluated with for the minimal drag trim configuration. So a nested optimization problem is set up that looks first for the minimal drag trim configuration and then for the optimal twist. Initially the Free Mesh Morphing technique is reviewed. Then in the following section the optimization problem is presented and solved by coupling the MLS morphing technique with a surrogate model approach based on response surface to accelerate the optimization procedure, a technique commonly adopted in CFD to replace the typical costly aerodynamics evaluations with cheaper models [11].

2 Shaper: a mesh free morphing tool

Strictly speaking Free Mesh Morphing is a general purpose technique used to deform an existing mesh — or better the nodes upon which mesh connectivity is defined — without requiring explicit information on the geometry initially used to create the mesh. More in particular Free Mesh Morphing allows to modify the position of

a set of target nodes in the space by defining a continuous and smooth deformation field to be added to the original position of the nodes. Morphing algorithms can thus be distinguished based on the way they define the deformation field to be applied. This paper focuses on the use of MLS as available in Shaper [12] - a commercial software for general purpose size-and-shape optimization. From a general standpoint, the MLS techniques belong to the field of data approximation form a scattered set of points. Applications of MLS can be found many different fields: computer graphics and visualization, image processing, regression models, supervised learning, mesh-free finite element methods are but few examples.

The problem can be generally formulated as follows [13]:

> Given N points located at positions $\boldsymbol{x}_i \in \mathbb{R}^3$ find a globally defined function $p(\boldsymbol{x})$ that approximates a set of given values $f(\boldsymbol{x}_i)$.

In order to have the maximum versatility, no special structure is supposed in the data point $\boldsymbol{x}_i$ distribution, which can be considered scattered in the space $\mathbb{R}^3$. This means that the method is essentially meshless, since no information about the nodes' connection is used. Consequently, it can be applied to any mesh type, either structure or unstructured, tetrahedral or hybrid and so on.

While this problem seems quite abstract, it may be easily applied to mesh morphing. In fact, by solving the function approximation problem it is possible to find the function that represent the mesh nodes displacement field given the displacement of few points, which will be called the *controlling handles*. A classical approach to solve this problem is represented by the Least Square ((LS) that looks for the solution of the following minimization problem

$$\min_{p \in \mathcal{C}} \sum_i \left(p(\boldsymbol{x}_i) - f(\boldsymbol{x}_i) \right)^2 , \qquad (1)$$

where $\mathcal{C}$ is a generic functional space. The problem is to find the appropriate functional space for the function $p(\boldsymbol{x})$ that represents a good global approximation of the real function $f(\boldsymbol{x})$. Usually the space of polynomial of maximum degree m is taken, i.e. $\mathcal{C} = \Pi^m$, however very often a global polynomial may not be able to represent complex local behaviors of the function to be approximated. To avoid the necessity to find a functional approximation space that is valid globally, Lancaster and Salkauskas proposed the MLS technique [2]. The idea is to perform individually a Weighted Least Square for each point $\boldsymbol{x}$ in the domain. So, the global approximation $p(\boldsymbol{x})$ is obtained from the union of a series of local functions i.e. $p(x) = \bigcup_{\boldsymbol{x} \in \mathbb{R}^3} p_{\boldsymbol{x}}(\boldsymbol{x})$, where the functions $p_{\boldsymbol{x}}(\boldsymbol{x})$ are obtained as solution of the following minimization problem:

$$\min_{p_{\boldsymbol{x}} \in \Pi^m} \sum_i \left(p_{\boldsymbol{x}}(\boldsymbol{x}_i) - f(\boldsymbol{x}_i) \right)^2 \theta \left(\| \boldsymbol{x} - \boldsymbol{x}_i \| \right) . \qquad (2)$$

At first glance, Eq. (2) looks very similar to (1); however, (2) contains a crucial difference: the approximation function $p_{\boldsymbol{x}}(\boldsymbol{x})$ depends continuously from the point at which is computed $\boldsymbol{x}$ thanks to the weight function θ.

The function θ is a non-negative weight function that depends only from the Euclidean distance between points in the solution space $r = \| \boldsymbol{x} - \boldsymbol{x}_i \|$. The approximation can be further localized if $\theta(r)$ is rapidly decreasing as $r \to \infty$. Extremely useful as weight functions in this case are the compact support RBF functions suggested by Wendland [14]. By looking at Eq. (2) it is easy to see that if $\theta(0) \to \infty$ the MLS fit will be forced to interpolate the prescribed values at known points $\boldsymbol{x}_i$; Levin in Ref. [13] proposes special RBF function that have this property.

The MLS approach is definitely more expensive than other approach like the LS of Eq. (1), because for every point $\boldsymbol{x}$ where the solution must be computed (which means for every grid node in our case) it is required to solve a Linear System of Equations (LSEs) to find the minimum of Eq. (2). However, the size of these LSEs can be kept small by using small support dimension for the weight functions.

In order to solve the MLS problem the user has to take a decision with respect to these parameters

- The functional space at which the local approximation functions $p_{\boldsymbol{x}}(\boldsymbol{x})$ must belong.

Usually the functions are polynomial, so the user must decide the maximum polynomial order.

- The type of weight functions. This choice may have an influence on the smoothness of the obtained approximation (see [13]).
- The size of the local support of the weight function. Usually a good strategy is to fix the number of points that must belong to each support, and then adapt the local size accordingly. In this way a natural adaptation of the algorithm is realized, since where there are few spread points the support is large, while for more dense areas the support becomes smaller.

2.1 Samples

As a first test a simple meshed sphere is built, as shown in Figure 2 with several handles associates represented by red dots. By moving in a different way the handles it is possible to modify the mesh of the sphere in a large number of new shapes.

Of course it is very easy to build a hierarchy of controlling handles. For instance it is possible to have a group of master handles that controls the displacement of a set of slave handles that in turns control the displacement of the mesh nodes.

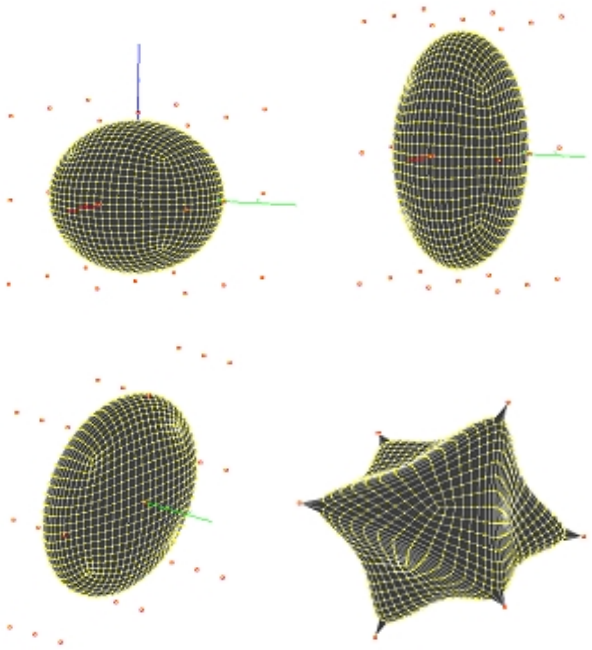


Figure 2: Mesh morphing with the MLS applied to a simple sphere. Top left: undeformed sphere. Top right: ellipsoid obtained by moving 2 handles. Bottom left: disc obtained by moving 6 handles. Bottom right: star obtained by moving 6 handles.

3 Drag minimization problem

In this section the optimization problem is described. The aircraft currently under optimization is the Target Aircraft (TA), an upscaled variant of the XDIA aeroelastic wind tunnel model. The TA three surface airplane features a 15 deg. negative sweep mid-wing with a T-tail configuration, see Table 3 for a summary of the geometric characteristics.

The goal of the present work is to optimize the twist of seven spanwise stations of the main wing of the TA and at the same time satisfy a lift and a pitching moment constraint. The cruise trim problem for classical configuration aircraft can be solved by writing the classical three longitudinal equilibrium equations: vertical and longitudinal equilibrium of forces and equilibrium of moments around the pitch axis. The solution of this problem leads to the determine the engine thrust, the angle of attack and the the elevator angle necessary. When a multiple control surface configuration is considered, like the one here investigate, the solution is in general not unique but there are several control surfaces set up that trim the aircraft. So, among all possible configuration it may be reasonable to look for the one with minimal global drag.

Based on this idea the mathematical formulation of the optimization problem may be expressed as follows, neglecting the thrust equation:

$$\begin{aligned} &\min C_D \\ &\quad \text{w.r.t.} \\ &C_L = C_{L_{design}} \\ &C_{M_{@CG}} = 0 \end{aligned} \tag{3}$$

In this way the optimization formulation Eq. (3)

Table 1: Geometric properties of Target Aircraft

Symbol	Meaning	Value
S [m^2]	Wing surface	75
b[m]	Wing span	30
AR	Wing Aspect Ratio	12
λ	Wing Taper Ratio	0.5
Λ [deg]	Wing Sweep Angle	-15
Root Airfoil	NACA	$63_2 - 215$
Tip Airfoil	NACA	$63_2 - 212$
MAC [m]	Mean Aerodynamic Chord	2.59
S_C [m^2]	Canard surface	25
b_c [m]	Canard span	14.14
AR$_C$	Canard Aspect Ratio	8
λ_C	Canard Taper Ratio	0.5
Λ_C [deg]	Canard Sweep Angle	0
Root Airfoil	NACA	$63_2 - 215$
Tip Airfoil	NACA	$63_2 - 215$
MAC$_C$ [m]	Mean Aerodynamic Chord	1.837
S_{HT} [m^2]	Tail surface	40
b_{HT} [m]	Tail span	11.8
AR$_{HT}$	Tail Aspect Ratio	7
λ_{HT}	Tail Taper Ratio	0.5
Λ_{HT} [deg]	Tail Sweep Angle	20
Root Airfoil	NACA	64 - 009
Tip Airfoil	NACA	64 - 009
MAC$_{HT}$ [m]	Mean Aerodynamic Chord	1.7519

deals explicitly with the trim problem so that its resulting wing is a wing that has the minimum trimmed drag coefficient. Equation (3) can be simplified if is interpreted as two nested optimization problems. The first is an unconstrained optimization of the twist distribution. The second, to be run every time a proposed twist distribution is computed, is a three parameters problem, namely angle of attack, canard and horizontal tail deflections, looks for the optimal trim.

Directly searching the design space would results in a huge amount of functions evaluation so that extensive use of meta-modeling techniques, namely the Kriging, has been made in this work. Three different response surfaces has been constructed, one each aerodynamic coefficient needed in optimization problem 3. Each aerodynamic coefficient is a function of ten variables, namely seven spanwise twist variables and three flight mechanics variables:

$$C_L = C_L\left(\boldsymbol{\xi}, \alpha, \delta_c, \delta_e\right) \tag{4}$$

$$C_M = C_M\left(\boldsymbol{\xi}, \alpha, \delta_c, \delta_e\right) \tag{5}$$

$$C_D = C_D\left(\boldsymbol{\xi}, \alpha, \delta_c, \delta_e\right) \tag{6}$$

$\boldsymbol{\xi}$ being the (1×7) vector of the twist design variables, α being the angle of attack and δ_c and δ_e being respectively canard and horizontal tail rotation angles. The response surfaces have been built upon a 200 sites Latin Hypercube (LH) design of experiment. A fourth Kriging of the function $F = C_D + 100(C_L - C_{L_{design}})^2 +$

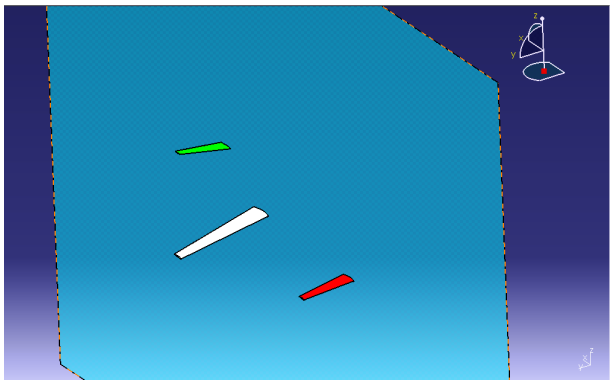


Figure 3: CAD model of the Target Aircraft

$50(C_{M_{@CG}})^2$ has been generated to add some infill point according to the Expected Improvement Approach (EIA). For the lift and pitching moment coefficients a linear mean model has been used, while for the drag coefficient the mean model was quadratic. For all aerodynamic coefficients a gaussian correlation has been used. The optimization has been performed using two numerical software: Nexus [**?**] and Matlab.

3.1 CFD and surrogate meta-models

Aerodynamic coefficients has been computed at each design sites using the Euler equations since for the minimization of the induced drag and interference drag caused by the canard on the main wing these physical model is expected to lead to correct results. Future developments that will consider also airfoil shape optimization for the reduction of drag will consider the employment of Navier-Stokes models. As CFD solver in this case as been chosen FLUENT. For the sake of simplicity the computational model, see Figure 3, consists of just the three lifting surface, and no fuselage model is included.

The surface mesh has been developed accordingly to the Drag Prediction Workshop prescriptions [5] but volume mesh is much coarser in order to gain computational speed. A run is deemed converged once at least a five order of magnitude residual drop has been achieved and the aerodynamic coefficients suffer only minor oscillations, as described in Figure 4.

Up to 800 handles have been used in the optimization. On the main wing for each of the

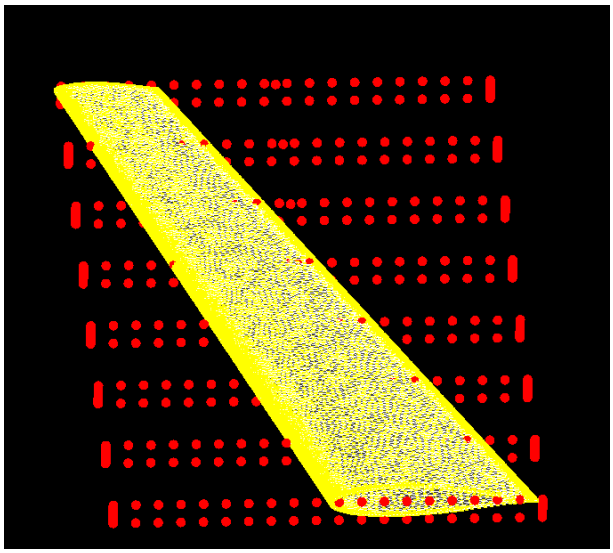


Figure 5: Shaper handles on the main wing

seven wing stations 12 handles have been distributed along the chord, equally subdivided between upper and lower surface. A thirteenth handle has been placed on each section's leading edge to define the axis of rotation of the section, which is an axis parallel to the y-axis of the global reference system passing through this handle. Hence, the rotation of this so-called master handle represent the twist of the section, i.e. one of the seven aerodynamic design variables; then the other twelve handles move rigidly following the master rotation. The handles distribution on the main wing is shown in Figure 5, while the deformed surface obtained by rotating the master handles is shown in Figure 6

For both the canard and the horizontal tail four handles have been distributed along the chord of twenty different spanwise station. A twentyfirst handle has been placed on the leading edge of the root of each control surface in order to define the axis of rotation of the whole surface as, an axis parallel to the y-axis of the global coordinate system through this handle. In this way the rotation of this handle is the rotation of the control surface, namely one of the three flight mechanics optimization variables. The remaining handles have been placed at the far field in order to prevent the far field to rotate. The global distribution of handles can be seen in Figure 7, while Figure 8 shows the de-

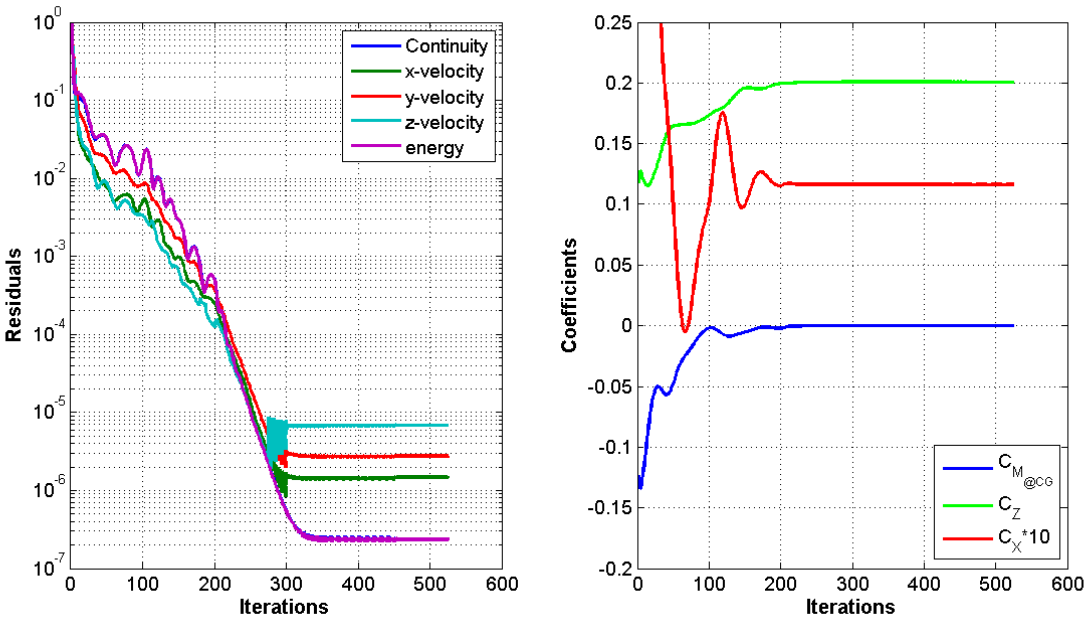


Figure 4: CFD computations convergence criteria

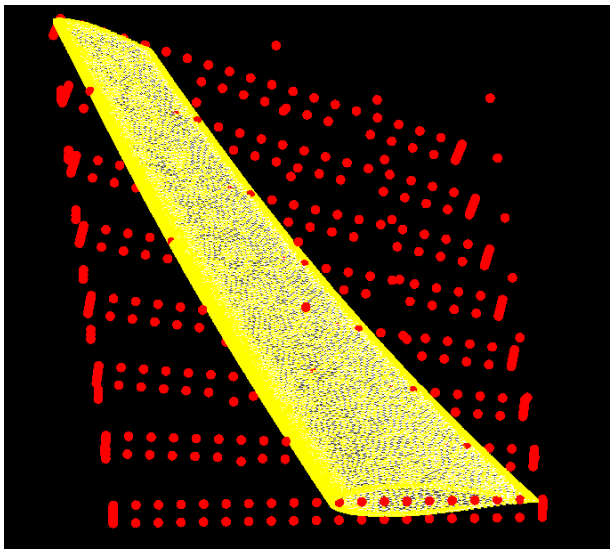


Figure 6: Deformed surface mesh of the main wing obtained by rotating the master handles

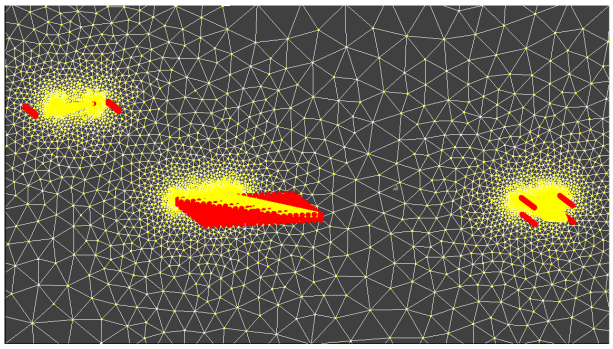


Figure 7: Shaper handles on the whole aircraft

formed surface mesh of the whole aircraft model obtained by applying different rotation to the handles.

A linear polynomial base with a 6 order RBF and a local support of 130 has been used to define the interpolator. These parameters are the best compromise between mesh quality preservation and mesh deformation accuracy.

The parameterized CAD model of the Target Aircraft has been used to validate Shaper mesh deformations by manually inspecting control surfaces' and wing sections' rotations showing good results.

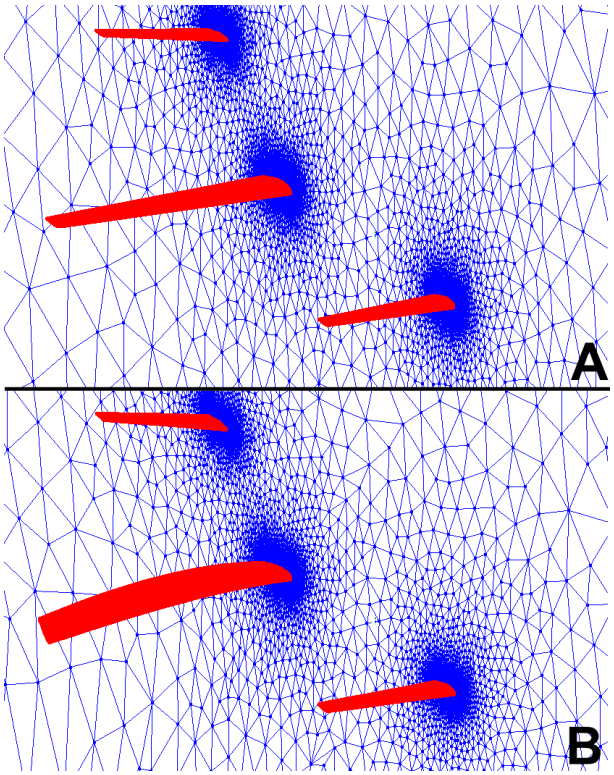


Figure 8: Deformed surface mesh of the aircraft obtained by rotating the master handles

Table 2: Cruise conditions for the target aircraft

Altitude [ft]	26000
Mach number [-]	0.67
C_L [–]	0.4
CG	MAX FWD

3.2 Results

The minimal drag configuration has been computed for the cruise condition described in Table 2. Once the three response surfaces in equation 6 have been built (see Figures 9 and 10) the aerodynamic optimization has been addressed with a genetic algorithm: each genetic-algorithm-proposed twist distribution has been trimmed with a gradient-based method. A multiple restart approach has shown little dependency on the initial guess of the flight controls design variables, suggesting a reasonably flat design space.

A comparison of the pressure coefficient distribution between the baseline and the Matlab optimized wing is shown in figure 11.

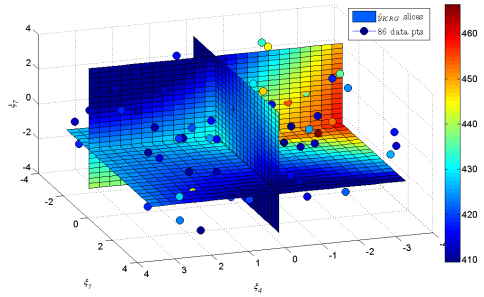


Figure 9: Slices of the flat trimmed-drag design space. Contours are in counts

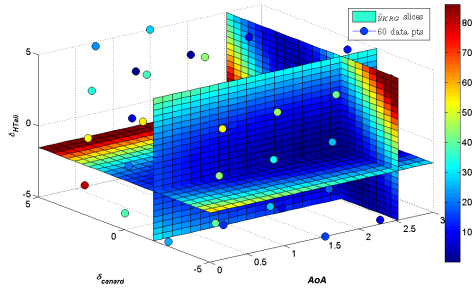


Figure 10: Slices of the flat design space of the trim function $F = C_D + 100(C_L - C_{L_{design}})^2 + 50(C_{M_{@CG}})^2$ for the optimum wing

As can be seen the baseline is itself a shock free wing so that only small improvements are possible, namely 3 counts. Comparison between the baseline and optimized twist distribution can be seen in Figure 12. Here also the twist distribution computed using Nexus is shown. The latter optimizer finds the same final drag, with a significantly different twist distribution, proving, once again, that the design space is flat.

Remarkably the optimum twist distribution is not linear. Control surfaces loading is shown in Figure 13.

It can be seen that the optimized wing gives a little more lift than the baseline and, consequently, the baseline configuration needs to use the canard more than the optimized configuration. The canard in turn gives rise to a higher positive pitching moment coefficient that need to be counteracted by the horizontal tail deflection, which again is higher for the baseline wing.

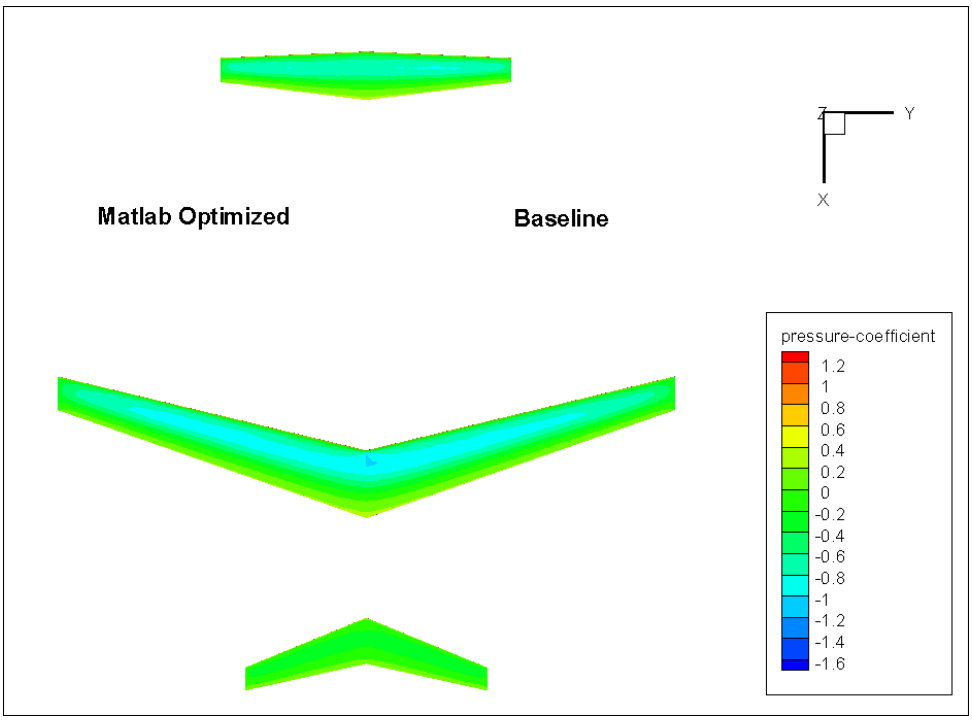


Figure 11: Comparison between the baseline and the optimized wing. A small low pressure region at the wing root has been cured by the optimization process

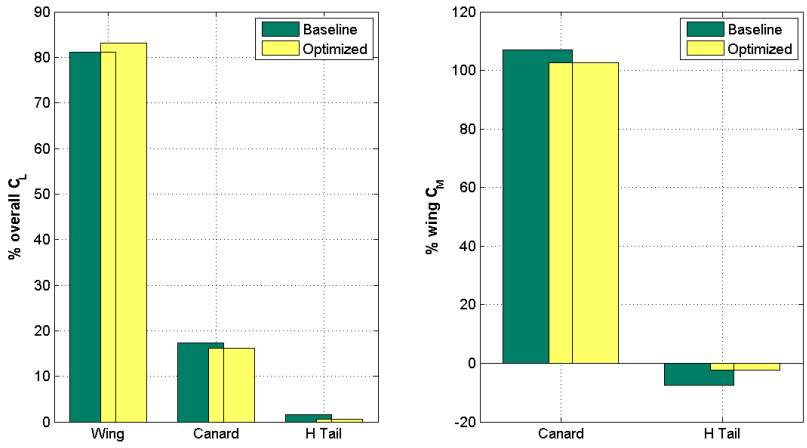


Figure 13: Control surfaces load distributions

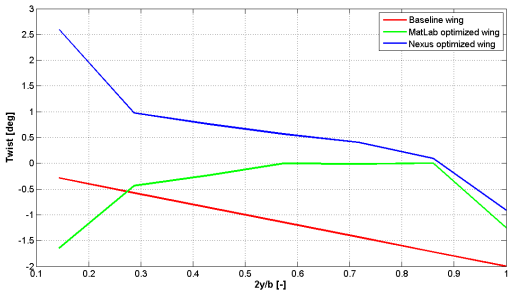


Figure 12: Comparison between the baseline and the optimized twist distribution

4 Final remarks

An innovative mesh morphing approach for shape optimization of grid-based numerical models has been presented. The application of this technique to a complex CFD problem showed a great potential of this approach given by the high flexibility joined with the large time saving in the generation of good quality meshes for modified geometry. However, a drawback that has to be considered is the fact that the final optimized geometry is only obtained in terms of meshed surface, so a post-processing procedure must be devised to transform back the obtained results into a CAD model.

Regarding the achieved results those small improvement seen are most likely due to the nature of the design variables and their limited number. Future work will consider the possibility to employ the handles to vary the airfoil shape. In this last case we believe there will be a much higher improvement once the optimization has been completed.

The optimum spanwise trimmed lift distribution is seen to load wing tips much more than wing root and is not elliptic as one might expect. The distance of each wing section from the center of gravity should be considered as an explanation. This load distribution is un- desirable from an airworthiness point of view for at high angles of attack wing tip will likely stall before the root and might lead to premature shock stall or buffet when the load is increased at high speed. However, it should be also considered that if the canard stalls before the wing then it will give rise to a nose down pitching moment that will reduce the angle of attack and prevent the wing to reach its stall angle, while still having the horizontal tail to provide for additional control.

References

[1] A. Jameson. A perspective on computational algorithms for aerodynamic analysis and design. *Progress in aerospace sciences*, **37**, 2001.

[2] P. Lancaster, K. Salkauskas. Surfaces generated by moving least squares methods. *Mathematics of Computation*, **37**, 141–158, 1981.

[3] G. Quaranta, P. Masarati, P. Mantegazza. A conservative mesh-free approach for fluid-structure interface problems. M. Papadrakakis, E. Oñate, B. Schrefler (eds.), Proceedings of the International Conference on Computational Methods for Coupled Problems in Science and Engineering. CIMNE, Santorini, Greece, 2005.

[4] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. Proceedings of the Computer Graphics (Proceedings of ACM SIGGRAPH 2001), 67–76, August 2001.

[5] E. N. Tinoco, M. Mani, O. P. Brodersen, R. A. Eisfeld, B. Whals, J. H. Morrison, T. Zickuhr, D. Levy, M. Murayama, D. J. Mavriplis, J. C. Vassberg. Grid quality and resolution issues from the drag prediction workshop series. *Journal of aircraft*, **46**(3), 2009.

[6] S. Ricci, M. Terraneo. Conceptual design of and adaptive wing for a three surfaces airplane. Proceedings of the Proceedings of the 46th AIAA Structures, Structural Dynamics and Materials Conference, AIAA-2005-1959. Austin,Texas, 18–21 April 2005.

[7] E. Kendall. The aerodynamics of three-surface airplanes. Proceedings of the AIAA, AHS, ASEE, Aircraft Design Systems and Operations Meeting. SanDiego, CA, October 31– November 2 1984. AIAA–1984–2508.

[8] E. Kendall. The minimum induced drag, longitudinal trim and static longitudinal stability of two surface and three surface aircraft. *AIAA*, (2164), 1984.

[9] E. Kendall. Performance trades of two surface and three surface configuration. *AIAA*, (2221), 1984.

[10] D. P. Raymer. *Aircraft design: a conceptual approach*. AIAA education series, 2006.

[11] A. J. Keane. Wing optimization using design of experiment, kriging model and data fusion methods. *Journal of aircraft*, **40**(4), 2003.

[12] iChrome Ltd. *Shaper: User Manuals - Keywords*. Bristol, UK, 2011.

[13] D. Levin. The approximation power of moving least-squares. *Math. Comp.*, **67**(224), 1517–1531, 1998.

[14] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.*, **4**, 389–396, 1995.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

CFD Hydrodynamic Analysis of an Ultralight Amphibious PrandtlPlane®

A. Frediani, M. Lucchesi and T. Lippi
University of Pisa, Department of Aerospace Engineering, Italy

Keywords: *PrandtlPlane, amphibious, CFD, hydrodynamics, hull design*

Abstract

This paper describes preliminary studies concerning the design of an amphibious ULM PrandtlPlane and, more in particular, the hydrodynamic optimization of the hull through a CFD analysis. The main parameters affecting the takeoff performances are: step depth, foreboby and afterbody lengths, maximum width at the chines, dead-rise angle, angle of afterbody keel, step height, step planform angle, curvature of forebody keel at the bow. The effects of these parameters are examined and a final hull configuration, taking hydrodynamic efficiency into account, is detected. The project is sponsored by the Regional Government of Tuscany (Italy).

1 Introduction

The market of amphibious aircraft is growing worldwide together with their performances; in the case of ultralight aircraft (ULM), new solutions able to improve the performances during takeoff on water have been proposed.

The regional government of Tuscany (Italy) has financed a project called IDINTOS, coordinated by the Department of Aerospace Engineering of Pisa University (Italy). This project aims to design and to manufacture a prototype of a ULM based on an innovative configuration, named PrandtlPlane ([1], [2], [3]), and to realize a scaled flying model to conduct preliminary flying tests; Figure 1 shows a very preliminary artistic view of the aircraft. The aircraft is two-seater, single-engine, with floating fuselage ("flying-boat" configuration) and retractable landing gear to fulfill the Italian and European flying regulations.

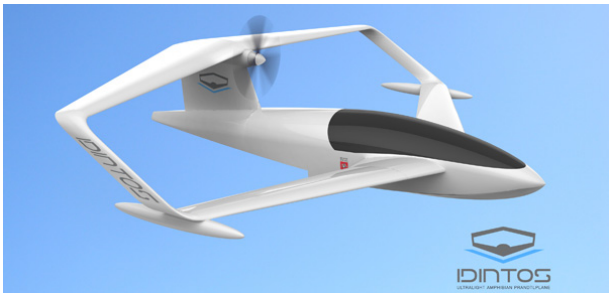


Fig. 1: PrandtlPlane amphibious ULM

The PrandtlPlane® (PrP) is the engineering application of the "Best Wing System" concept due to Ludwig Prandtl, who in 1924 ([4]) demonstrated that a proper box-like wing, made of two horizontal wings and two vertical tip wings, provides the minimum induced drag, for given wingspan and total lift. Previous studies have analyzed the application of this

configuration to commercial transport aircraft, land-based ULM and UAVs. The PrandtlPlane configuration applied to light aircraft could allow, in adjoin to high performances, to improve the safety in flight ([5]).

The present paper concerns the development of the conceptual design of an ULM amphibian aircraft and, in particular, of the fuselage, in order to optimize the takeoff capabilities. This analysis is conducted through a CFD code with the aim of carrying out the best configuration to be followed by tests on a flying scaled model.

2 Seaplanes characteristics

Amphibious aircraft are seaplanes able to operate indifferently on water or land, being provided with retractable landing gear. Seaplanes can be divided in two main categories (Fig. 2): floatplanes and flying-boats; the present PrandtlPlane is a flying-boat.

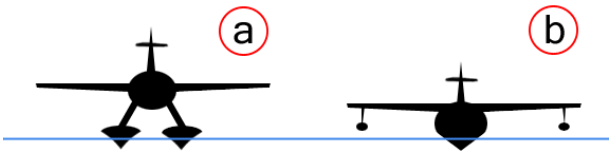


Fig. 2: (a) floatplane; (b) flying-boat

2.1 Hull design

A flying-boat fuselage (and also a single float) has a hull with a keel and a pair of hard chines, which form longitudinal edges and generate a V-shaped bottom transversal section (Fig. 3). The hard chines are introduced to create a marked line of separation between water and hull sides; the V-shaped section gives lateral stability and, also, reduces the impact with water during landing.

A fundamental element of a seaplane hull is the transversal step (named "Redan"), which creates a sharp change in the keel and the chines lines; the function of this element will be explained afterwards. The two parts of the hull separated by the step are named "forebody" and "afterbody".

Figure 3 illustrates the geometric parameters which describe a hull, namely: length (1), step height (2), step planform angle (3), dead-rise angle (4), maximum width at the chines (or maximum beam, 5), angle of afterbody keel (6), curvature of forebody keel at the bow.

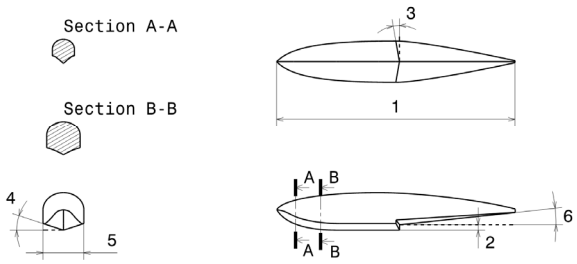


Fig. 3: design parameters of a hull

In the forebody, the dead-rise angle grows from the step section to the bow; it is well known that a smaller angle near the step determines a reduction of water drag (hereafter "resistance") during planing (when the contact with water is confined in this zone) and, also, that a higher dead-rise angle at the bow improves the ability of the hull to cut through waves in rough water at low speed (in the pitch-down asset).

2.2 Hydrodynamic behaviour

In the takeoff run of a seaplane, different phases can be identified, during which speed, trim angle and vertical displacement vary in relation to the forces applied by water on the hull and by air on the wing system (Fig. 4).

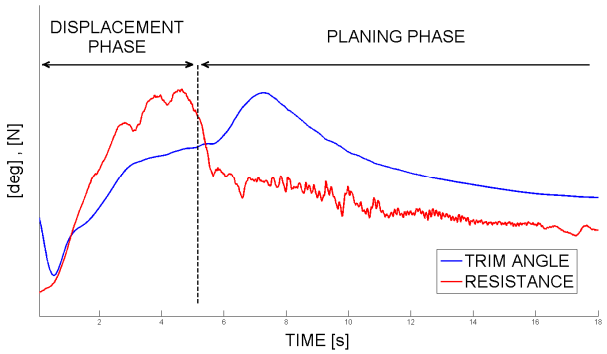


Fig. 4: typical resistance and trim angle during takeoff

At rest, the trim angle and the draft are determined by buoyancy. After the engine ignition, the trim angle is decreased by the nose-down moment introduced by the thrust; during

this phase, known as "displacement phase", the angle of attach reaches a minimum. The nose-down moment is sudden contrasted by the hydrostatic pressures (the only external forces acting during this phase) which increases the trim angle after the minimum. During the displacement phase, the resistance increases (Fig. 4); the hull accelerates as long as resistance remains smaller than thrust. As speed is increased, the water flow under the hull is not able to follow the sharp outline of the step edge and an increasing separated flow forms behind the step section, where water is replaced by air.

The wetted region moves towards the stern and the contact area is reduced consequently; thus, resistance rate diminishes until resistance reaches a maximum, called "hump" and, after that, the hull "rises on the step", the trim angle reduces as shown in Fig. 4 and resistance drops rapidly according to the decrease of the wetted surface.

The fundamental effect of the transversal step is now clear: it constitutes a discontinuity to prevent water to remain attached on the rear hull and it worths noting that, without the step, the flow would generate a resistance on the afterbody, delaying or even forbidding takeoff.

Under the increase of the speed and the variation of the trim angle, the hydrostatic pressure becomes smaller and smaller meanwhile a new phase of motion occurs, named "planing phase", where weight is progressively sustained by aerodynamic lifts on both wings and hydrodynamic lift; the hull emerges progressively. The stability of equilibrium of the aircraft when moving "on the step" is more and more influenced by the aerodynamics and the stability of flight of the wing system.

Under some particular combinations of trim angle and speed a motion of the hull, called "porpoising", may occur ([6]). Porpoising is characterized by coupled oscillations in pitch and vertical displacement of sustained or increasing amplitude occurring even in smooth water: this phenomenon causes a resistance increase, a worse aircraft's controllability and may lead to catastrophic failure. A parameter that mainly influences the occurrence of porpoising is the distance between the center of gravity (CG hereafter) and the step section.

The spray is another issue to be considered during the takeoff run. Spray are generated from the stagnation line on the forebody and are deflected laterally due the V-bottom effect. The effects on resistance connected to the spray generation cannot be predicted through the CFD analysis with an appropriate level of confidence. The effects of spray on the takeoff performances will be analyzed by means of flying tests to be conducted on a scale model of the aircraft.

2.3 Conventional and Planing-Tail hulls

In the literature ([7], [8], [9]) a hull configuration, named "Planing-Tail" is studied; this configuration differs from those commonly used (hereafter called "conventional") for the following aspects: the CG is positioned after the step section rather than forward, the afterbody is much longer, the step is deeper while the angle of afterbody keel is lower.

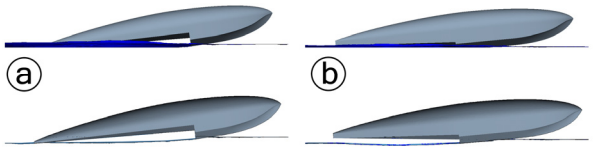


Fig. 5: lateral and section view of planing-tail (a) and conventional (b) configurations

The distinctive feature of planing-tail hull lies in its behaviour during the planing phase: contrary to a conventional configuration, where the contact with water occurs only on the step edge, there is a second contact area at the afterbody tip, as shown in Fig. 5.

Past NACA studies highlight important advantages, as: a lower planing resistance and a wider range of trim angle and speed conditions free from porpoising.

3 IDINTOS configuration

The IDINTOS project has been financed by the Regional Government of Tuscany with the aim to design and manufacture an amphibious ULM PrandtlPlane; the artistic preliminary

configuration depicted in Fig. 1 is just a starting point of the project.

As mentioned above, in a seaplane, in addition to the stability of flight, the stability during takeoff on water must be accomplished; the positions of the CG on water is a fundamental parameter in the optimization procedure of the hull.

3.1 Regulations

Italian Regulations on ULM Amphibians require:

- maximum operative weight of 495 Kg;
- buoyancy guaranteed for a weight equal to 180% of MTOW;
- wing tip floats able to provide a stabilizing moment that is at least 1.5 times the overturning moment due to lateral tilt of the aircraft;
- watertight compartments with similar internal volume;
- plane must not sink with 2 compartments flooded.

3.2 Weights distribution

As shown in [10], a correct flight stability for a PrP configuration can easily be achieved with a CG roughly placed in an intermediate position between the wings.

2 Pilots – Zero Fuel	3.108 m
2 Pilots – Full Fuel	3.113 m
1 Pilot – Zero Fuel – Ballast	3.164 m
1 Pilot – Full Fuel – Ballast	3.163 m
0 Pilots – Zero Fuel	3.318 m
0 Pilots – Full Fuel	3.295 m

Table 1: operating conditions and CG positions

A preliminary estimation of weight and inertia is based on a point-like masses simplified model. The masses considered are: fuselage, wing system (divided into front, rear and bulkhead), two pilots (placed side by side), 60 litres of fuel in main tank, 18 litres of fuel in reserve tank, landing gear, engine, transmission, propeller, board instrumentation, parachute, battery, internals.

The total number of configurations to be analyzed is 10, considering: the number of pilots (0, 1 or 2), fuel volume (full or empty) and operating conditions (floating or flying) .

The CG at full operative weight results at 3.11 m from the bow; in Table 1 all the operating conditions are shown.

3.3 Preliminary design

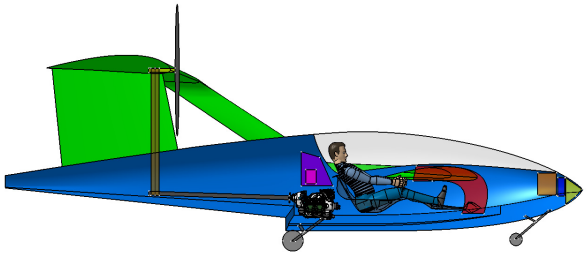


Fig. 6: sketch of lateral view of the aircraft

An artistic lateral sketch of the aircraft which meets all regulations and design requirements is shown in Fig. 6.

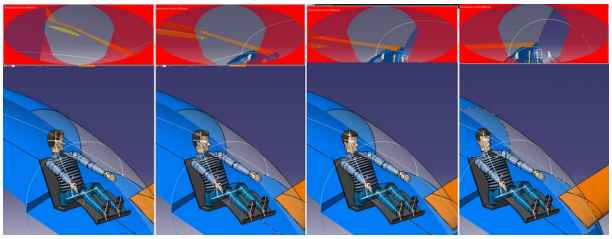


Fig. 7: visibility check of the pilots

Using code CATIA, the following investigations are performed:

- ergonomics of interiors and pilots visibility, to optimize the internal cabin and the external fuselage (Fig. 7);
- hydrostatic stability, which occurs when the CG of the displaced volume of fluid and that of the airplane are aligned along the vertical (Fig. 8);
- lateral stability, in order to design wing tip floats (Fig. 9);
- buoyancy in the case of failure of 2 watertight compartments.

All parts which have to be accessible in floating must stay always above the water level.

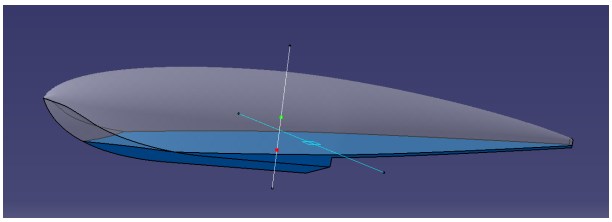


Fig. 8: hydrostatic stability condition

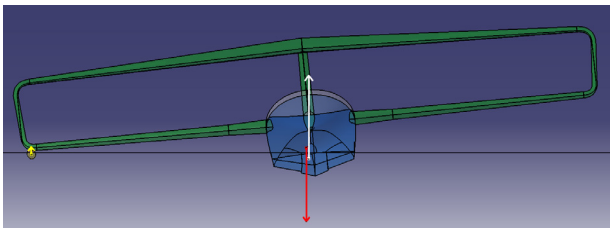


Fig. 9: lateral stability on water

4 Hydrodynamic numerical analysis

The main part of the present paper concerns the numerical simulations of the effects that the principal hull shape geometric parameters produce on takeoff run characteristics, that are: resistance (at the hump and in the planing phase), trim angle, speed and stability.

The results of this analysis will be used to select the hulls which will be more in depth analyzed with flying scale models and also verified into a towing tank.

4.1 Analysis of hull configurations

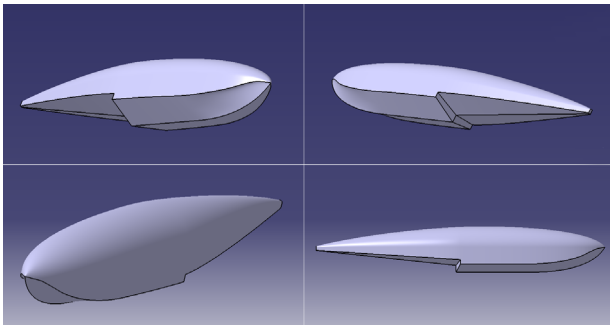


Fig. 10: basic fuselage shape

The basic fuselage has the following dimensions (Fig. 10):

- length: 7 m;
- step depth: 150 mm;
- maximum beam at chines: 1.2 m;
- angle of afterbody keel: 6°;
- step distance from the bow: 3.5 m;
- CG distance from the bow: 3.15 m;
- dead-rise angle at step section: 18°.

As said before, moving toward the bow, the dead-rise angle increases and the V-bottom sides vary from straight to concave.

Starting from this basic configuration, the main geometric characteristics of the hull are varied with the aim of investigating their influence on the takeoff performances. The part above the hull is fixed and it does not represent the actual top fuselage but a simplified one in order to reduce modeling and computational times.

The geometric parameters shared by all the models are:

- maximum beam, resulting from the definition of the internal cabin layout;
- dead-rise angle at step section.

The following parameters are varied:

- afterbody keel angle;
- step depth;
- distance between CG and step section;
- afterbody length;
- longitudinal steps called “spray rails”;
- step planform angle;
- curvature of forebody keel at the bow.

Spray rails are realized in two different configurations (Fig. 11).

In total, 12 hull variants, including 2 regarding planing-tail configurations, are analyzed.

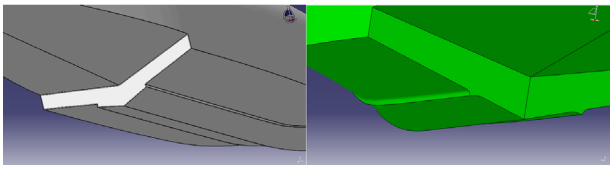


Fig. 11: spray rail analyzed

4.2 Numerical code, models and solvers

Numerical analyses on hulls are carried out with CD-Adapco Star-CCM+ (versions 5 and 6) CFD software.

The integration of the two longitudinal differential equations give the following results: resistance vs time, aerodynamic lift forces and

moments vs time, pitch damping vs time, acceleration, speed, distance from start.

The VOF (Volume Of Fluid) model is used to simulate the multiphase flow; this model, described in details in [11], is recommended for typical free surface problems.

The two phases are treated as an equivalent fluid, assuming common velocity, pressure and temperature fields and solving the same set of governing equations for a single-phase flow. The physical properties of this fluid depend on those of constituent phases and, also, on the values of the so-called volume fraction:

$$\alpha_i = V_i / V \qquad (1)$$

where V_i is the volume of the ith phase and V is the total volume. This model utilizes the HRIC (High Resolution Interface Capturing) convection discretization scheme to follow the evolution of the interface between the two phases.

The multiphase flow is solved with incompressible RANS equations; both air and water densities are assumed as constant. The turbulence model chosen is the *Realizable Two-Layer* κ-ε; the boundary layer is modeled by means of the *Two Layer All y+* wall treatment.

4.3 Body modeling

The hull is modeled as a rigid body with pitch rotation, vertical displacement and horizontal displacement degrees of freedom; these quantities vary versus time in response to pressure and shear actions exerted by the multiphase fluid, and to external forces and moments. The grid is fixed with respect to the body and so the whole computational domain rotates and moves according to the rigid body motion of the same.

To reduce the computational cost, the aircraft is represented only by the hull. The wing system is taken into account by the aerodynamic derivatives, which have been calculated before and, in the present analysis, are assumed as inputs.

Aerodynamic lift L and pitching moment M (halved, as well as all the other external actions, because only half of the hull is simulated, as will be seen afterwards) are inserted through the following expressions:

$$L = 1/4 \cdot \rho_0 \cdot S \cdot V_x^2 \cdot \cdot (C_{L0} + C_{L\alpha} \cdot (\alpha - \tan^{-1}((dV_z/dt)/V_x))) \qquad (2)$$

$$M = -1/4 \cdot \rho_0 \cdot S \cdot V_x^2 \cdot c \cdot \cdot (C_{m0} + C_{m\alpha} \cdot (\alpha - \tan^{-1}((dV_z/dt)/V_x))) \qquad (3)$$

where ρ_0 is the density at sea level, S is the wing system surface, c is the mean aerodynamic chord, V_x and V_z are the horizontal and vertical components (in wind-axis system) of the hull speed, α is the hull trim angle, C_{L0} and C_{m0} are lift and pitching moment coefficients at $\alpha=0$ with high lift devices extended, $C_{L\alpha}$ and $C_{m\alpha}$ are lift and pitching moment α-derivatives, assumed to be independent from any ground effect. V_x, V_z and α are updated at every time step.

Lift is applied in the CG, assuming, for simplicity sake, that it is located in the mean position between the two wings. In principle, the actual positions of the aerodynamic center of pressure could introduce a pitching moment, but, in the framework of this preliminary analysis, it is reasonable to assume that the optimum hull characteristics are independent from this moment. This hypothesis will be verified by means of flying tests to be conducted on scale models.

Another action exerted by the wing system is the moment M_q dependent on the pitching angular velocity q; this is applied on the body as a *Damping Moment*, which requires to specify only the damping constant:

$$M_q / q = 1/4 \cdot \rho_0 \cdot S \cdot V_x^2 \cdot c \cdot C_{mq} \qquad (4)$$

where C_{mq} is the pitching moment q-derivative, whose value is assessed through an analysis based on a panel method.

The other external forces acting on the hull are gravitational force and thrust. The latter is inserted as a function of the hull speed, obtained through the actuator disk theory: this model is very simple, but, at this preliminary stage, it allows to make a quick estimate with an adequate level of accuracy, since only the engine power, the air density and the propeller disk surface shall be specified. In the speed range of interest for the takeoff run (up to 25 m/sec), thrust is approximated as a linear

function of the hull speed. This force is applied in the point of installation of the propeller, in *x* (body-axis system) direction.

The aircraft mass, the CG position, the moments of inertia (based on the weights distribution of the entire aircraft) and the directions of the initial body-axis coordinate system are also specified.

4.4 Spatial discretization

Figure 12 illustrates the control volume (CV) for a hull, where *L* represents the hull length: dimensions are chosen considering past experiences on planing hulls.

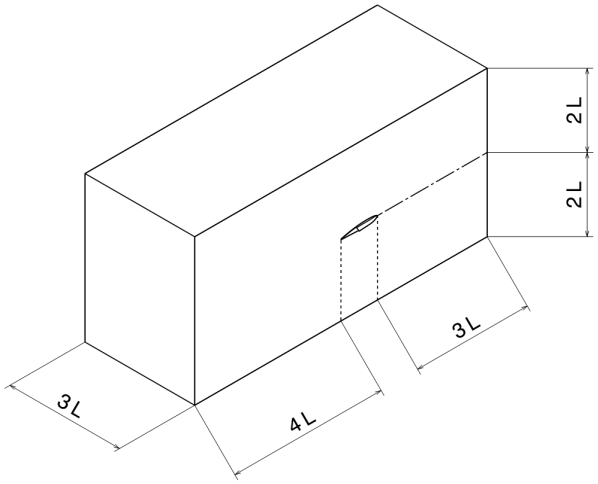


Fig. 12: control volume of the analysis

Since a takeoff run without yaw and roll angles is simulated, the *xz* symmetry plane is utilized and only a half hull is studied.

Unstructured meshes are generated in Star-CCM+; the elements used are hexahedra, whose dimensions are refined in proximity of the hull surface.

Eight layers of prismatic cells are extruded on the hull surface, in order to enhance the alignment with the local flow direction. In addition, 2 volumetric blocks with increasing anisotropic refinements (higher along the normal to the initial water level) are created to improve the grid resolution around the free surface. Figure 13 shows an example of mesh used for simulations.

The resulting number of cells for each grid is about of 280,000; considering the number of cases to simulate, the model is judged to be adequate at this level of analysis, where the interest is basically focused on comparing the performances of the candidate hulls.

Analyses of some configurations have been performed with 1,200,000 and 2,500,000 cells; a mesh refinement should be associated to a proper time step reduction because the local Courant number must be limited to provide simulations stability. This would require great computational resources, which are not justified during a comparative study.

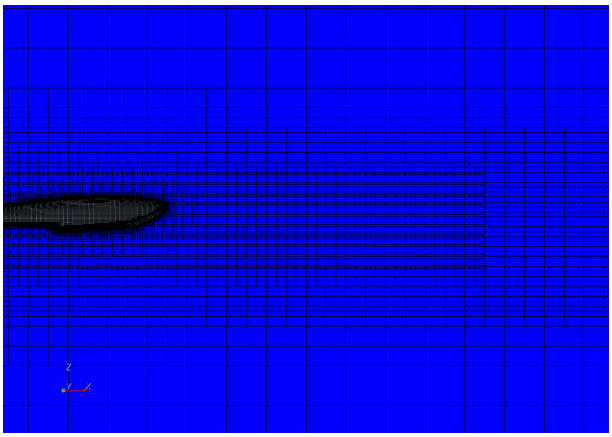


Fig. 13: example of mesh refinement

4.5 Time integration.

Time integration is conducted using an implicit unsteady solver with a first order temporal scheme; the calculations are carried out with time steps of 0.01, 0.005 and 0.001 seconds.

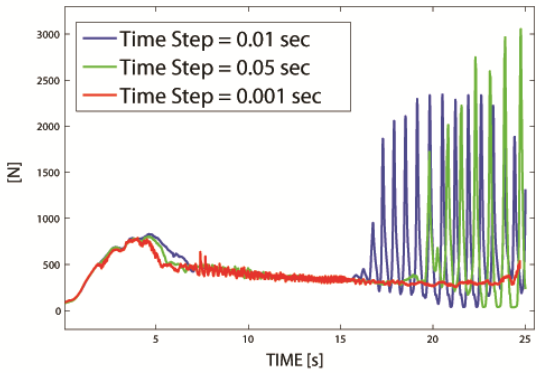


Fig. 14: effects of integration time steps

The 0.01 sec. time step gives unsatisfactory results: some simulations diverge and, in the

other cases, unexpected vertical oscillations occur, meanwhile the hull is planing, the speed is increasing and the trim angle is reducing. These draft oscillations are coupled with quick increase of trim angle and resistance; resistance becomes unstable and assumes very high values. Hull acceleration reduces to zero and the hull oscillates with a speed smaller than that required to takeoff (Fig. 14); then, the aircraft is unable to takeoff.

These effects are not connected to any physical phenomenon, e.g. porpoising, but depend on numerical problems of convergence: in most cases, when reducing the time step to 0.005 sec., the dynamics of the hull remains unaltered, but the draft oscillations described before disappear or are considerably delayed so that the takeoff speed is reached. In some cases the 0.005 sec. time step is still insufficient and, hence, the smaller of 0.001 sec. is adopted; in this case, the oscillatory motion disappears, as Fig. 14 clearly shows. All the solutions in Fig. 14 show the following characteristics: the initial behavior before the hump is the same; some small differences could be present after the hump and, finally, the planing phase is the same, apart from the possible final numerical oscillations, discussed before.

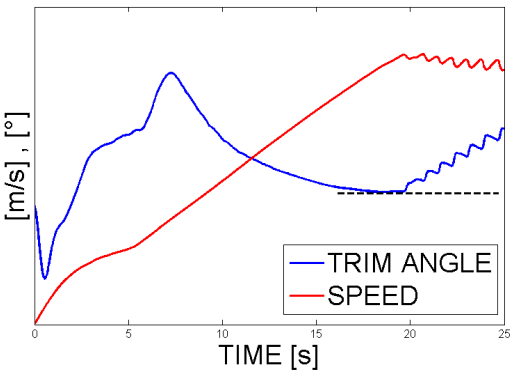


Fig. 15: oscillations in trim angle and hull speed

The variations of the angle of attach and speed vs time are deployed in Fig. 15. These results are the solutions of the two longitudinal equations of motion; the aerodynamic characteristics are taken into account as said before. In particular, the aerodynamic damping C_{mq} of the present PrandtlPlane is much higher than a conventional aircraft. When this damping is not introduced, the oscillatory motion happens very soon; when it is introduced, motion becomes stable but, as shown in Fig. 15, the stability of motion of the aircraft becomes critical when $d\alpha/dt = q$ tends to zero. In other words, the oscillations are numerical but they occur when all the damping effects on motion vanish.

4.6 CFD results

CFD simulations confirm that planing-tail hulls present a lower resistance than a conventional one during takeoff; the drawback of this configuration, however, is the pilots visibility, because the trim angle remains higher than a conventional one throughout the complete run.

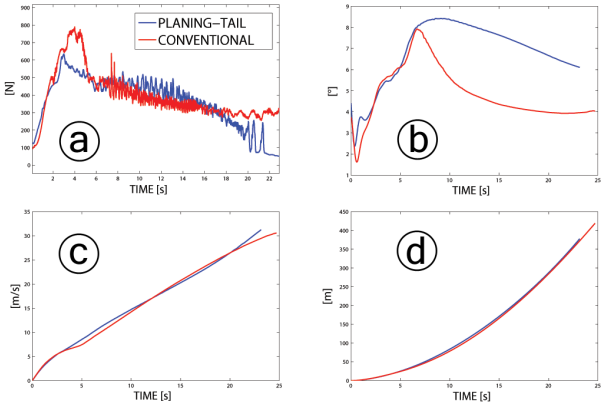


Figure 16: comparison between planing-tail and conventional hulls

Figure 16 shows a comparison between a planingtail and a conventional solution; the planingtail presents a lower resistance hump in the first phase (Fig. 16a), a trim angle with a higher peak and also higher values along the complete run (Fig. 16b) with a consequent limitation of visibility during the complete takeoff maneuver. The aircraft speed vs time is deployed in Fig. 16c; it is quite evident that the speed vs time is not very different and the takeoff distance of the two aircraft is quite the same (Fig. 16d). A conventional hull is then adopted for the present ultralight amphibious aircraft because an excellent visibility for almost all maneuvering conditions is guaranteed without significant performance penalty. The

planing-tail hull could be adopted in the case of bigger aircraft provided with sophisticated control instrumentations.

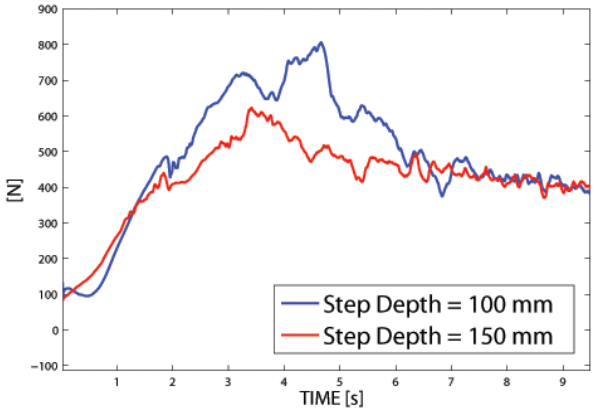


Fig. 17: effect of step depth on resistance

The effect of step depth is described in Fig. 17. A deeper step results into a reduction of the resistance hump, which is also anticipated; this positive effect is the consequence of the onset of separation between afterbody and water, (anticipated by a deeper step). On the other side, a deeper step reduces the aerodynamic efficiency during flight; thus, we need to achieve a compromise, by adopting the minimum step depth which permits a safe and acceptable takeoff run.

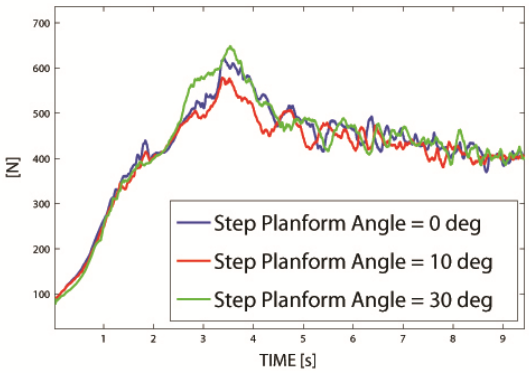


Fig. 18: effect of step planform angle on resistance

The effect of step planform angle is shown in Fig. 18. An angle of 10° represents a good compromise between easy manufacturing and good performances.

The effect of spray rails is depicted in Fig. 19. The use of spray-rails leads to lower trim angles during the planing phase (Fig. 19a) with a better visibility during the takeoff maneuver and, also, a higher acceleration achieved (Fig. 19b).

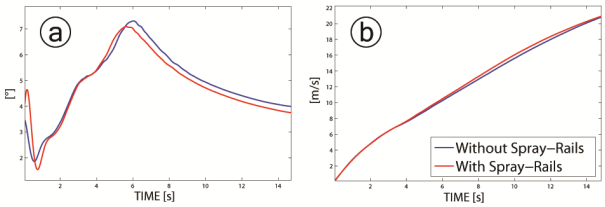


Fig. 19: effect of spray rails on trim angle (a) and hull speed (b)

The effect of longitudinal distance between step and CG can be observed in Fig. 20; an intermediate position of the CG produces the best performances and the higher stability in the takeoff run.

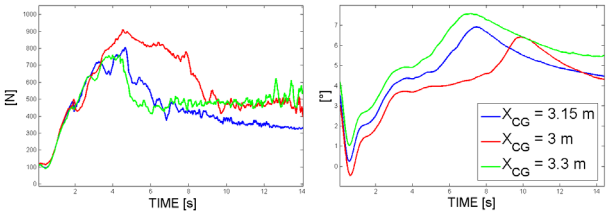


Fig. 20: effect of CG-step distance on resistance (a) and trim angle (b); X_{STEP} = 3.5 m from the bow

The effects of a shorter afterbody are a faster transition between the displacement and the planing phases but also an increase of the trim angle peak (Fig. 21).

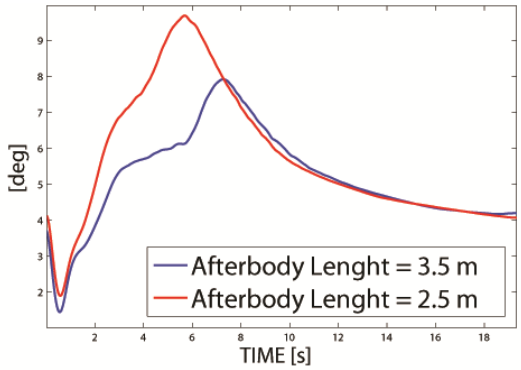


Fig. 21: effect of afterbody length on trim angle

5 Conclusions

A CFD analysis have been performed on 12 hull configurations of a ULM amphibious

PrandtlPlane; results show that a configuration exists with a high hydrodynamic efficiency coupled with a trim angle history that permits an excellent visibility to the pilots.

References

[1] A. Frediani, L. Balis Crema, G. Chioccia, G. L. & L. M. Ghiringhelli, "Development of an innovative configuration for transport aircraft", XVII AIDAA Congress, 2003, pp. 2089-2104.

[2] A. Frediani, "The Prandtl wing", Lecture series on Innovative Configuration and Advanced Concepts for Future Civil Aircraft, 2005.

[3] A. Frediani and G. Montanari, "Best Wing Sytem: An Exact Solution of the Prandtl's Problem", Variational Analysis and Aerospace Engineering, 2009, pp. 181-211.

[4] L. Prandtl, "Induced drag of multiplanes", NACA n°182, 1924.

[5] A.Frediani, E. Rizzo, V. Cipolla, L. Chiavacci, C. Bottoni, J. Scanu and G. Iezzi, "Optimization methods for ULM PrandtlPlane design and preliminary flight tests on scaled models", CEAS 2008, London, 2008.

[6] D. Savitsky, "Hydrodynamic Design of Planing Hulls", Marine Technology, Vol. 1, No. 1, 1964, pp. 71-95.

[7] J. R. Dawson, R. C. Walter and E. S. Hat, "Tank test to determine the effect on planing-tail hulls of varying length, width and plan-form taper of afterbody", Langley Memorial Aeronautical Laboratory, 1946.

[8] C. C. Yates and J.M. Riebe, "Aerodynamic characteristics of three planing-tail flying-boat hulls", Langley Memorial Aeronautical Laboratory, 1948.

[9] R. McKann and H. B. Suydam, "Hydrodynanic characteristics of an aerodynamically refined planing-tail hull", Langley Memorial Aeronautical Laboratory, 1948.

[10] V. Cipolla, S. Giuffrida, "Utilizzo di codici a pannelli nel progetto preliminare di velivoli PrandtlPlane ultraleggeri; applicazione a nuove configurazioni", M.Sc. Thesis, Department of Aerospace Engineering, University of Pisa, Italy, 2006.

[11] STAR-CCM+ 5.04.006, "User Guide", CD-adapco.

Synthesis and Topology Study of HLFC System Architectures in Preliminary Aircraft Design

T. Pe and F. Thielecke
Hamburg University of Technology,
Institute of Aircraft Systems Engineering, Germany

Keywords: *Hybrid Laminar Flow Control, HLFC, Suction System Architectures, Sizing Methodology, Topology Study, Preliminary Aircraft Design*

Abstract

Hybrid Laminar Flow Control (HLFC) is a very promising technology with respect to high fuel savings for future aircraft. Great advances have been made regarding the ability to specify the aerodynamic requirements. However, for the design of an integrated HLFC aircraft, also structure and system technologies have to be understood and optimised.

This paper presents a detailed system sizing methodology with the objective to determine main system parameters such as power consumption and mass in an early design phase.

This methodology is then applied for a suction system architecture study for a long range research aircraft. In this study different architectures with varying numbers of compressors and plenum chambers are analysed. The results are recommendations about number and locations of the compressors and the main parameters of the systems that can be used in preliminary design at overall aircraft level.

1 Introduction

The increase of the lift-to-drag ratio (L/D) is an important objective towards the design of future eco-efficient aircrafts. Regarding the drag breakdown of a typical twin-jet in cruise flight, the friction drag contributes with nearly 50% to the overall drag [1], see Fig. 1. Therefore, its reduction is an important topic for scientists and aircraft industry.

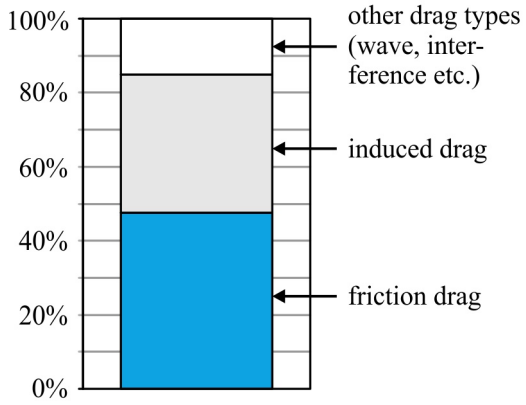


Fig. 1 Drag breakdown of a typical twin-jet in cruise flight [1]

On conventional swept wings of transport aircraft the boundary layer (BL) transition occurs very close to the leading edge [2], due to three distinct mechanisms: Tollmien-Schlichting instability, cross flow instability and attachment line transition [3]. Since the friction drag of the turbulent BL is up to ten times higher than for the laminar BL [1], it is desirable to maintain laminarity as long as possible.

1.1 HLFC Principal

The BL can be stabilised with the following techniques: acceleration of the flow through a negative pressure gradient or suction of the BL [1]. The first technique is called *Natural Laminar Flow* (NLF) if it is used solely. It is limited to smaller sweep angles and Reynolds numbers [4] and therefore not applicable for today's transport aircraft. The second is called *Laminar Flow Control* (LFC). Laminarity can be maintained for higher sweep angles and Reynolds number, but it has the disadvantage of a high energy effort and additional and large systems.

Hybrid Laminar Flow Control (HLFC) combines suction near the leading edge with a negative pressure gradient, see Fig. 2, and overcomes the disadvantages of NLF and LFC.

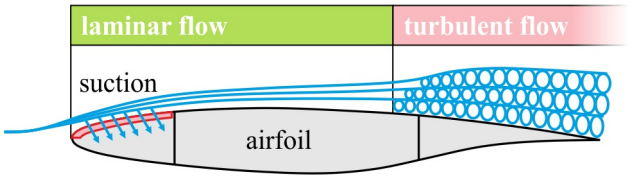


Fig. 2 Laminar-turbulent boundary layer transition of an HLFC airfoil

Applying HLFC on wings, tailplanes and nacelles, an overall drag reduction of approximately 15% is expected in [5]. However, the range of predicted overall drag reduction is large in publications. It depends on the assumptions made for the laminar extent and applied surfaces, but usually lies in a range between 10% and 20%.

The design of an integrated HLFC aircraft includes the optimisation of the aircraft configuration, airfoils, suction system, leading edge high-lift device, suction structure and surface. Compared to conventional aircraft, the suction system is a completely new system that adds extra mass and power consumption and has a negative contribution to the net benefit (fuel reduction) of the HLFC technology.

An important degree of freedom for the system design is the topology of the suction system (including number and location of compressors). This aspect is investigated in this paper.

1.2 Suction System and Structure

In the ALTTA programme (*Application of Hybrid Laminar Flow Technology on Transport Aircraft*) a concept was developed that simplifies HLFC structures and systems [5]. The ALTTA double structure in Fig. 3 consists of a porous outer skin and an inner skin with metering holes. Stringers in between the skins form closed suction channels. The size of the metering holes and the pressure in the plenum chamber below the double structure defines the suction distribution over the surface. The complexity of ducting and valves can be reduced compared to former HLFC experiments, where the adjustment of the suction distribution was realised by ducts and valves connected to a certain number of suction chambers.

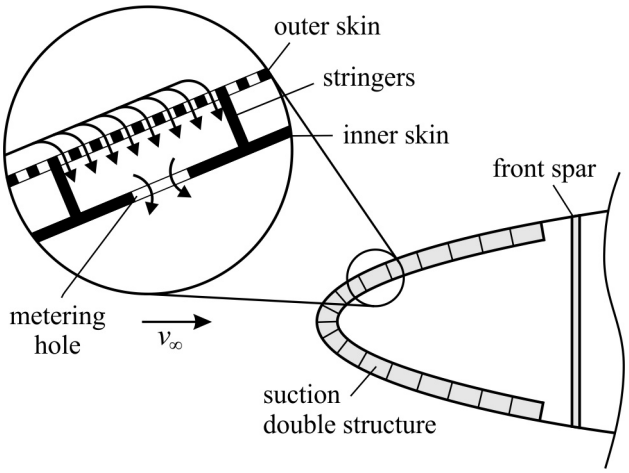


Fig. 3 Principal of suction double structure, developed in the ALTTA programme [5]

The schematic layout of a simple suction system is shown in Fig. 4. The main component is the compressor which has the task to generate an underpressure in the plenum chamber in order to realise a mass flow through the double structure.

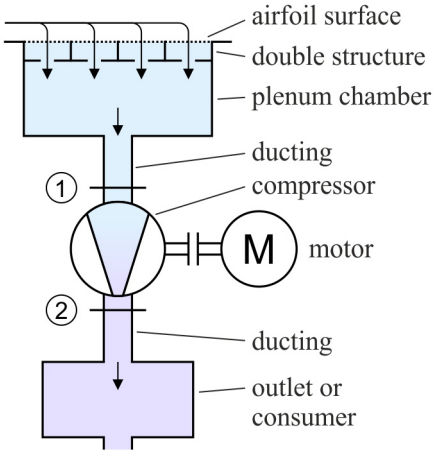


Fig. 4 Schematic layout of a simple suction system

The requirements at the outlet depend on the type of usage of air, e. g. blowing out or usage in the environmental control system (ECS). Therefore, the compressor has to generate a desired pressure and flow speed at the system outlet. Also pressure losses in the ducting down- and upstream of the compressor have to be overcome.

The compressor's shaft power is provided by a motor. Regarding the trend of the *More Electric Aircraft* (MEA), the use of an electric motor is probable, but in general other motor types can be used.

2 Sizing Methodology

An important step is the calculation or estimation of component and system parameters for a given architecture. Fig. 5 shows an overview of the sizing methodology.

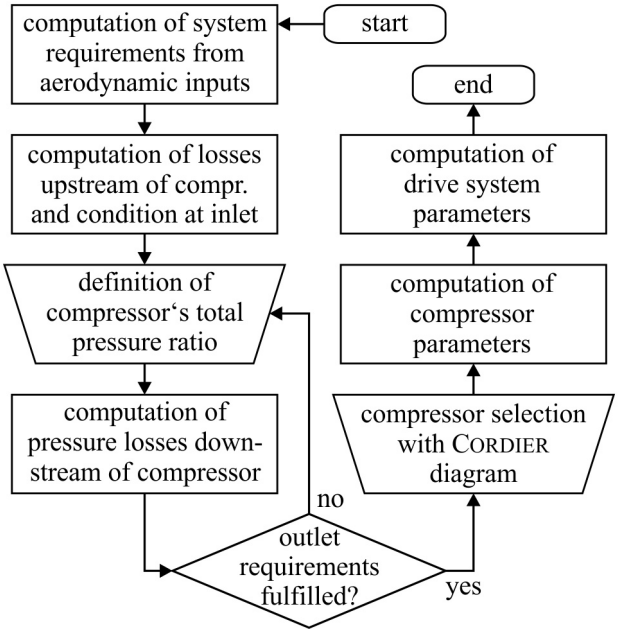


Fig. 5 Flow chart of sizing methodology for HLFC suction systems

The objective of this methodology is to estimate power consumption, mass and size of the components during aircraft preliminary design phase. It is not necessary to design each component in detail, if accurate information about these parameters can be generated by other means.

The main step is the preliminary design of the compressor, which is done using the CORDIER diagram. It allows the determination of the best (dynamic) compressor type and important dimensionless parameters, such that the efficiency is maximised. Besides the requirements on the suction surface and of the system outlet, the compressors are sized through internal pressure losses in the ducting network and plenum chambers. If the compressor parameters are known, the drive system can be designed.

Wing and airfoil geometry, the pressure coefficient C_p [6] and the suction coefficient C_q [7] are needed as input data, where:

$$C_p = \frac{p_s - p_\infty}{\frac{1}{2}\,\rho_\infty\, v_\infty^2}\ ,\quad C_q = \frac{w_s}{v_\infty}\ . \tag{1}$$

The following sections give an overview of the formulas used in the sizing methodology. A more detailed description can be found in [8].

2.1 Losses in Double Structure

The air flow in the double structure passes the porous outer skin, enters into the chamber and then passes the metering hole, as illustrated in Fig. 6.

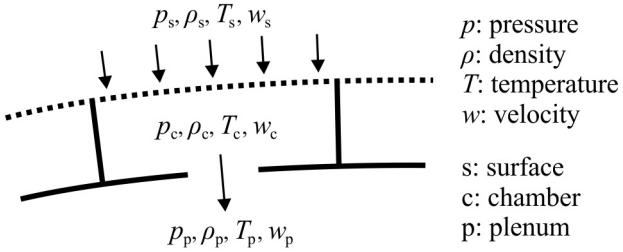


Fig. 6 Flow conditions in the double structure

Pressure and velocity on the surface can be calculated with Eq. (1). The temperature on the surface is determined through the boundary layer temperature [9]:

$$T_s = T_\infty + r \cdot T_\infty \cdot \frac{\kappa - 1}{2} \cdot Ma_\infty^2\ , \tag{2}$$

where r is the recovery factor and κ the isentropic exponent. The density can be estimated with the ideal gas law. The flow condition on the surface is then fully known.

The micro-sized holes of the outer skin are usually produced with laser or electron beam drilling that produces approximately conical forms [10]. The pressure losses can be modelled by:

$$\Delta p_{sc} = A \cdot \mu_s \cdot w_s + B \cdot \rho_s \cdot w_s^2 \ , \qquad (3)$$

where A and B reflect the loss characteristics of the hole geometry and drilling matrix [10].

In early design phases often the problem arises that the geometry of the double structure chambers and the metering hole diameters are not yet defined. Regarding the fact that the plenum chamber pressure p_p has to be smaller than the smallest pressure in the double structure p_c, one can use an estimation:

$$p_p = \min(p_s - \Delta p_{sc}) - \Delta p_{cp} \ , \qquad (4)$$

using a reasonable value for Δp_{cp} .

If the geometry is known, a mean pressure in the chamber p_c and the losses through the metering holes have to be calculated.

The mass flow of a suction area A is calculated by the following surface integral:

$$\dot{m}_A = \iint_A \rho_s \, w_s \, \mathrm{d}s \, \mathrm{d}y \ . \qquad (5)$$

This implies the knowledge of the airfoil geometry and the suction and pressure coefficient at each point of the suction area. Often only few 2D profiles are given, which require an interpolation.

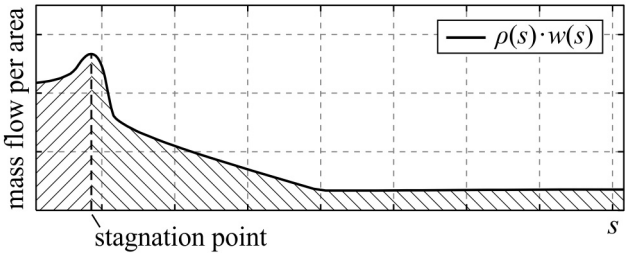


Fig. 7 Typical 2D suction mass flow profile

A typical mass flow profile is given in Fig. 7. Suction begins below the stagnation point in order to account for shifting of the stagnation point in off-design cases.

2.2 Losses in Plenum Chamber

After passing the double structure, the air flows in plenum chambers through the leading edge (LE). The plenum chamber could be a separated closed structure or the LE box itself, if it can be achieved to seal it, in particular with respect to moving high lift systems. A continuous mass flow is entering into the chamber in spanwise direction, which causes the overall mass flow in the chamber to increase in flow direction. Therefore, it is reasonable to expand the chamber cross section from outboard to inboard, which corresponds to expanding LE box geometry. To withstand aerodynamic loads, the surface has to be supported by ribs that the air flow has to pass through.

A pressure loss model was developed in order to account for pressure losses in the plenum chamber and ribs. The principal is shown in Fig. 8. The chamber is divided into n parts, where n is the number of ribs that the flow is passing through.

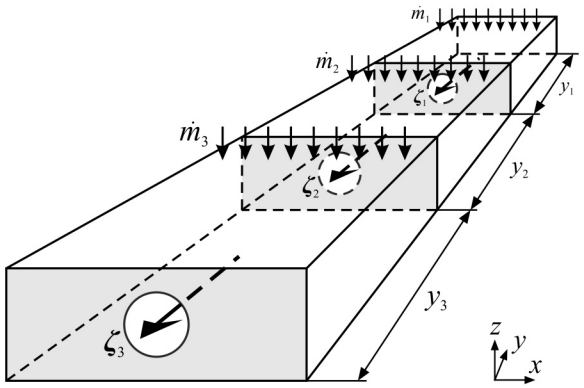


Fig. 8 Pressure loss model of plenum chamber

Since the divergence angle of the chamber is very small between two ribs, the losses are modelled using a formula for a rectangular tube [13]. The whole mass flow of one part is assumed to enter directly after a rib without flow mixing losses.

Pressure losses at ribs are calculated using a pressure loss coefficient ζ [13]:

$$\Delta p_{\mathrm{rib}} = \zeta \cdot \rho \cdot \frac{w^2}{2} \ . \qquad (6)$$

ζ depends on the geometry of the ribs and can only be roughly estimated in this design phase. It is assumed that the rib is designed as an open truss. In [13] formulas and diagrams are given to calculate the pressure loss through "barriers uniformly distributed over a duct cross section". Although this is not exactly the case, it can be used for a first estimation of the rib losses, but experiments or flow simulations should verify the real losses, when exact rib geometry is known.

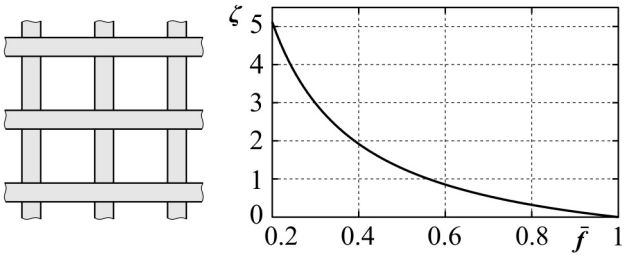


Fig. 9 Pressure loss coefficient through two-plane screen [13]

For this work the loss coefficient for a flow through a two-plan screen is used, see Fig. 9. The coefficient ζ is calculated by:

$$\zeta = 1.28 \cdot \frac{1-\bar{f}}{\bar{f}} \ , \qquad \bar{f} = \frac{F_0}{F_1} \ , \tag{7}$$

where F_0 is the free area and F_1 is the whole area. A ratio $\bar{f}$ around 0.9 is assumed in this work, which means that 10% of the cross section is occupied by rib beams.

2.3 System Outlet Condition and Losses in Ducting

The compressors have to be designed such that the required system outlet condition is fulfilled. A simple reverse-computing from system outlet to compressor outlet is not possible. This can be seen in the formula for the pressure loss of a compressible fluid in a pipe [11]:

$$\Delta p_{\text{pipe}} = p_{\text{in}} - p_{\text{out}} \ ,$$

$$p_{\text{out}} = p_{\text{in}} \sqrt{1 - \lambda \frac{L}{d} \frac{\rho_{\text{in}}}{2} \bar{w}_{\text{in}}^2 \frac{2}{p_{\text{in}}} \frac{T_{\text{in}} + T_{\text{out}}}{2\, T_{\text{in}}}} \ . \tag{8}$$

For a given flow condition at the pipe outlet, it is not possible to determine directly the flow condition at the inlet.

The solution is an iterative process, where the condition at the inlet is varied until the required outlet condition is reached. For the here presented methodology this means that the compressor outlet condition has to be varied during the iteration. The main influence parameter is the pressure ratio p_2/p_1 of the compressor, whereby the total pressure at the outlet can be changed. The diameter of the ducting influences the flow speed (and the pipe losses).

The procedure includes the following steps: definition of the ducting geometry, definition of the compressor pressure ratio, calculation of the losses in the ducting, comparison between the actual and the required outlet condition. If the difference is too large, the procedure is repeated using adapted parameters.

Depending on the accuracy of the duct routing definition, several components like bends, diffuser, junctions and valves can be considered. The pressure losses of these components can be estimated using empirical data which is available for example in [12] and [13]. If a detailed routing is not possible, due to very early design phases, mark-up factors can be applied to estimate such components.

Special attention has to be given to combining junctions of a collective duct, see Fig. 10. The outlet of a junction is influenced by both flow conditions of the inlet ducts. A required outlet condition can therefore be reached with various combinations of inlet conditions.

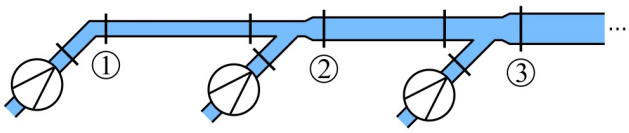


Fig. 10 Several compressors feeding into a collective duct via combining junctions

With the requirement that the flow conditions at both inlet ducts of a junction should be equal, the mixing losses become small and the calculations are simplified. An inside-out procedure can be conducted, beginning with the compressor nearest to the system outlet. The subsequent compressor has to be designed such that the already given flow condition at the junction inlet is fulfilled.

It has to be noted, that the effort to estimate pressure losses for complex ducting is very high. If possible, the ducting should be kept simple and mark-up factors should be applied in preliminary aircraft design. For complex ducting an internal flow simulation software should be prioritised.

2.4 Compressor Selection and Parameters

The required flow conditions at the inlet and outlet of the compressors are now known, so the compressor design process can begin. The suction system application requires a high constant flow with moderate pressure ratios. Size, mass

and power consumption of the compressors should be minimised. The dynamic compressors are suitable for this application; especially the radial, diagonal (or mixed flow) and axial compressors, due to their high efficiencies.

The objective is to find a high efficient compressor. The parameters of such compressor depend on the requirements for volumetric flow $\dot{V}$, pressure ratio p_2/p_1, rotational speed n and blade diameter d. The preliminary design is conducted using the CORDIER diagram [14], which is based on the dimensionless parameters specific speed σ and specific diameter δ:

$$\sigma = n \cdot \frac{\sqrt{\dot{V}}}{(2\,Y)^{\frac{3}{4}}} \cdot 2\sqrt{\pi}\ , \tag{9}$$

$$\delta = d \cdot \left(\frac{2\,Y}{\dot{V}^2}\right)^{\frac{1}{4}} \cdot \frac{\sqrt{\pi}}{2}\ , \tag{10}$$

where Y is the specific energy transfer [15].

Assuming an isentropic compression, Y can be computed with [15]:

$$Y_{\text{is}} = \Delta h_{\text{is}} + \frac{w_2^2 - w_1^2}{2}\ , \tag{11}$$

$$\Delta h_{\text{is}} = \frac{\kappa}{\kappa - 1}\, R\, T_1 \left(\left(\frac{p_2}{p_1}\right)^{\frac{\kappa-1}{\kappa}} - 1\right)\ . \tag{12}$$

CORDIER calculated for a high number of very efficient turbomachines their σ and δ values and the efficiency curves and displayed them in a (σ, δ)-diagram. Fig. 11 shows the original diagram from [14], the so called CORDIER diagram. It can be seen, that the points are located in a narrow band around the CORDIER line. The diagram can be used for the compressor design: when either σ or δ is fixed, the other value can be determined, such that compressor efficiency is maximised.

Different impeller types (radial, diagonal, axial) are optimal for specific ranges of σ. According to [15] these ranges are:

- Radial: $\sigma = 0.06 \ldots 0.32$,
- Diagonal: $\sigma = 0.25 \ldots 1.0$,
- Axial: > 0.8 .

The efficiency is of course not only influenced by the specific speed and diameter, but by a number of other variables. Therefore, the efficiencies in the diagram present rather upper limits that are possible when choosing the optimal σ and δ values.

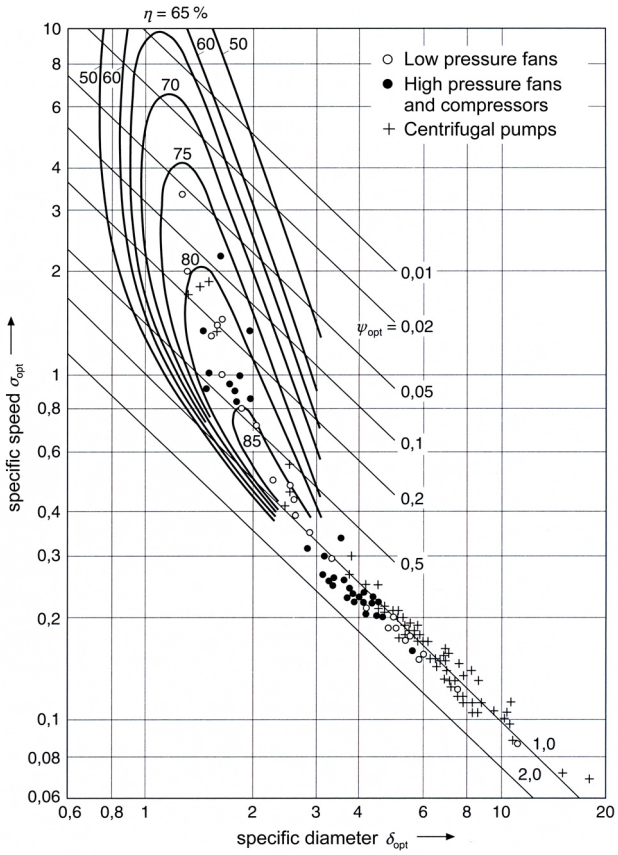


Fig. 11 CORDIER diagram for turbomachines [14]

The shaft power P_{cmp} is calculated by [15]:

$$P_{\text{cmp}} = \frac{\dot{m} \cdot Y}{\eta_{\text{cmp}}}\ . \tag{13}$$

2.5 Drive System Parameters

For this project an electric motor, driven by a frequency converter to control the speed, is used. A detailed design is not necessary for these components, since there is reliable data available to estimate their efficiencies.

The efficiency of an induction motor is estimated using the standard IEC 60034-30 [16], which defines different efficiency classes, see Fig. 12. Efficiencies are defined on the rated output power. In the European Community the IE2 class shall be applied to all motors sold or traded since June 2011. The IE3 class will be mandatory from 2015 on [17].

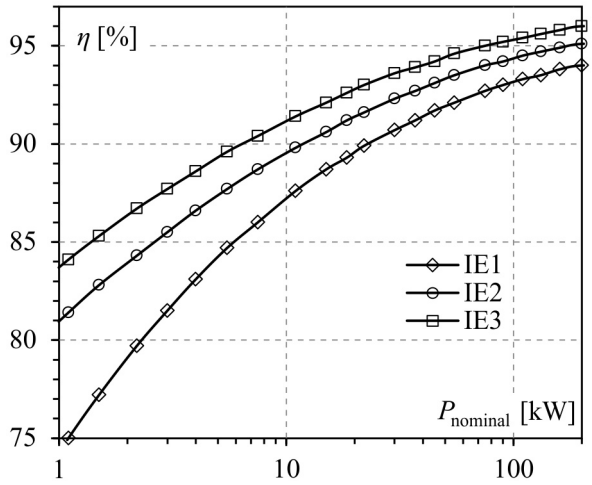


Fig. 12 IEC 60034-30 efficiency classes of 4-pole, 3-phase cage-induction motors [16]

Eq. (13) provides the required power of the compressor in nominal operation mode. With this value, a minimal efficiency η_{mot} of an IE2 or IE3 induction motor can be determined with Fig. 12.

Based on different manufacturer data, the efficiency η_{cnv} of frequency converters is in general 95% or higher. The electric power consumption for the nominal operation can then be estimated by the following formula:

$$P_{\mathrm{el}} = \frac{P_{\mathrm{cmp}}}{\eta_{\mathrm{mot}} \cdot \eta_{\mathrm{cnv}}} \ . \tag{14}$$

2.6 System Mass Estimation

The focus of the preceding sections was on the determination of the system power consumption. Another important parameter for the aircraft design is the system mass, which has to be determined for competing architectures. A direct physical calculation based on volume and density is only in rare cases possible. For the other cases two methods can be applied: 1) a regression analysis based on existing mass data or 2) a physical-technical model from KOEPPEN [18]. The former has the disadvantage that the model is only valid for similar components, which decreases the validity range. The latter can be very accurate, but could require very detailed and complex physical formulas.

The methodology from KOEPPEN [18] consists of three steps, see Fig. 13. The first step is the identification of the parameters importances with respect to the mass of the component or subsystem. Only those with a large influence on the mass will be used. The second step is the forming of physical or technical motivated functionals for the mass computation. The last step is the adaption of this functionals to existing data of built aircraft by scaling factors.

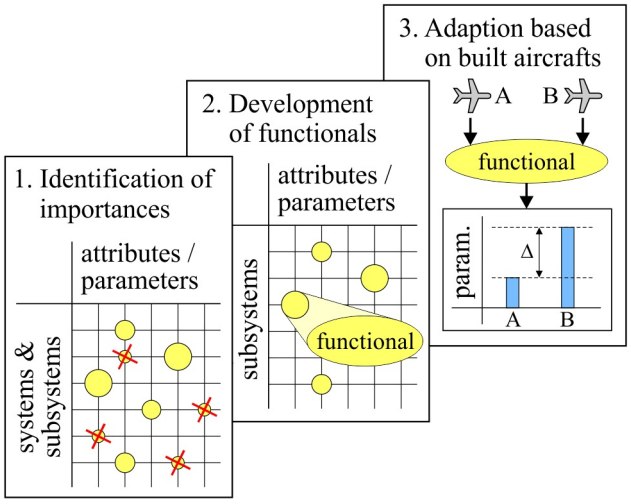


Fig. 13 Physical-technical mass estimation methodology from KOEPPEN [18]

For the estimation of the ducting and wiring network mass the methodology from KOEPPEN is used. Compressor, induction motor and frequency converter masses are estimated with linear regression models.

For the determination of the ducting mass, it is necessary to decide on the material and wall thickness. Lengths and diameters are given from the architecture sizing methodology. The mass of a single duct is calculated by:

$$\begin{aligned} m_{\mathrm{d}} &= A_{\mathrm{d}} \cdot \ l_{\mathrm{d}} \cdot \rho_{\mathrm{d}} \\ &= \frac{\pi}{4}(d_{\mathrm{o}}^2 - d_{\mathrm{i}}^2) \cdot l_{\mathrm{d}} \cdot \rho_{\mathrm{d}} \ . \end{aligned} \tag{15}$$

This mass reflects the pure duct mass. To account for brackets, couplings etc. an adaption factor is applied.

The mass of a single wire can be obtained in a similar way by:

$$m_{\mathrm{w}} = A_{\mathrm{w}} \cdot \ l_{\mathrm{w}} \cdot \rho_{\mathrm{w}} \ . \tag{16}$$

The cross section has to be sized according to its load, voltage and other parameters. According to [19] the cross section of the wire is optimal, if:

$$A_{\mathrm{w}} = \frac{\rho_{\mathrm{w,el}} \cdot P}{P_{\mathrm{w,sp}} \cdot U^2} \ . \qquad (17)$$

Again, an adaption factor that reflects the difference to built aircraft is applied.

Masses of the electric motors and inverters are estimated using linear regression models:

$$m_{\mathrm{mot}} = a \cdot P_{\mathrm{mot,nominal}} + b \ , \qquad (18)$$

$$m_{\mathrm{cnv}} = c \cdot P_{\mathrm{cnv,nominal}} \ . \qquad (19)$$

The regression coefficients a, b and c were derived from internal project data in the context of electrical aircraft system investigations.

The estimation of the compressor mass is a very difficult task, because it is influenced not only by the compressor type (axial, diagonal or radial), but also significantly by several parameters such as the blade diameter, the rotational speed or the number of stages. A model that takes these parameters into account is not available at the moment.

The data of known compressors, built or computed for HLFC projects, showed that a relation of the compressor mass to the casing volume V and the (design point) mass flow $\dot{m}$ can be used for a linear regression model:

$$m_{\mathrm{cmp}} = d \cdot V_{\mathrm{casing}} + e \cdot \dot{m} + f \ , \qquad (20)$$

where d, e and f denote the regression coefficients. Because only few supporting points are available, this model contains some uncertainty, especially because the compressor speed range of the here investigated compressors is much larger than for the given compressors. Nevertheless, the model allows showing mass tendencies for the compressors.

3 Reference Aircraft and Requirements

The investigated aircraft in Fig. 14 is a large long-haul twin-jet with 80 m span and 85 m length and has a design range of 8150 nm.

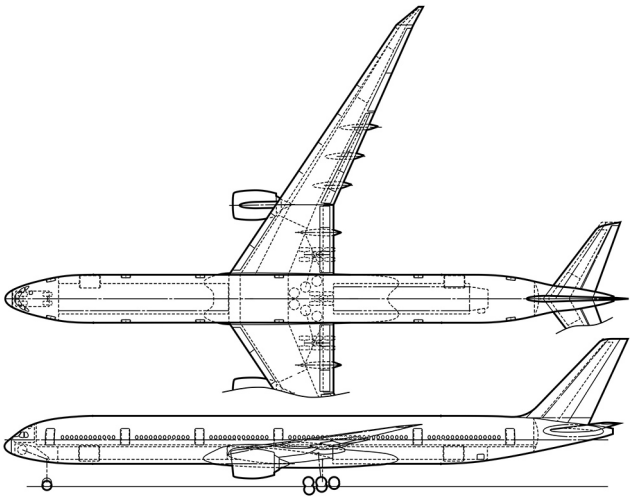


Fig. 14 HLFC research baseline aircraft (turbulent reference) [20]

It is designed as a conventional (turbulent) aircraft and serves as the reference aircraft. For the investigation of the HLFC technology, various characteristics are changed, such as the wing, tailplanes, (suction) systems and also single top level aircraft requirements (TLAR).

In a first step, the HLFC technology is investigated for a flight condition with cruise speed of Ma = 0.8 at 35000 ft altitude.

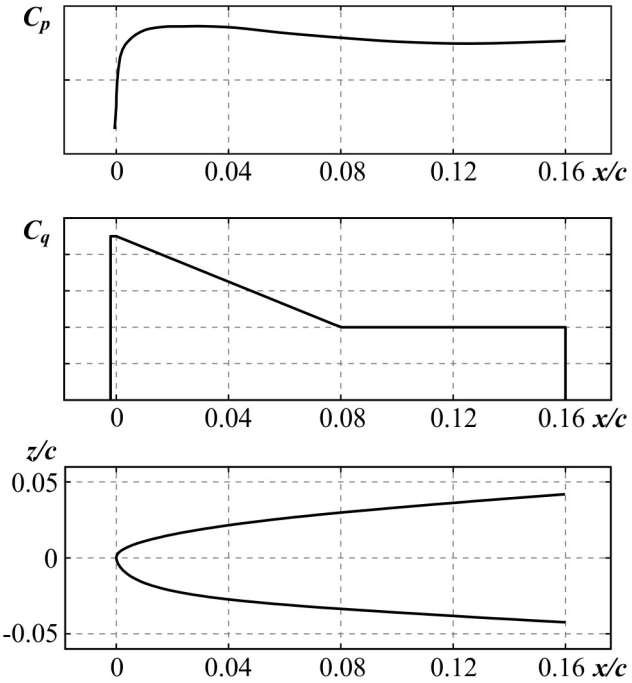


Fig. 15 Aerodynamic inputs for 2D section cut at η = 0.375 [22]

The aerodynamic inputs are given for 2D profiles as shown in Fig. 15. This data is converted to mass flow profiles and plenum cham-

ber pressures and interpolated over the wing planform.

Detailed data is only available for the wings. For the tailplanes the data is estimated using information from the former HLFC project ALTTA [21]: mass flow per m^2 suction area and plenum chamber pressures. For the definition of the available space, scaled profiles from actual aircraft are used.

4 System Architecture Studies

The objective of this work is the investigation of different suction system topologies, basically in terms of number compressors. The set of aerodynamic requirements and the aircraft configuration stay fixed.

4.1 Requirements

In the investigated configuration in Fig. 16, suction is applied to the wings (upper side only) and the tailplanes (both sides). The air is blown out through outflow valves, whose positions are shown in Fig. 16. The frequency converters are located in the E/E bay (electrical and electronics bay) near the landing gear bay. This configuration is fixed for this work, but of course different configurations are possible.

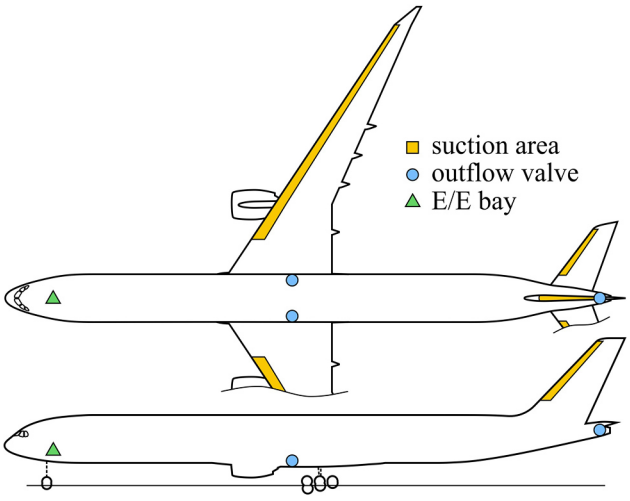


Fig. 16 Suction areas and position of outflow valves and E/E bay

As wing high lift device a Krüger flap with additional shielding functionality is assumed. The flap and its kinematics reduce the available space in the leading edge box. A space allocation model based on Fig. 17 is used, where the upper part of the leading edge box, up to the front spar, is available for the system and plenum chamber. Actuators and kinematics are located at dedicated spanwise stations. So the minimum available space for long spanwise extended elements (such as ducts) is reduced by their occupied space included their travel paths, whereas short elements can use the full space, when they are displaced in spanwise direction.

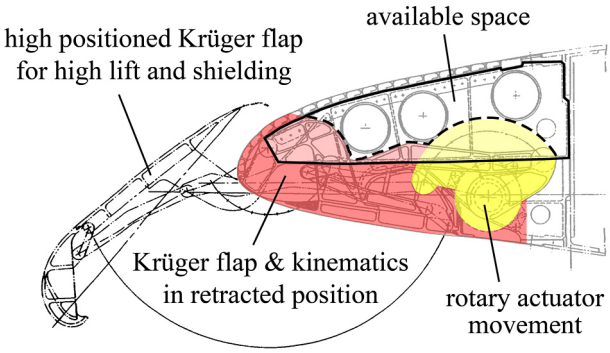


Fig. 17 High positioned Krüger flap for HLFC leading edge [23]

This space allocation model was geometrically simplified and scaled to the used airfoils.

The tailplanes do not employ high lift systems, so it is assumed that the whole leading edge box is available for system installations.

4.2 Topology Study

For the wings the following configurations are investigated: 1, 2, 4, 6 and 8 compressors. For the tailplanes 1, 3 and 6 compressors are analysed, see Fig. 18. Previous computations showed that a connection between wing and tailplane suction system is not reasonable due to very high mass of the connection ducts. Therefore, both systems are investigated as separated systems and can be combined arbitrarily.

In this study each compressor is connected via a separate duct to the nearest outflow valve. In terms of mass, collective ducts for several compressors are advantageous, but failures could lead to the disconnection of several compressors. Future studies should determine the required reliability of the suction system and conduct an architecture study taking reliability analysis into account.

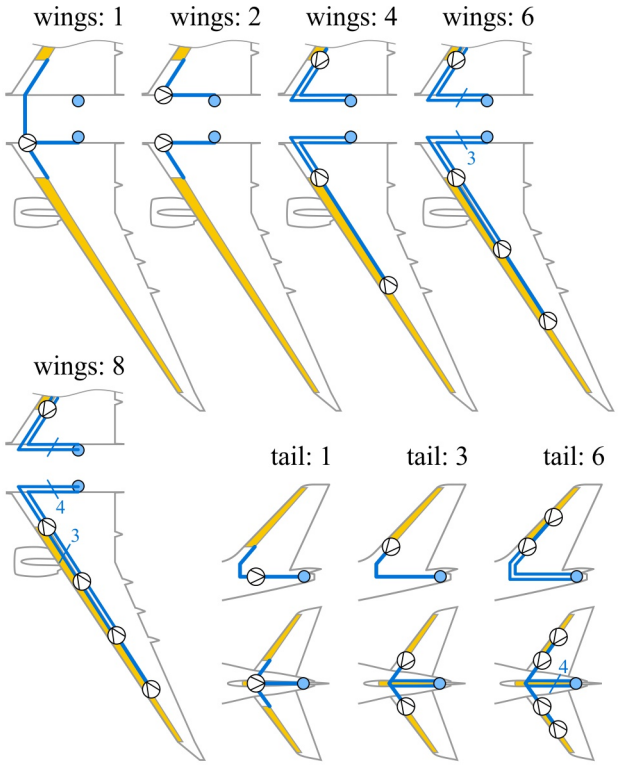


Fig. 18 Investigated number of compressors on wings and tailplanes

The compressors are designed such that at the system outlet a flow speed slightly above $Ma = 0.2$ and a static pressure equal to the ambient pressure is reached.

The main output of this study is the power consumption and mass breakdown for different topologies, see Fig. 19 to Fig. 22.

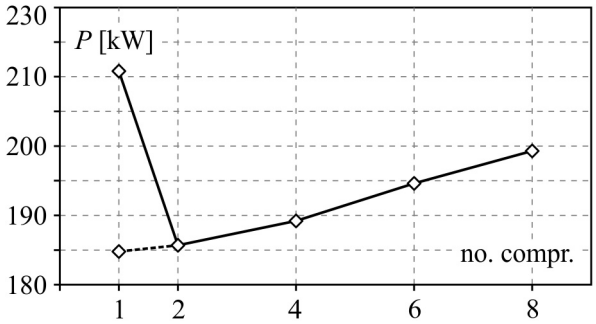


Fig. 19 Overall electrical power of the wing suction system architectures

The power consumptions for the wing architectures in Fig. 19 show a considerable higher power for the 1-compressor configuration. The reason is that the blade diameter had to be decreased so that the compressor fits into the available space. This caused a decrease of the efficiency (according to the CORDIER diagram in Fig. 11). The dashed line represents a compressor design, if no space limitation is given. In this case all compressor efficiencies are nearly equal and the higher power consumption of the architectures with a higher number of compressors is the result of higher pressure losses in the longer ducts.

The mass breakdown is shown in Fig. 20. Again, the dashed lines represent the compressor design without space limitations. Motors, frequency converters and wiring is summarised as “electrical equipment” and is basically dependent on the power consumption.

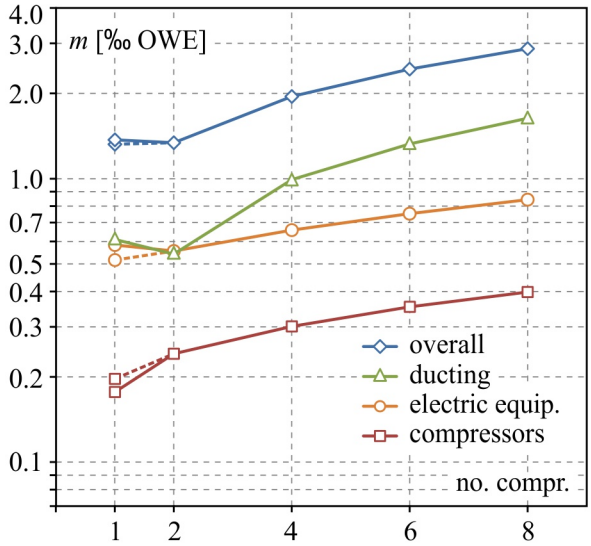


Fig. 20 Mass breakdown of the wing suction system architectures relative to Operating Weight Empty (OWE)

The ducting network contributes with a considerable part to the overall mass, especially for architectures with a higher number of compressors. Of course, this is influenced by the constraint that in this study each compressor has a separate duct to the outflow valve. A collective duct would decrease the slope of the ducting mass curve. Since the duct lengths to the outboard compressors are long, even with collective ducts, the ducting mass would stay quite high.

As mentioned in Sec. 2.6, the compressor mass calculation contains uncertainties. Because the compressor mass is small compared to the ducting and electrical equipment, the influence on the overall mass is small.

Power consumption and mass breakdown of the tailplane architectures are shown in Fig. 21 and Fig. 22. Here, all compressors fit into the assumed space allocations. The increase in

power and mass is caused mainly by the increase in overall duct lengths.

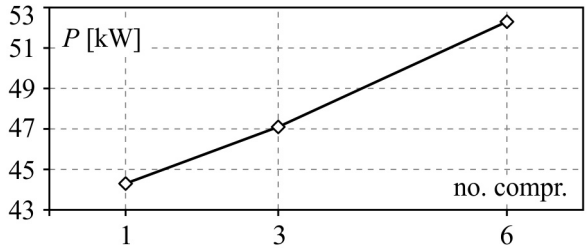


Fig. 21 Overall electrical power of the tailplane suction system architectures

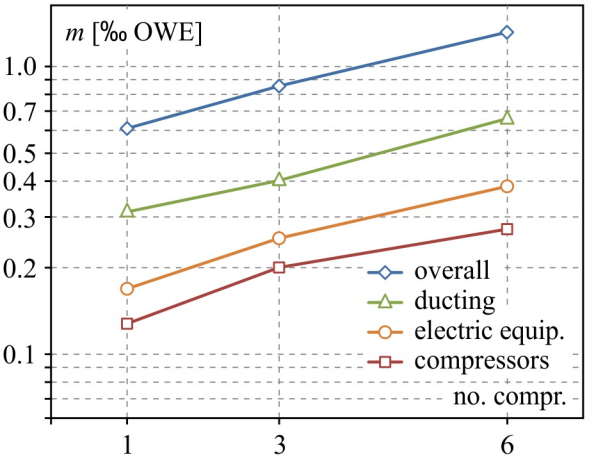


Fig. 22 Mass breakdown of the tailplane suction system architectures relative to Operating Weight Empty (OWE)

Using the CORDIER diagram for compressor design yields high efficiencies, which are in this work nearly equal, except for the 1-compressor wing architecture. The specific speed σ was chosen to be around 0.5, resulting in an efficiency of around 0.8. For σ = 0.5 a diagonal compressor is optimal. Recalling the definition of the specific speed [15]

$$\sigma = n \cdot \frac{\sqrt{\dot{V}}}{(2\,Y)^{\frac{3}{4}}} \cdot 2\sqrt{\pi}\ , \qquad (21)$$

one recognises that a decreasing volumetric flow rate (higher partition of the suction areas) results in an increasing speed to keep σ at the same value. This effect appears in this topology study results: the maximal rotational speed of the compressors increases with increasing number of compressors, see Fig. 23.

Such high rotational speeds are not unusual for compressors (e.g. compressors in aircraft environmental control systems), but could be difficult for electric motors and frequency converters. Further investigations should include an analysis of reasonable speeds for the electric drive system.

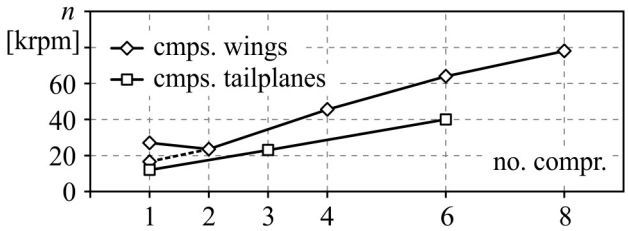


Fig. 23 Maximal compressor rotational speeds of system architecture

5 Conclusion and Outlook

The results of this study in terms of mass and power consumption suggest using an architecture with few compressors. One reason is that fewer and shorter ducts can be used, which have a considerable contribution to the overall mass. Furthermore, slower rotational speeds (with high compressor efficiencies) are possible, which eases the drive system design and control. For the wings architecture the asymmetric 1-compressor configuration is not recommended due to the necessary cross duct.

For the wings a 2-compressor configuration and for the tailplanes a 1-compressor configuration is recommended (for an integrated HLFC aircraft). It should be noted that no safety and reliability issues have been included into this assessment. In terms of reliability it is favourable to have a higher partition into subsystems. The probability that several (independent) subsystems fail is lower than the probability that a single (similar) system fails. This characteristic could be used to guarantee that a certain percentage of the suction system is functional.

A major problem for the suction system is the available space. This applies especially for the wing LE box, where the suction system has to share the space with high lift devices, anti- or de-icing systems and other systems. The here used space assumptions are rather optimistic. For a more accurate assessment of the integration into the LE box, a feasible high lift and anti-/de-icing concept have to be available. If possible, the compressors should be moved into the belly fairing or to a part of the LE box near the wing root, where no high lift system is located.

Because the tailplanes do not use high lift devices, the available space is larger, but a 1-compressor configuration would lead to problems anyway. It is recommended to locate the suction system in the tailcone, where more space is available.

The here presented sizing methodology and architecture studies are intended to be used in preliminary aircraft design. Besides the system data, of course detailed aerodynamic information regarding the airfoil geometries and drag reductions, as well as structural information are necessary.

From a systems point of view the next important step is the linking of this architecture studies with reliability analysis. The reliability of the system (and also of the structure and surfaces) is of great importance. If laminarity is lost, the aircraft operates with considerable higher drag and carries the additional mass of the suction system. This could lead to a flight diversion, if the destination airport cannot be reached anymore. To encounter this, additional reserve fuel could be carried with, but this extra mass decreases the aircraft's efficiency. The other possibility is to design suction systems and structures such that (a certain percentage) of laminarity can be guaranteed with high probability. For the system this could be realised by redundancies, but causing extra mass. It is necessary to find an optimal configuration taking these considerations into account and assess the technology on mission level with respect to mission fuel consumption.

References

[1] Robert J.-P., "Drag Reduction: An Industrial Challenge", AGARD Special Course on Skin Friction Drag Reduction, Rhode-Saint-Genèse, 1992, pp. 2.1–2.15.

[2] Anderson J. D., *Aircraft Performance and Design*, McGraw-Hill Boston, 1999.

[3] Green J. E., "Laminar Flow Control – Back to the Future?", AIAA 38th Fluid Dynamics Conference and Exhibit, Seattle, 2008.

[4] Schrauf G., "Status and perspectives of laminar flow", The Aeronautical Journal, Vol. 109, No. 1102, 2005, pp. 639–644.

[5] Horstmann K. H., Schrauf G., Sawyers D. M., Sturm H., "A Simplified Suction System for an HLFC Leading Edge Box of an A320 Fin", CEAS Aerospace Aerodynamics Research Conference, London, 2002, pp. 91.1–91.7.

[6] Anderson J. D., *Fundamentals of Aerodynamics*, 3rd ed., McGraw-Hill Boston, 2001.

[7] Schrauf G., "A Program to Compute Velocity and Temperature Profiles for Local and Nonlocal Stability Analysis of Compressible Conical BL with Suction", ZARM Technik Report, 1998.

[8] Pe T., Thielecke F., "Methodik zur Leistungsabschätzung von HLFC-Absaugsystemen im Flugzeugvorentwurf", DGLR Deutscher Luft- und Raumfahrtkongress 2010, Hamburg, 2010.

[9] Schlichting H., Gersten K., *Grenzschicht-Theorie*, 10th ed., Springer Berlin, 2006.

[10] Bieler H., "HLFC for Commercial Aircraft – First ELFIN Test Results", 1st Euro. Forum on Laminar Flow Technology, Hamburg, 1992, pp. 193–200.

[11] Wagner W., *Strömung und Druckverlust*, 6th ed., Vogel Buchverlag Würzburg, 2008.

[12] Miller D. S., *Internal Flow Systems*, 2nd ed., BHRA Cranfield, 1990.

[13] Idelchik I. E., *Handbook of hydraulic resistance*, 3rd ed., CRC Press Boca Raton, 1994.

[14] Cordier O., "Ähnlichkeitsbedingungen für Strömungsmaschinen", *Brennstoff, Wärme, Kraft (BWK)*, Vol. 5, No. 10, 1953, pp. 337–340.

[15] Bohl W., Elmendorf W., *Strömungsmaschinen 1*, 10th ed., Vogel Buchverlag Würzburg, 2008.

[16] International Electrotechnical Commision, "Efficiency classes of single-speed, three-phase, cage-induction motors (IE-code)", IEC 60034-30, 2008.

[17] Commission of the European Communities, "Commission Regulation (EC) No 640/2009", 2009.

[18] Koeppen C., "Methodik zur modellbasierten Prognose von Flugzeugsystemparametern im Vorentwurf von Verkehrsflugzeugen", Ph.D. Dissertation, Institute of Aircraft Systems Engineering, Hamburg University of Technology, Germany, 2006.

[19] Schubert H., "Modellierung und Bewertung der charakteristischen Eigenschaften elektrischer Netzwerke zur Abschätzung der Systemmasse im Vorentwurf von Verkehrsflugzeugen", diploma thesis, University of Stuttgart, Germany, 2008.

[20] Heidmann P., Schäufele S., "HLFC research baseline", Airbus Operations GmbH, 2010.

[21] Schrauf G., Meyer P., Sawyers D., "Application of Hybrid Laminar Flow Technology on Transport Aircraft, Final Report", Airbus Deutschland, 2003.

[22] Rodax B., "HLFC Profiles", Deutsches Zentrum für Luft- und Raumfahrt (DLR), 2011.

[23] Rudolph P. K. C., "High-Lift Systems on Commercial Subsonic Airliners", NASA Contractor Report 4746, 1996.

CEAS 2011

Preliminary Design Optimization Methods

3rd CEAS Air&Space Conference
21st AIDAA Congress

Maneuver Loads Calculation via Optimization Techniques

Ignacio Ruiz Juretschke+, Maria Calvo Blanco#, Juan Román García* and Javier Hilario Montes§

+Aerospace Engineer at Airbus Military Static Loads Dept. PhD student at ETSIA-UPM, Spain
#Aerospace Engineer at Airbus Military Static Loads Dept. HO Loads Models, Spain
**Aerospace Engineer. HO Airbus Military Product Certification, Spain*
§Aerospace Engineer. HO Airbus Military Static Loads Dept. Lecturer of Algebra and Calculus at the Fundamentals of Mathematics Dept. of ETSIA-UP, Spain.

Keywords: *Optimization, Maneuver, Loads*

Abstract

EADS-CASA, now Airbus Military, Static Loads Dept. has been working on the definition of a new methodology for fast loads calculation based on optimization techniques. This project is, oriented to time and cost reduction in the field of load calculation.

The main target of this new methodology, christened Aircraft Response Confinement (ARC), lays in the de-coupling of Flight Control Laws design from Loads Analysis, something that has become necessary due to the evolution of Flight Control Systems incorporating more complex features and not being available at the time Design Loads are required, as well as in the enormous advantage in terms of time cost and computational effort implied in the avoidance of flight simulations for loads analysis.

The present paper develops an application exercise of this methodology to a medium range UAV (Unmanned Aerial Vehicle) currently under design, comparing the results against those obtained via classic simulation.

1 Introduction

Current A/C Load Calculation is based on maneuver simulation, i.e. the calculation of the A/C response to control surface deflection (in terms of flight parameters) and of the loads appearing at each A/C section for the previously obtained flight parameter evolution.

This usually means performing a huge number of simulations to guarantee structural integrity at all flight conditions required by the international regulations (CS, FAR, JAR...) and/or the Structural Design Criteria (SDC). The simulation being rather time costly by itself, it also implies the availability of a flight simulator of the aircraft by the time Load Analysis begins, something that is not usual, especially in modern fly-by-wire or fly-by-light A/Cs with complex guidance laws, which development generally spans all through the design phase and is performed in parallel to Load Analysis by the Flight Mechanics Department.

Adding the fact that the standard procedure for Static Loads Calculation within a full A/C project requires of several load calculation loops (Preliminary Design, Design, Check-Stress,

Certification, etc.), a methodology that significantly reduces the load calculation time and simultaneously frees the Loads Department from its dependency of the Flight Mechanics Department would be of great value.

The present paper demonstrates the validity of the optimization based load calculation procedure by comparing the results achieved by this methodology with those obtained via classic simulation based methods. It has been conceived as a follow up of the work presented at CEAS 2009 by J.M. Roman "*A New Approach to the Aircraft Manoeuvre Loads Problem*" [1], where the optimization based approach to the aircraft maneuver loads problem is described theoretically for a standard A/C equipped with an advanced flight control system (FCS).

2 Loads Model

The methodology described in this paper is built around the A/C's Loads Model [4]. The Loads Model is the foundation of any loads calculation procedure, whether via classic simulation approach, via optimization or any other. It is a dataset relating loads at different aircraft stations, called Monitor Stations (MS), with the variable (aircraft parameters) set used to describe the A/C's state (attitude, configuration, speed, acceleration, etc.), whether in flight or on the ground.

The Loads Model is structured at EADS-CASA in three main parts:

- Aerodynamic Loads Model
- Inertial Loads Model
- Miscellaneous Loads Model

The Aerodynamic Loads Model relates Loads at MS with a set of aerodynamic variables, such as Mach, Dynamic Pressure, angle of attack (α), side-slip (β), control surface deflection (δ's) etc... It is constructed with the aerodynamic unitary effects for each variable provided by the Aerodynamic Dept., which are obtained generally through fluid dynamic computations and corrected with Wind Tunnel Test data.

The Inertial Loads Model relates Loads at MS with the inertial variables, usually 12, including load factors (Nx, Ny, Nz), angular accelerations ($\dot{p}, \dot{q}, \dot{r}$), and the cross products of the A/C's angular rates (p^2, q^2, r^2, pq ...) for each considered mass state of the A/C.

The Miscellaneous Loads Model is created to consider any other load affecting the A/C, such as thrust, propeller 1P loads or loads introduced by the landing gear. For the present exercise only thrust and 1P effects have been considered.

Thus a generic formulation of the Loads Model is:

$$L_i = f\left(\vec{V}\right) \qquad (1)$$

Where L_i represents forces and moments at each MS and $\vec{V}$ includes aerodynamic, inertial and miscellaneous variables.

The Loads Model is formulated as linear as possible, but it is non-linear at least for the quadratic angular rates terms in the inertial model. Some aerodynamic effects are treated as linear in the model, but corrected with efficiencies to take into consideration possible non-linearities (section wise linear).

3 Classic approach to Maneuver Load Calculation

The classic approach to Maneuver Load Calculation is generally based on maneuver simulation to obtain the variation of the A/C parameters $\vec{V}$ in the time domain. This allows calculating via Loads Model the evolution of loads at all MS. This procedure must be repeated to cover all maneuvers required by the certification standards (FAR, CS, DEFSTAN, STANAG etc.) [5], [6], [7], [8] or the high level Design Criteria. (An exception to this standard procedure are those A/C designed by trimming the A/C at the borders/corner points of the flight

envelopes defined in the SDC, under the assumption that those are the worst possible conditions. Still an usually complex trimming software is required). Some typical maneuvers, classified in longitudinal or lateral are:

Longitudinal:

- Pull-Up and/or Push-Down, balanced quasi-steady conditions.
- Coordinated Turn, balanced steady condition.
- Pitch maneuvers, including Pull-Up/Push-Down from 1g, pitch recoveries from max./min. Nz etc.

Lateral

- Yaw maneuvers, including sudden rudder action, overswing to maximum side-slip, steady side-slip conditions and centering of yaw control.
- Roll maneuvers considering limiting combinations between Nz and roll rate/acceleration

These maneuvers have to be repeated for each considered mass state across the flight envelope to ensure compliance with the requirements.

The loads obtained are afterwards post processed in accordance with a selection criteria defined by the Stress Dept. to determine the critical load cases for each MS, usually in form of maximum and minimum of each load function (1D envelopes, i.e. Fx, Fy, Fz, Mx, My and Mz at each MS) and of convex envelopes (2D envelopes), i.e. critical combination of two functions, like Fy vs. Fz at a specific station.

The described 1D and 2D envelopes are the basis for design and structural sizing of the aircraft.

This overall process is subject to the availability of an adequate flight simulator, which is to allow a simulation as similar as possible to real flight. As already mentioned, the flight control laws are developed by the Flight Mechanics Dept. all across the A/C's design phases. Unfortunately many times this implies that these control laws are not available in a representative state until the latter phases of the overall process.

This means that load calculation has to begin with generic, provisional, little representative versions of the flight simulator, which usually renders less accurate loads and can lead to erroneous dimensioning. This has to be corrected on later stages, representing a considerable amount of effort for all departments involved (Loads, Flight Mechanics, Stress etc..). This is especially dramatic for A/Cs with natural instabilities along the flight envelope, which cannot be controlled without a system with the proper stabilization functions and control schedules.

Here is where the new methodology sets in, offering a fast, non-expensive and simple way to estimate loads at all monitor stations without the need of a flight simulator or complex trimming software.

4 Structural Design Criteria

The main requisite for the proposed methodology is the clear definition of the design criteria, i.e. a clear definition of the borders of the flight and/or maneuver envelopes in terms of flight parameters that the A/C will have to endure and sustain without damage.

This means defining the intended operational envelope before starting the design process in terms of parameter response boundaries. It represents an agreement between Loads and Flight Mechanics teams, where the former will analyze the boundaries of the defined envelopes for structural design and the latter will ensure that the final flight control laws do not let the A/C leave these areas, in terms of parameters, as depicted in Fig.1. This is why the new methodology has been named "Aircraft Response Confinement" (ARC). The definition of such a detailed Structural Design Criteria is a task that requires important amounts of engineering expertise and a very clear idea of

what the final product is expected to do in terms of maneuver performance.

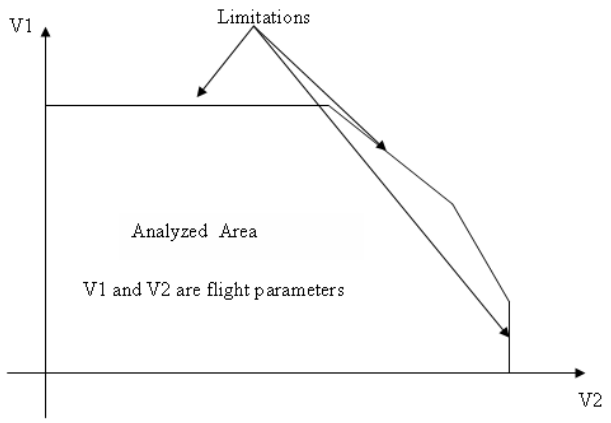


Fig.1 ARC concept

For the present exercise an aircraft, an Unmanned Aerial Vehicle (UAV) currently under development, with a relatively simple Loads Model and Flight Control System has been chosen. Official design is being performed applying the classic, simulation based method, which means that there is already sufficient data and knowledge on the A/C's operation to contrast the results obtained via the proposed optimization methodology. It also means that the SDC had to be defined purposefully for this exercise, as none was available. This was not a simple task, and most envelopes are based on the applying regulations, on the future mission of the aircraft and on previous experience with some more complex A/Cs where a SDC was developed.

The following graphs show the final applied SDC. The first one depicts the maneuver envelope, a single flight point (marked in Fig.2) having been chosen to develop the optimization exercise, as for the initial stage of the study it was decided to consider speed and altitude as known parameters instead of variables. The chosen flight point is *Vc=60m/s* EAS (Equivalent Air Speed) and *hc=20000ft*, where *Vc* and *hc* are cruise speed and altitude. Also a single mass configuration was chosen, in this case the ZFW (Zero Fuel Weight) of the baseline configuration.

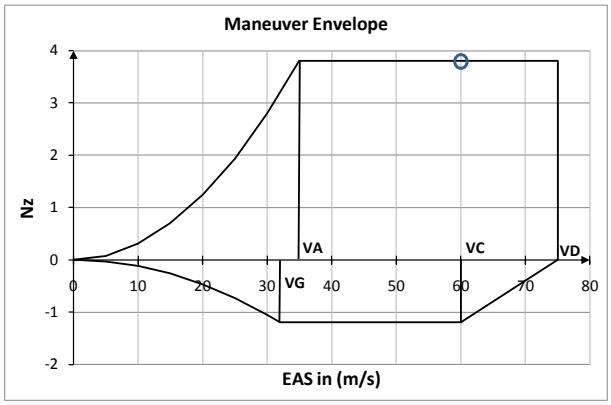


Fig.2 Maneuver Envelope

Figure 3 shows the limitations for combinations of vertical load factor and the lateral parameters, i.e. roll and yaw rate and acceleration. As can be seen, the A/C is expected to sustain maximum Nz up to 50% of roll and yaw rates/acceleration.

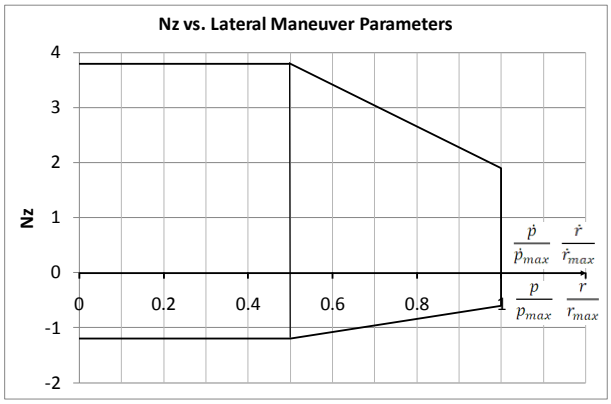


Fig.3 Nz vs Max. Lateral Maneuver Response Parameters

Figure 4 depicts the restrictions imposed on lateral load factor and yaw acceleration.

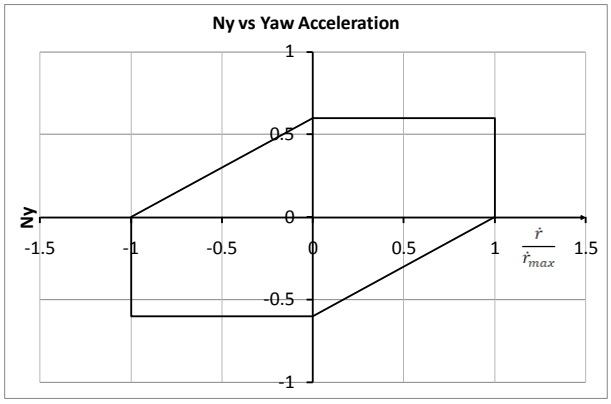


Fig.4 Ny vs Yaw Acceleration

Figure 5 shows the last restriction imposed on angular rates and accelerations, specifying that at a 100% of any of them, a maximum of a 33% of the other can be reached. This restriction deserves a special mention in the chapter addressing the optimization constraints, as it is none-linear.

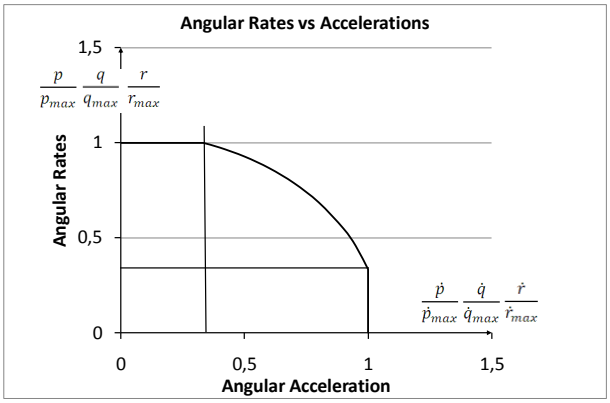


Fig.5 Angular Rates vs Accelerations

Apart from these high level restrictions each variable is assigned a maximum and a minimum value, as described in the next section.

5 Loads Calculation via Optimization

The new load calculation methodology is based on the Loads Model, as it is the model that describes the loads at each MS as a function of the flight parameters. Thus each function of the Loads Model is taken *sequentially* as target of an optimization process, to determine the variable vectors that render its maximum and minimum value. The word *sequentially* has been highlighted to emphasize the fact that in this initial exercise each function's maximum and minimum has been searched independently, and not in a multi-objective way, something that would be of great interest when trying to define 2D load envelopes (combination of two loads). Multi-objective optimization has inherent complexities that have not been addressed in this study yet, although at the end of this paper some methodologies are proposed to obtain 2D load envelopes based only on single objective optimization.

The Loads Model of the UAV considered for the present exercise accounts for 15 independent variables, given that for simplicity some variables were taken as constant parameters. These parameters are listed below:

- *hc=20000ft*
- *Vc=60m/s EAS*
- *Mass State*
- *Thrust & Torque*

Thrust and Torque require special considerations. Both are defined for this UAV in a Thrust and Torque Map, which is function of speed and altitude, as well as operation condition, which can be either maximum thrust or braking. Speed and altitude being parameters, the optimization procedure is performed for both conditions each time, and the worst resulting load (maxima & minima) for each section taken.

Table 1 depicts the independent optimization variables, its description and also its maximum and minimum value.

Variable	Description	Max.	Min.	Units
AOA	Angle of Attack	30	-30	deg
QND	Pitch Rate*	60	-60	deg/s
SDP	Side-Slip	25	-25	deg
PND	Roll Rate*	60	-60	deg/s
RND	Yaw Rate*	30	-30	deg/s
NX	Long. Load factor	0.7	-0.2	-
NZ	Vert. Load Factor	3.8	-1.2	-
NY	Lat. Load Factor	0.6	-0.6	-
PD	Roll Accel.	140	-140	deg/s²
RD	Yaw Accel.	130	-130	deg/s²
QD	Pitch Accel.	230	-230	deg/s²
DR1_ef	Left Rudder	17.25	-17.25	deg
DR2_ef	Right Rudder	17.25	-17.25	deg
DA1_ef	Left Aileron	9.4	-20	deg
DA2_ef	Right Aileron	9.4	-20	deg
* Angular Rates are non-dimensioned in the Loads Model with Mean Aerodynamic Chord and True Speed				

Table 1 Loads Model Variables

These variables are either applied directly or combined to form other ones, such as symmetric and antisymmetric rudder deflection, which are the ones actually used for the Aerodynamic Model. Regarding the variable set someone might miss elevator deflection, but the UAV taken for this exercise has a V-shaped tail,

which means that rudder and elevator are a same surface that couples both effects.

The optimization target function L_i, where L_i can be any of the Loads Model functions, is in all cases non-linear, as the Inertial Model is quadratic with respect to angular rates. The Aerodynamic Model has a certain level of non-linearity, as some variables (e.g. asymmetric rudder deflection or tail side-slip) are corrected with different effectiveness values depending on the angle of attack.

The optimization is performed attending to all restrictions mentioned in section 4 and to the variable limits in table 1. These last ones are included as upper and lower boundaries for the variables:

$$\vec{L}_B \le \vec{V} \le \vec{U}_B \qquad (2)$$

The constraints in Fig.3 & Fig.4 are introduced as linear inequalities, i.e.:

$$A \cdot \vec{V} \le b \qquad (3)$$

The ones in Fig.5 are introduced as non-linear inequalities.

$$c\left(\vec{V}\right) \le 0 \qquad (4)$$

But the most important constraints of all are those that ensure that the aircraft is trimmed in its movement as a rigid body in all conditions. The first six functions in the Loads Model describe forces and moments in the centre of gravity of the aircraft and these must be zero to fulfil this vital condition.

$$L_i\left(\vec{V}\right) = 0 \quad i = 1 \text{ to } 6 \qquad (5)$$

If the problem weren't complex enough, one final consideration has to be done on the calculation of the 1P loads. 1P loads are those loads deriving from the aircraft's propeller in its own rotation plane, i.e. Fy, Fz, My and Mz, as can be seen in Fig.6.

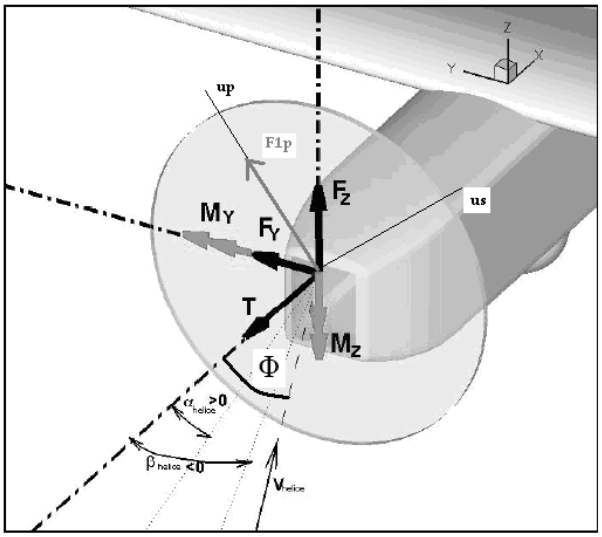


Fig.6 1P Loads

For the classic simulation based method the 1P loads were formulated in accordance with ESDU-89047 [2], which means that the derivatives of forces and moments with respect to the local inflow angle Φ were interpolated as a function of Thrust Coefficient (CTH) and Advance Ratio (J), which are function of thrust and speed. The same formulation was used for the optimization problem, but with a simplification in the decomposition of the angle of attack and side-slip of the propeller. This was necessary due to convergence problems that appeared when using the original decomposition. Still the 1P loads depend in the end on the angle of attack, side-slip, and all three angular rates. The error introduced is small, but still sufficient to explain small deviations in the final results, as will be explained later on.

The optimization code was also prepared to simultaneously render all correlated loads L_i for each different variable vector V.

The whole exercise was developed in a commercial optimization software suite with the capability of solving non-linear problems subject to linear and non-linear constraints. More information on this software can be found in [3].

Though the optimization code has not yet been optimized for fast computation, the overall optimization process of all functions and the simultaneous calculation of the correlated loads took just a few minutes. This gives an idea of the potential benefit in terms of computational effort this methodology may offer.

6 Results

The evaluation of the results for all load functions of this model is an exhausting exercise, because even though this is a rather simple model there are 204 of them. Therefore this chapter will focus on a representative subset, composed of the foremost fuselage monitor station (FUS_C01), the wing root section (WR_0300), and the V-tail stabilizer (EL_ESB0), depicted in Fig.7 (for visual reasons ER_ESB0 has been graphed, it is the symmetric right hand station to EL_ESB0). This figure also shows the local reference systems in which loads are calculates. Maximum positive force is along the axis, while maximum negative (minimum) one is in negative axis direction.

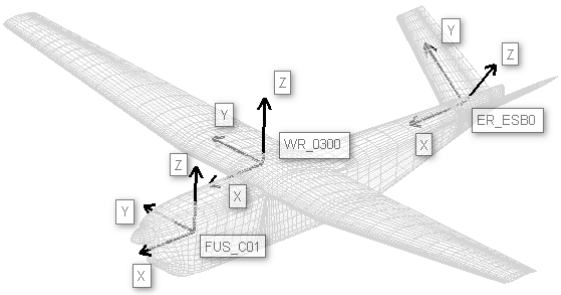


Fig.7 Monitor Stations and Axes

In the following table 2 a summary of the results obtained for these stations is given and compared to the results achieved via flight simulation.

	Simulation		Optimization		Difference (%)	
Function	Max	Min	Max	Min	Max	Min
FUS_C01_FY	0.683	-0.709	0.795	-0.658	16.4	-7.2
FUS_C01_FZ	0.955	-3.323	1.017	-3.448	6.5	3.7
FUS_C01_MX	0.259	-0.182	0.252	-0.189	-2.6	4.0
FUS_C01_MY	0.520	-1.225	0.736	-1.607	41.4	31.1
FUS_C01_MZ	0.442	-0.442	0.550	-0.650	24.6	47.1
WR_0300_FZ	6.456	-1.738	6.566	-1.861	1.7	7.1
WR_0300_MX	10.623	-3.049	11.280	-4.287	6.2	40.6
WR_0300_MY	1.105	-1.353	1.234	-1.475	11.7	9.0
EL_ESB0_FZ	0.731	-1.384	1.031	-1.891	41.0	36.6
EL_ESB0_MX	0.711	-0.962	0.780	-1.200	9.8	24.8
EL_ESB0_MY	0.157	-0.263	0.208	-0.289	32.0	10.0
	Results in kN and kNm					

Table 2 Results & Comparison

As can be seen optimization results are in general conservative, something that was expected and intended, but not overly so. It all depends on the quality of SDC. In this case a rather simple set of restrictions was imposed, but a more extensive one would probably have narrowed the margins. Notice that the two greatest loads, wing shear force and bending moment are rather well approximated.

Worrisome are those cases where the methodology was not conservative, like minimum fuselage Fy and maximum Mx, even if it is a very small percental difference. These two stations presented the greatest under-estimations of the whole study, and considering that they are directly affected by the 1P-loads and the thrust and torque modeling, this level of error is considered acceptable.

Thus it may be concluded that this methodology is capable of offering in a short time a conservative but rather realistic approximation of the loads.

Now the question arises if some information can be gained in terms of convex envelopes of two forces or moments, and not only maximum and minimum of each. With the information currently available, a coarse approximation to this 2D problem would be plotting the original, (simulation based) envelopes against the rectangle shaped by the maximum and minimum of each load.

The following graphs, Fig.8 to Fig.10, show several examples.

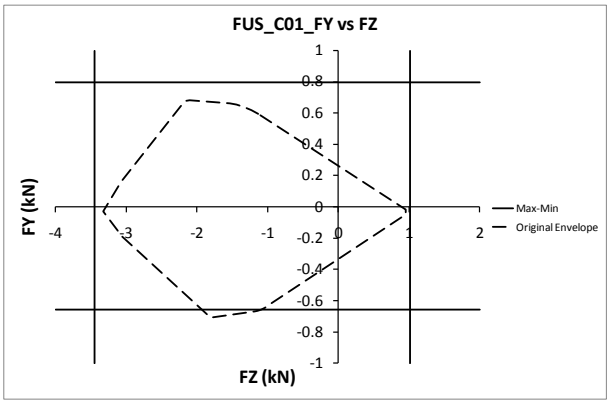


Fig.8 Fuselage FY-FZ. Original envelope vs. MAX/MIN

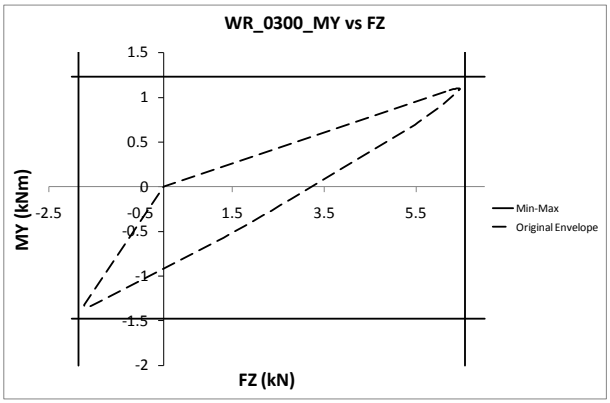


Fig.9 Wing Root MY-FZ. Original envelope vs. MAX/MIN

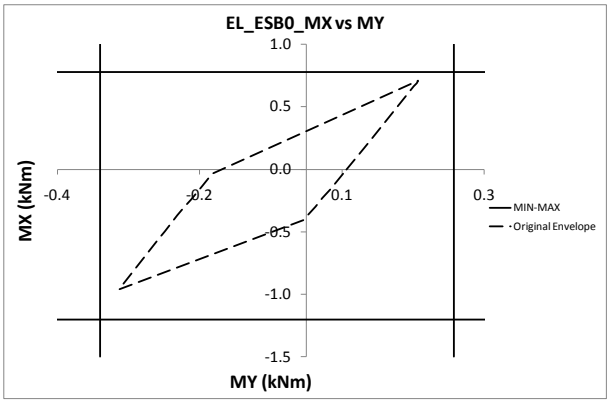


Fig.10 Stabilizer MX-MY. Original envelope vs. MAX/MIN

It is evident that such a coarse approximation to the 2D envelope is of little use. But this doesn't mean that nothing better can be achieved with the single objective optimization. Two different methods are proposed to obtain more accurate convex envelopes with the same optimizer.

The first method is of pure geometrical nature. Given two loads (L_1 and L_2), each with its maximum, minimum and correlated loads, one can define four points of the plane, which means having at least eight conditions, four related to position of the points and four related either to horizontal or vertical tangency. Eight conditions are too many to determine a conic section. Thus the problem has to be split in four sub-problems, one per arch section between two of these points. Then there are four conditions for each elliptical section, i.e. two points to pass through and two tangency conditions. The fifth condition necessary to define a single ellipsis is chosen to be the selection of the ellipsis, which principal axes also pass through the two available points.

Fig.11 tries to give a graphic explanation of the method.

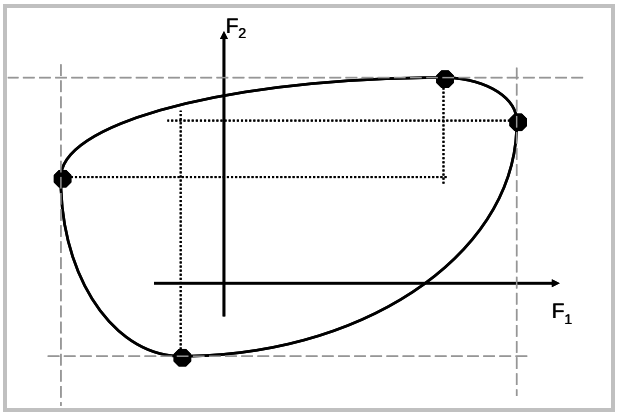


Fig.11 Graphic explanation of Method 1

Another approach to the convex envelope problem using single objective optimizers is the definition of a target function for the optimization problem that is a balanced combination of both loads, L_1 and L_2.

The combination is done through the components of a rotating vector, i.e. the maxima are projected on this vector, while the equation is balanced between both loads by dividing each term by its absolute maximum. Mathematically the function is:

$$F = \frac{L_1}{\max abs(L_1)}\cos\theta + \frac{L_2}{\max abs(L_2)}sen\theta \qquad (6)$$

For each value of θ an optimization process of F has to be performed. For this exercise θ was varied in five degree steps, and each step takes about a minute to converge. Also the absolute maximum of each load has to be known beforehand, which means that this process can only be applied after the original optimization process.

Figures 12 through 14 compare the results obtained with both methods, in comparison with the original convex load envelopes resulting from simulation.

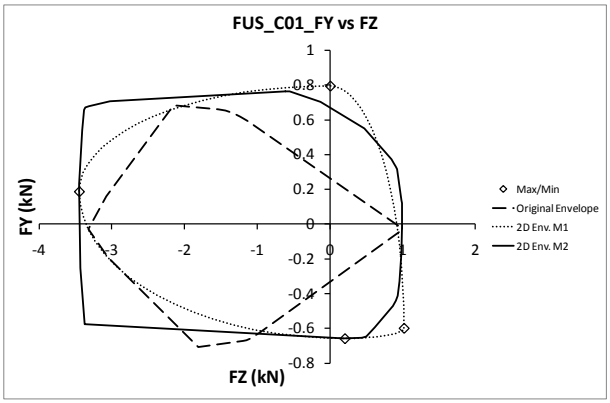


Fig.12 Fuselage FY-FZ. Original envelope vs. Methods 1 & 2

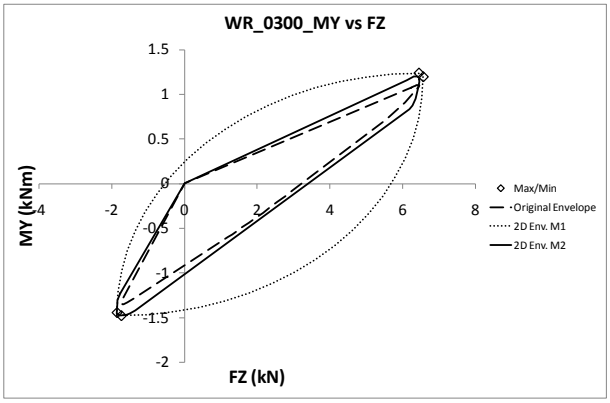


Fig.13 Wing Root MY-FZ. Original envelope vs. Methods 1 & 2

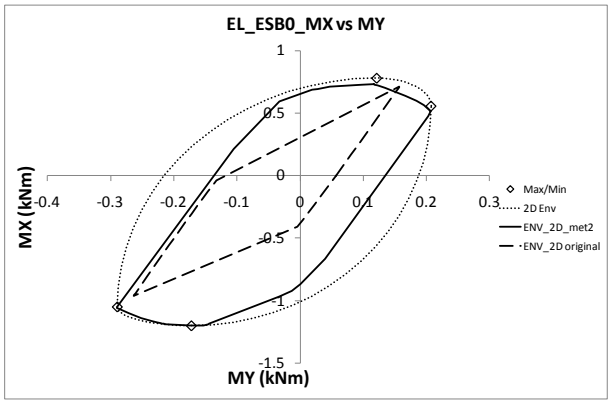


Fig.14 Stabilizer MX-MY. Original envelope vs. Methods 1 & 2

As can be seen both methods show an appreciable reduction of the envelopes area if compared to simple MAX/MIN approach. Yet Method 1 seems more case dependant, as can be seen comparing Fig.12 and Fig.14. In the former Method 1 results rather non-conservative, especially in the left lower corner, while in Fig. 14 it results overall conservative.
Method 2 gives very good approximations in Fig.13 and 14, while in Fig.12 it is overall conservative (remembering that the overall procedure was non-conservative in terms of negative FY by a 7%).

Therefore it seems that the effort of doing an extra set of optimization loops to apply Method 2 is worthwhile, even if Method 1 is probably more robust and reliable, being based on simple geometrical considerations. Another advantage of method 2 is its capability to render correlated loads along the A/C for each one of the corner points forming the 2D convex envelope, something that method 1 can't do.

7 Conclusions

The present paper proposes a new methodology for loads calculation, based on the Loads Model and single objective optimization.

Starting from the work done for a UAV currently under development, the methodology has been proven to give valid, generally conservative results, accurate enough to be used

as Preliminary Design Loads, if not directly Design Loads.

The optimization based calculation process is fast, in the order of minutes, and has no need for a simulation tool. These two properties enable the Loads Dept. to overcome two major bottlenecks in the overall load calculation process, i.e. the dependency on the Flight Mechanics Dept. to deliver updated simulators in time and the considerable cost in terms of computation of the simulation of all required maneuvers.

The methodology only requires the clear and adequate definition of a Structural Design Criteria, something that isn't trivial either.

Finally two methods are proposed to obtain convex load envelopes using only single objective optimization. Both methods result satisfying, especially the second one, based on the optimization of a function combination of two loads.

8 References

[1] Román J.M., "A New Approach to the Aircraft Manoeuvre Loads Problem," CEAS 2009, Manchester UK

[2] ESDU 89047, "In-plane forces and moments on installed inclined propellers at low forward speeds," Issued December 1989, with Amendments A to C August 2006, RAES.

[3] Matlab Optimization Toolbox, R2009a, The Mathworks.

[4] Ruiz I. and Calvo M., "ATLANTE Loads Model", EADS-CASA 2010.

[5] Federal Aviation Authority, "PART 25 AIRWORTHINESS STANDARDS: TRANSPORT CATEGORY AIRPLANES," Federal Aviation Regulations, www.faa.gov

[6] European Aviation Safety Agency, C25 and C23, www.easa.europa.eu

[7] UK Defence Standardization, "DEFENCE STANDARD 00-970", Ministry of Defence, UK

[8] NATO STANDARIZATION AGENCY, STANDARIZATION AGREEMENT (STANAG) 4671, "UNMANNED AERIAL VEHICLE SYSTEMS AIRWORTHINESS REQUIREMENTS (USAR)", NATO.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Multi-objective and Multidisciplinary Optimization of Wing for Small Aircraft

W. Stalewski, J. Żółtak
Instutut Lotnictwa (Institute Of Aviation), Poland

Keywords: *multi-criteria design, multidisciplinary design, genetic algorithm, aerodynamic of wing*

Abstract

During this work a cost efficient methodology of small wing design and optimisation was developed. A parametric model of aircraft/wing has been designed. In this model both clean wing and wing with high lift system were parameterised. The multi-disciplinary, multi-objective optimisation technique was applied to design a turbulent wing for small aircraft. Finally the properties of designed wing have been tested for achieving given requirements.

1 Introduction

The costs of air transport on small and medium-sized distances are significantly greater than the costs of alternative modes of transport such as road or rail transport. This is one of the important barriers to improve the competitiveness of this sector. To reduce the costs generated in the different phases of aircraft's life cycle is the only way to change this situation. The Cost-effective Small Aircraft (CESAR)[1] project supported by the EU in 6th Frame Program was exactly focused on the development of technologies that reduce time to market entry and decrease the cost of such process for small aircraft. Two basic Aircraft Configurations (AC) [1] were selected as demonstrators: AC1 – low speed aircraft (turboprop) and AC2 – transonic speed very light business jet. For both configuration the turbulent and natural laminar wing technologies were developed.

In this paper the results of work performed at the Institute of Aviation (Instytut Lotnictwa) in the development of numerical methods of wing design for a small aircraft will be presented. The proposed methodology is based on parametric modelling of complex aerodynamic objects [2],[3] and numerical multi-criteria and multi-disciplinary optimization taking into account a various types of design constraints [4]. The method has been used to design the AC2 turbulent (AC2T) wing with high lift system for very light business jet.

[1] This research was supported by the FP6-AIP5-CT-2006 Integrated Project CESAR (Cost-effective Small Aircraft), Contract No.: 030888, *www.cesar-project.eu*

2 Design and Optimisation Methodology

The method for multidisciplinary and multi-objective optimisation, based on a genetic algorithm was adapted to small aircraft design. The general scheme of the method is presented in Fig. 1. The methodology couples two approaches: "*automatic*" numerical optimization (upper part of scheme) and "*manual*" direct-optimization (lower part of scheme). The "*automatic*" optimisation is fully controlled by *CODA3dgenetic* (generic algorithm) software calling other subsidiary codes. The program *CODA3Dparam* (parametric modelling of wing/ /aircraft) works in batch mode, delivering for given input - genotype, the real geometry of parameterised wing/aircraft, written in different formats readable by other software. The optimisation method used the typical objectives and constraints for aerodynamic design. Besides quantitative objectives additional qualitative criteria were applied. The aerodynamic performance was determined using 3D full potential method (*CODA3Dfps3dbl*) and 3D panel method (*CODA3Dpanel3dbl*) both coupled with 2D integral boundary layer solution and Vortex Lattice Method taking into account viscous effects (*Coda3Dvlm2)*. The box-beam model [5] of the wing structure was used to estimate a weight of the wing (*Coda3Dstruct*) . When it is possible, with respect to cost of the calculation, the objectives are obtained directly by CFD and/or CSA software. In other cases the Response Surface Method is used to evaluate objectives. The scheme of multidisciplinary analyses is shown of Fig. 2.

The result of numerical-optimisation process is a Pareto Set containing Pareto-optimal solutions. Final down selection of potentially "*best*" solutions from Pareto Set is done directly by designer. Next, selected solutions are tested in detailed analysis process. If needed, the designer, basing on his experience, may "*manually*" modify designed object to improve its properties. At this stage of design process, usually more sophisticated software is used for evaluating physical features of designed object.

The parametric modelling method, developed earlier in Institute of Aviation was adapted to small aircraft wing description. It uses a system of design components: scalars, points, curves, surfaces and solids. Curves, surfaces and solids are described using NURBS (Non-Uniform Rational B-Splines) [3]. Each component can be modified by change of value attributed design parameters set. To calculate

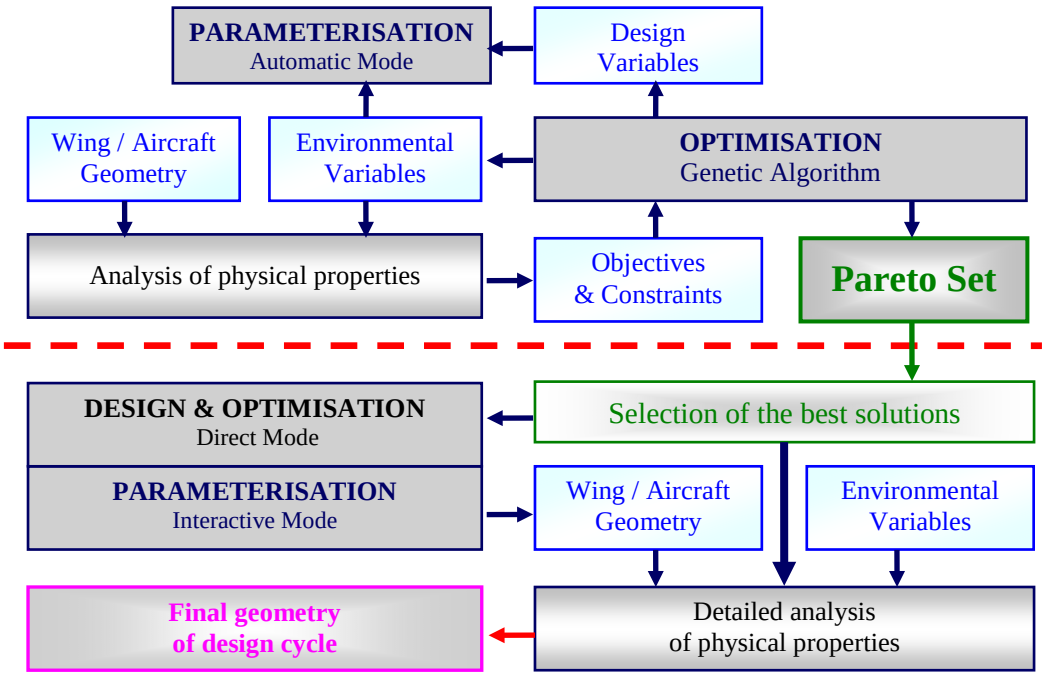


Figure 1. Design and optimisation methodology.

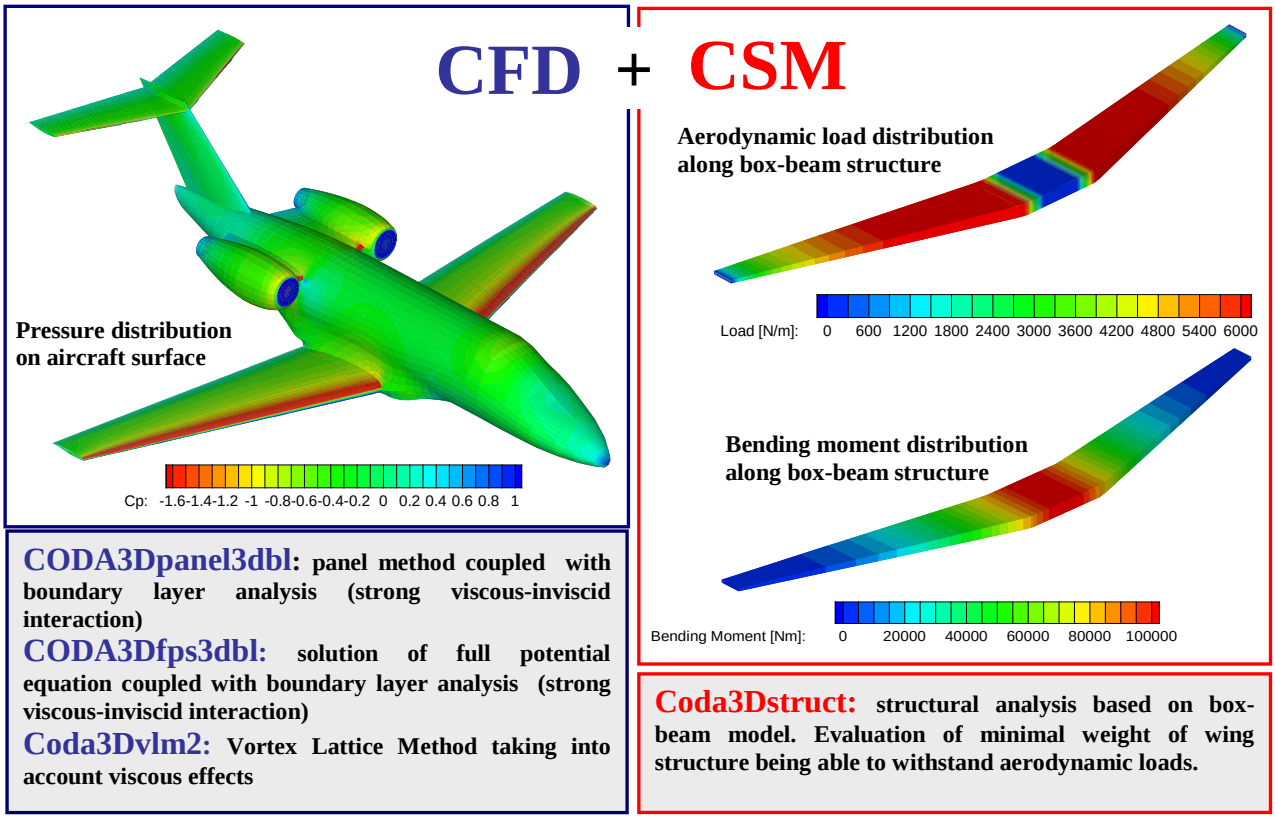


Figure 2 Scheme of simplified CFD and CSA codes used in turbulent-wing-optimisation process.

new values of design parameters some standard actions (e.g. translation, rotation, scaling) are used. For specific procedure (e.g. calculation of wing area or fuel tank volume) additional functions can be implemented and used.

Finally the generic parametric model of small aircraft wing geometry was proposed. In this model a relatively small number of parameters describe wide class of wing geometries without and with high lift devices. Parametric model of the wing/aircraft enables to parameterise both: the external surface of the wing/aircraft (for Computational Fluid Dynamic (CFD) analysis) and the internal structure of wing box (for simple Computational Structure Analysis (CSA) analysis).

3 Parametric Model of AC2-T Wing

Top Level Aircraft Requirements for AC2T aircraft concept was defined by Piaggio Aero and EADS-Socata [6].

Based on them it was assumed that wing platform was trapezoidal in form and fixed. It was considered that wing panels should have a classical, cost saving oriented manufacturability. It was assumed that the wing would consist of two segments having panels modelled by ruled surfaces:

- ✓ Flap Zone (FZ)
- ✓ Aileron Zone (AZ)

The span of AZ was established to 32,3% of wing semi span [6].

Three end-sections of wing segments completely defined the wing surface: Root Section (RS), Middle Section (MS) – border section between FZ – AZ and Tip Section (TP)

As basic airfoils, two airfoils [7] ,[8] were chosen:

- ✓ ILM115 (with relative thickness 15%c)
- ✓ ILM111 (with relative thickness 11%c)

All airfoils having other relative thicknesses were created by mixing two above basic airfoils.

The wing box was parameterised because its geometry was the input data for wing weight evaluation. Following positions of spars were established:

- ✓ Front Spar: 10% of the wing chord
- ✓ Rear Spar: 63% of the wing chord

For fuel capacity purpose it was assumed that the wing box total volume should not be lower than 0.6205 m^3 on each semi-wing. Parametric model of the wing enabled to control the volume of fuel tank. It was assumed that it would not be possible to create any geometry of the wing having this parameter smaller than acceptable minimum.

Finally, six design parameters were defined for the design and optimisation of clean wing:

- ✓ Parameter 1: Thickness of MS
- ✓ Parameter 2: Twist angle of MS
- ✓ Parameter 3: Twist angle of TS
- ✓ Parameter 4: Camber of mean line of RS
- ✓ Parameter 5: Camber of mean line of MS
- ✓ Parameter 6: Camber of mean line of TS

According to the design of high lift system for basic airfoil ILM115 [7], it was assumed that within the FZ the designed wing would be equipped with Fowler Flap (FF). The chord of the flap was established to 30% of wing chord.

Taking into account cost-saving-oriented manufacturability, it was assumed that the flap panel would be a ruled surface defined by its limiting cross sections: Flap Inner Section (FiS) and Flap Outer Section (FoS). The twist axis of the wing is placed at 36.5% chord, which is an equidistant point from front and rear spar positions.

A way of FF positioning in given cross-section of wing FZ is explained in Fig. 3.

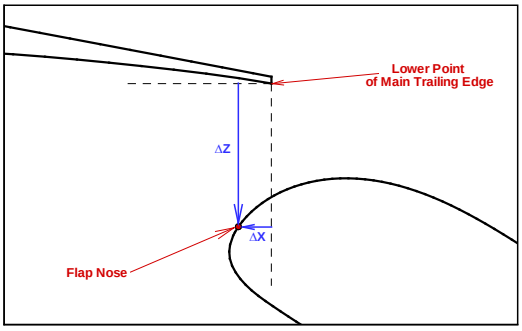


Figure 3 The definition of flap position in given wing cross-section.

It leads to following set of five parameters defining the position and deflection of flap:

- ✓ Parameter A: ΔX_i flap-nose position in FiS
- ✓ Parameter B: ΔZ_i flap-nose position in FiS
- ✓ Parameter C: ΔX_o flap-nose position in FoS
- ✓ Parameter D: ΔZ_o flap-nose position in FoS
- ✓ Parameter E: flap-deflection angle δ_{FL}

4 Multi-Disciplinary Optimisation of AC2-T Clean Wing

At the first stage of the design process, the clean wing was designed. It was achieved by solving appropriate multi-disciplinary, multi-objective optimization problem. The main goal of optimization was to design possibly low-drag and low-weight wing, having possibly high values of CLmax.

4.1 Definition of optimisation problem

In design process some Flight Conditions (FC) were considered –Table 1

Table 1. Flight condition [6]

	Mach number	Reynolds number	CL	Altitude [ft]
FC1	0.716	$10.0 \cdot 10^6$	0.16	31000
FC2	0.610	$10.0 \cdot 10^6$	0.50	41000
FC3	0.610	$6.0 \cdot 10^6$	0.75	41000
FC4	0.136	$5.7 \cdot 10^6$		

Base on then four following objectives were established

$$F_1 = C_{L\max} \qquad F_2 = \frac{L_1}{D_{w1}}$$

$$F_3 = \frac{L_2}{D_{w2}} \qquad F_4 = \frac{L_3}{W_{w3}}$$

Where:

CL_{max} – maximum of CL @ FC4
L_1 – lift of the wing @ FC1
D_{W1} – drag of the wing @ FC1
L_2 – lift of the wing @ FC2
D_{W2} – drag of the wing @ FC2
L_3 – global lift of aircraft @ FC3,
W_{W3} – minimal weight of the wing structure being able to withstand an aerodynamic load @ FC3

Additionally some aerodynamic constraints were defined. They expressed the following requirements:

- ✓ wing pitching moment coefficient must be higher than -0.08
- ✓ maximum of local Mach on the wing upper/ lower surface must not exceed assumed limits

No geometrical constraints were established because all of them were taken into consideration within the parametric model of the wing.

The goal of the optimisation was to find solution having possibly largest values of objectives F1, F2, F3, F4 and which fulfils all assumed constraints. The objectives and constraints were evaluated using cost effective, simplified codes – Fig. 2.

4.2 Solution

The multi-objective genetic algorithm was used to solve defined above optimization problem. Starting from random population of optimized wings, the genetic algorithm generated subsequent generations of wings which were generally better and better fitted to design criteria. A population of each generation was constant and set to 380 individuals. The solution process was stopped after achieving the 180-th generation.

The solution of multi-objective-optimisation problem was the Pareto set, including 2559 Pareto-optimal wings. Projections of Pareto set on 2D spaces of selected pairs of objectives are presented in Fig. 4.

Finally from obtained Pareto set, the optimal solution was selected. The following order of criteria was used:

Priority 1: F3 = L2/DW2
Priority 2: F2 = L1/DW1
Priority 3: F1 = CLmax
Priority 4: F4 = L3/WW3

Taking into account above priorities and all pros and cons, the wing denoted as AC2T-IoA-08 was selected form then Pareto-set as the final version designed AC2-T wing.

4.3 Geometric properties of wing

Geometry of the wing AC2T-IoA-08 fulfils all geometrical constraints and preferences, particularly concerning cost saving oriented classical manufacturability. According to assumed geometrical constraints, the wing ACT2-IoA-08 has fixed trapezoidal planform. The wing AC2T-IoA-08 consists of two segments, corresponding to FZ and AZ zones. The surface of each segment is a ruled surface, defined by its end-sections: RS, MS and TS. These section were obtained by some modifications and combinations of airfoils: ILM115 and ILM111.

Spanwise distributions of sectional: relative thickness, camber of mean line and twist for the wing AC2T-IoA-08 are shown in Fig. 5. The RS, MS and TS sections of the wing, projected on symmetry plane are shown in Fig. 6.

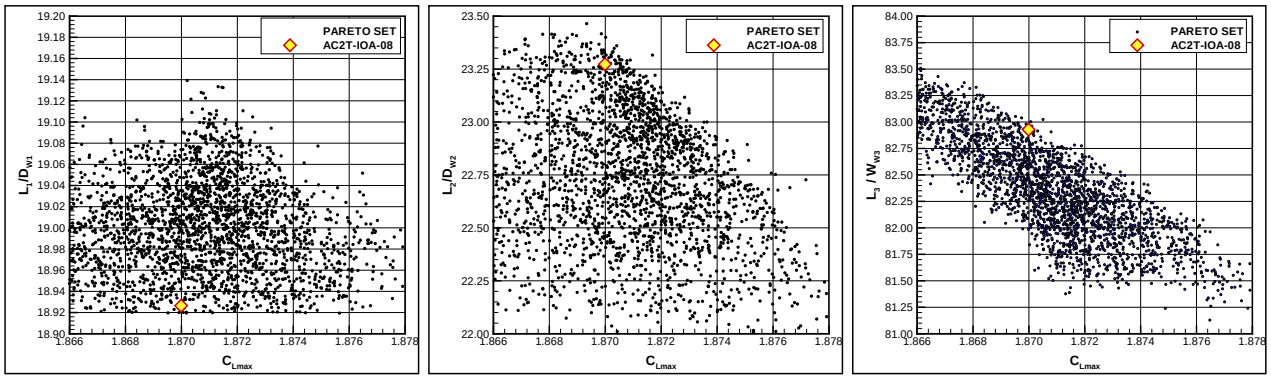


Figure 4 Projection of Pareto Set on F1-F2 (left), F1-F3 (middle) and F1-F4 (right) spaces.

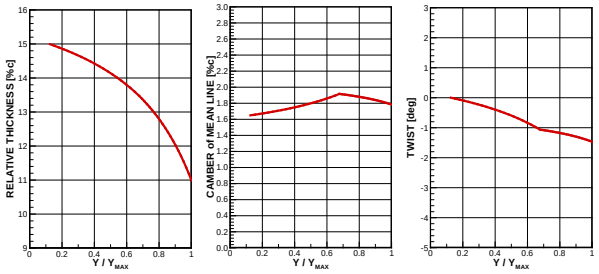


Figure 5 Spanwise distributions of: relative thickness (left side), camber of mean line (middle) and twist (right side) for the AC2T-IoA-08 wing

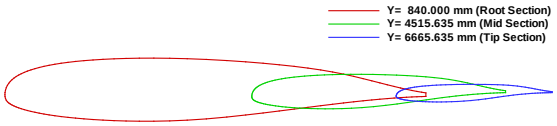


Figure 6 Root, Mid and Tip sections of wing AC2T-IoA-08. Projection on aircraft plane of symmetry.

4.4 Aerodynamic properties of wing

The main goal of optimization was to design possibly low-drag and low-weight wing, having possibly high values of CL_{max}. To check whether the goal was achieved, a wide spectrum of aerodynamic characteristics for the wing AC2T-IoA-08 was calculated.

Cruise Properities (CP)

Ones of most important aerodynamic characteristics concerned CP1 and CP2 for which the objectives (FC1 and FC2 respectively) of optimisation were formulated:

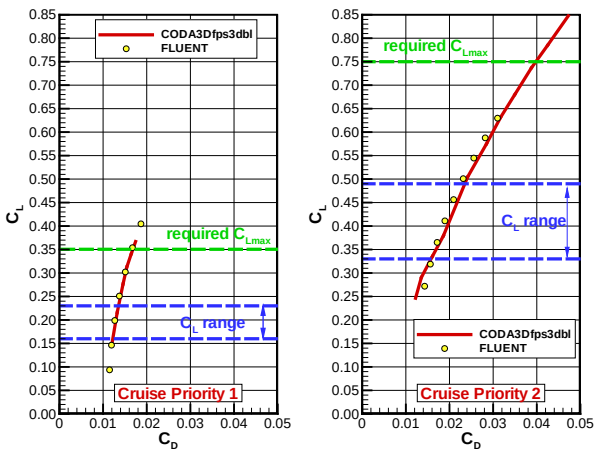


Figure 7 Drag polars for cruise CP1 &CP2. Results of computations performed using FLUENT and CODA3Dfps3dbl codes.

CP1: M=0.716, Re=$10 \cdot 10^6$, $0.16 \le CL \le 0.23$
CP2: M=0.610, Re= $6 \cdot 10^6$, $0.34 \le CL \le 0.50$

For above CP, flow calculations were performed for isolated wing AC2T-IoA-08. Two independent CFD codes were used: *CODA3Dfps3dbl* and *FLUENT* [8] to check whether the simplified code *CODA3Dfps3dbl* (used in optimisation process), gives reliable results in comparison with Navier-Stokes-Equations solver.

Figure 7 shows the drag polars calculated for wing AC2T-IoA-08 for CP1 and CP2. Relatively good agreement can be seen between two, completely different, CFD codes. Concerning calculated values of drag, which has been minimised during optimisation process, it may be noticed, that for turbulent wing the drag level is quite good. Particularly, the drag-divergence phenomenon has not appeared within assumed range of flight conditions.

Figure 7 also shows, that required levels of C_{Lmax} for cruise CP1 and CP2 are achievable taking into account results of CFD computations.

Figure 8 presents the analysis of pitching moment coefficient for CP1 and CP2. Although the differences between two CFD codes are quite big, it is worth to underline that both codes gave results confirming that pitching moment coefficient of the wing AC2T-IoA-08 fulfils defined limitations [6].

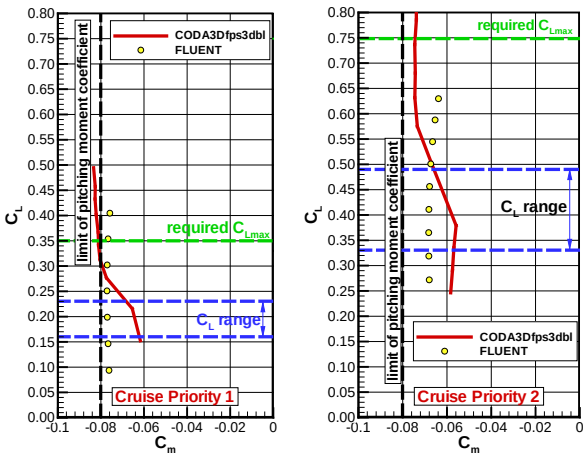


Figure 8 Dependencies CL vs. Cm for CP1 &CP2. Results of computations performed using FLUENT and CODA3Dfps3dbl codes.

Wing aerodynamic properties concerning quality of flow are well illustrated by flow parameter distributions in wing cross-sections. Figures 9 and 10 show distributions of local Mach number in selected wing cross-sections for CP1 and CP2 respectively. Both figures are illustrating results of calculations performed using *CODA3Dfps3dbl* code.

Generally, the presented graphs show good quality of flow, particularly good transonic properties, without shock waves and without drag divergence. Within assumed ranges of lift coefficient limiting cruise flight conditions, there are only weak supersonic regions on the wing surface.

Low Speed Property(LSP)

According to [6], the analysis of maximum lift coefficient was performed for

LSP: M=0.136, Re= $5.7 \cdot 10^6$

Figure 11 shows dependencies CL versus α calculated by different codes at above flight conditions. In the graph there are compared results of program *CODA3Dpanel3dbl* obtained from calculations of flow around complete aircraft AC2 with the wing AC2T-IoA-08 (in this case aerodynamic characteristics of separated wing were appropriately extracted from results concerning complete aircraft) and results of simplified analysis performed using

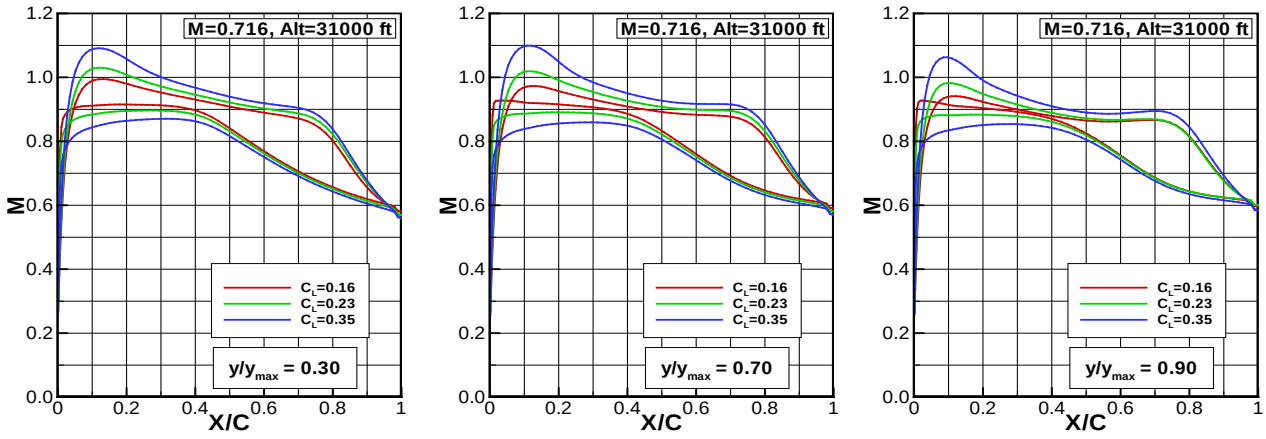


Figure 9 Distribution of local Mach number in selected wing section calculated for CP1.

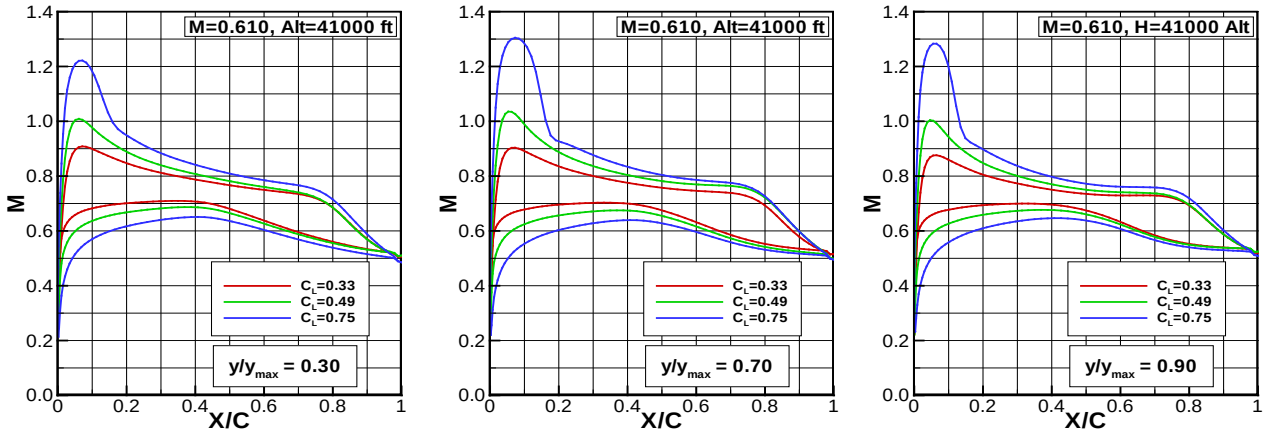


Figure 10 Distribution of local Mach number in selected wing section calculated for CP2.

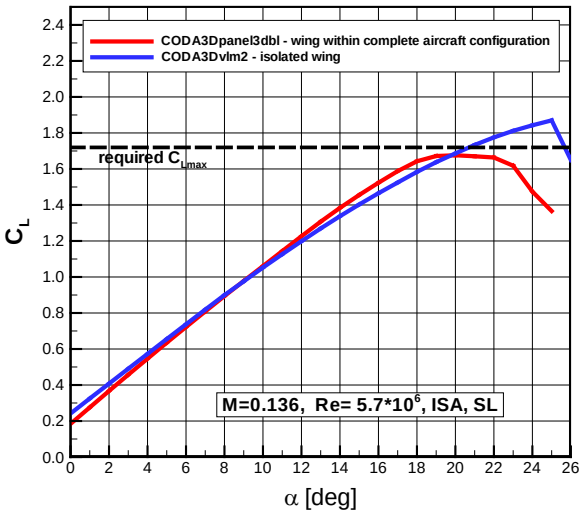


Figure 11 Comparison of curves CL vs. α calculated for the wing AC2T-IOA-08 by codes: CODA3Dpanel3dbl and CODA3Dvlm2.

code *CODA3Dvlm2* for clean wing (this code was used during numerical-optimisation process) Above characteristics are compared with required value of maximal lift coefficient ($CL_{max} \geq 1.72$) according to [6].

Generally it seems that code *CODA3Dvlm2* considerably overestimates CL_{max}. Partially it may result from overestimated (probably) 2D characteristics of wing sections calculated by MSES [9] and used by the *CODA3Dvlm2* program as input data.

The results of code *CODA3Dpanel3dbl* seems to be more accurate. Presented results lead to the conclusion that required value of CL_{max} is rather unreachable by the wing AC2T-IoA-08 in considered flight conditions. The open question is whether the lack of CL_{max} is considerable or slight.

4.5 Design and Optimisation of High Lift System

In the second stage of design process high list system for AC2T-IoA-08 was optimised. The positions and deflections of FF for takeoff and landing configurations were established by Direct Optimisation method.

Starting from flap position and deflection being optimal for 115 airfoil [7], some alternative solutions were taken into consideration by changing appropriate design parameters concerning wing high lift system. For each configuration, calculations of flow around complete aircraft were performed to assess expected value of CLmax. The calculations were performed using *CODA3Dpanel3dbl* code for the following flight conditions:

- ✓ Takeoff (FC5) : M=0.125, $Re=5.2 \cdot 10^6$,
- ✓ Landing (FC6) : M=0.104, $Re=4.4 \cdot 10^6$.

From among considered sets of design parameters defining flap deflection and position, the most promising solutions for both takeoff and landing were chosen and assumed as "*optimal*".

4.6 Geometric properties of high lift configurations

Table 2 shows established optimised values of parameters defining flap positions and deflections for cruise, takeoff and landing configuration.

Table 1. Established values of parameters defining optimised positions and deflections of flap for cruise, takeoff and landing configuration.

	ΔX_1	ΔZ_1	ΔX_2	ΔZ_2	δ_{FL}
Cruise	20.00		20.00	3.22	0.00
Takeoff	7.00	3.90	7.00	3.35	15.00
Landing	1.00	3.00	1.00	2.50	37.50

The limiting cross-sections FiS and FoS of designed high lift system are shown in Fig. 12.

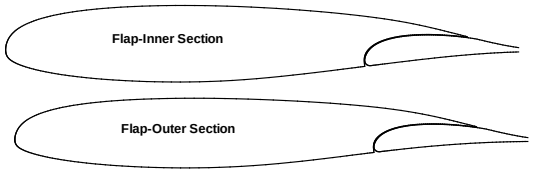


Figure 12. Flap-Inner and Flap-Outer cross-sections of wing AC2T-IoA-08 with high lift system.

4.7 Aerodynamic properties of high lift configurations

CFD calculations for high lift configurations of complete aircraft AC2 with wing AC2T-IOA-08 equipped with Fowler flap, were performed using code *CODA3Dpanel3dbl*,

taking into account following flight conditions [6]:

- ✓ Takeoff: M=0.125, Re=$5.2 \cdot 10^6$
- ✓ Landing: M=0.104, Re=$4.4 \cdot 10^6$

From obtained results, global aerodynamic characteristics concerning separated wing (segment of the wing from RS to TS) with FF were extracted. These characteristics are presented in the Fig. 13.

Table 3 shows basic aerodynamic properties of wing AC2T-IOA-08 with high lift system.

Table 2. Basic aerodynamic properties of wing AC2T-IoA-08 with high lift system.

	Calculated	Requited
Takeoff		
CL_{max}	2.389	≥ 2.1
CL_{max}/1.44	1.652	
CL/CD at CL_{max}/1.44	13.794	
CM at CL_{max}/1.44	-0.085	> -0.2
Landing		
CL_{max}	2.739	≥ 2.8
CL_{max}/1.69	1.620	
CL/CD at CL_{max}/1.69	10.020	
CM at CL_{max}/1.69	-0.246	> -0.4

Summarising presented results, the following conclusions concerning aerodynamic properties of wing AC2T-IoA-08 with Fowler flap may be formulated:

Takeoff configuration

- ✓ Wing lift coefficient CL_{max} fulfils assumed requirement.
- ✓ Wing pitching moment coefficient fulfils the requirement also
- ✓ At CL=CL_{max}/1.44 the lift to drag ratio reaches satisfactory value: C_L/C_D=13.794

Landing configuration:

- ✓ Wing lift coefficient reaches maximal value CL_{max} which is slightly below the requirement. The lack of CL_{max} is relatively small, so maybe further improvement of high lift system would solve this problem.
- ✓ Wing pitching moment coefficient fulfils requirement

5 Conclusion

In the paper the cost efficient methodology of small aircraft/wing design and optimisation has been presented. The methodology was adapted to specific design process of AC2T wing for small aircraft. To improve efficiency of this process, a parametric model of AC2 aircraft/wing has been developed. Within the model both clean wing and wing with high lift system were defined. The Multi-Disciplinary Optimisation technique was applied to design a turbulent wing for AC2 concept of small aircraft. The final result of performed design process is the wing AC2T-IoA-08 without and with the high lift system (FF). For selected fight conditions the aerodynamic characteristics of the wing were obtained. The results of calculation shows, that the wing AC2T-IoA-08

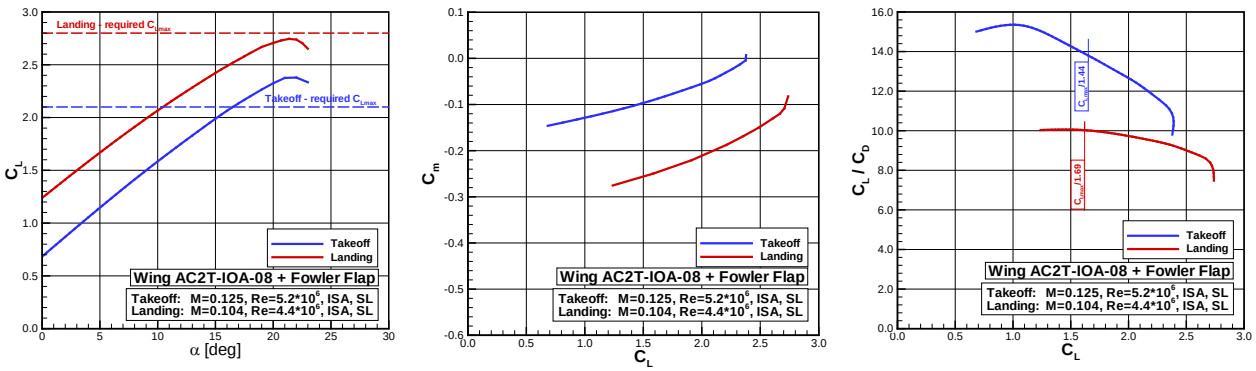


Figure 13 Takeoff and landing configurations of wing AC2T-IOA-08 with Fowler flap aerodynamic characteristics CL vs. α – left side, Cm vs. CL – middle part and CL/CD vs. CL – right side

fulfils most of defined objectives and constraints, and seems to be interesting solution of turbulent wing for high-speed, cost-effective small aircraft.

Nomenclature

CL	lift coefficient
CD	drag coefficient
CM	pitching moment coefficient
CLmax	maximum lift coefficient
M	Mach Number
Re	Reynolds Number

References

[1] "Report on base model definition full aircraft – AC-1 & AC-2",CESAR Project, Deliverable Report No. D1.1.1-1

[2] Stalewski W., Rarata Z., "PARADES – the CAD software for Air Intake and Bypass Parameterisation", IoA Internal Report TN-16/BA/05/M

[3] Stalewski W., "Parametric Modelling of Aero-dynamic Objects - The Key to Successful Design and Optimisation", CEAS-2011

[4] Rokicki J., Stalewski W., Żółtak J., "Multi-disciplinary Optimisation Of Forward-Swept Wing", in Evolutionary Methods for Design, Optimization and Control", eds.: T. Burczynski and J. Périaux, © CIMNE, Barcelona, Spain (2009)

[5] Oyama A. "Multidisciplinary Optimisation of Transonic Wing Design Base on Evolutionary Algorythm Coupled with CFD Solver", ECCOMAS 2000, September, 2000

[6] "Report on wing design objectives AC2, turbulent wing", in CESAR Project, Deliverable Report No. D1.2.2-5

[7] Stalewski W., Żółtak J., Kania W., Wysocki Z., Suchożebrska A., "Airfoil Design and Analysis AC2, Turbulent AFT,HLT", in CESAR Project Deliverable Report D1.2.1-09

[8] Krysztofiak G., Kania W., "Report on wind tunnel tests AC1/AC2, turbulent wing AFT, HLT; Part2 High speed wind tunnel tests of turbulent airfoil for AC2 aircraft wing", CESAR Project Deliverable Report D1.2.1-10

[9] Stalewski W., Żółtak J., "Wing Design And Analysis Report AC2, Turbulent Wing, AFT, HLT", in CESAR Project Deliverable Report D1.2.2-9

[10] FLUENT: *www.ansys.com*

[11] MSES: *www.am-inc.com*

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

POD and Domain Decomposition Hybrid Method for Compressible Aerodynamics Computation

D. Cinquegrana[1], R.S. Donelli[2], D. Quagliarella[2] and A. Viviani[1]
1)Seconda Universitá di Napoli - Dip. Ing. Aerospaziale e Meccanica (DIAM), Aversa (Italy)
2)Centro Italiano Ricerche Aerospaziali (CIRA), Fluid. Lab., Capua (Italy)

Keywords: POD, Domain Decomposition, Optimization, CFD, Transonic Flows.

Abstract

A transonic aerodynamic optimization problem of industrial interest may require several generations of a classical genetic algorithm, with population sizes that may demand tremendous amounts of CPU time, because the evaluation of the objective functions requires the usage of a full-order numerical flow solver. This work presents a different approach aimed to reduce the computational effort. In particular, it relies on a hybrid method to solve, rapidly and efficiently, the aerodynamic flow field while ensuring at the same time the aerodynamics accuracy. A RANS solver is coupled with a Reduced Order Model, based on Proper Orthogonal Decomposition Technique, through a Domain Decomposition algorithm, in order to solve the flow field around an airfoil with a reduced number of Degrees of Freedom (DOFs) and an application of the method to the airfoil shape optimization problem by using a genetic algorithm will be presented.

1 Introduction

Multi-objective aerodynamic shape design is often approached with methods based on evolutionary computing as this class of techniques may offer better flexibility and robustness if compared to other more classical approaches like gradient based optimization techniques. On the other hand, a transonic aerodynamic optimization problem of industrial interest may require several generations of a classical genetic algorithm with population sizes that may demand tremendous amounts of CPU time because the evaluation of the objective functions requires the usage of a full-order numerical flow solver. In these cases it becomes mandatory to find a procedure able to reduce the required CPU-time to make evolutionary computing techniques practically applicable to a wide range of aeronautical problems of industrial interest, even with todays computing power. A common way to overcome this limit, is the introduction of surrogate models in the design loop [1]. These models significantly reduce the CPU time for each objective function evaluation at the cost of an increased approximation error in the objective function evaluation. This work presents a different approach to reduce the computational effort, it relies on a hybrid method to solve, rapidly and efficiently, the aerodynamic flow field around an airfoil while ensuring at the same time the aerodynamics accuracy. The method is here limited to 2D bodies and illustrated through the application to a wing section design problem. The POD is a technique introduced in aeronautics by Lumley [2] to extract coherent structures from turbulent flow. In 1987 Sirovich [3] built a ROM of a dynamical system using POD in order to reduce the degrees of freedom of a system. Since from then, numerous works on ROM/POD are

published in various disciplines of the aeronautic world such as experimental [4], flow control [5], fluid-structure interaction[6] and design[7]. In aerodynamic design, a ROM technique based on POD and a collocation least square over the whole domain for airfoil shape optimization has been investigated by LeGresley et al. [8]. Domain Decomposition [9] technique is widely used since from the 1980s when parallel computing is introduced to numerical methods. Lucia [10] performed the flow analysis around a 2D blunt nose in high-speed conditions by coupling ROM and full-order models, tanks to Domain Decomposition technique. LeGresley [11] developed a more sophisticated approach by applying a Dynamic DD technique. In this case, the domain is divided in several subregions and an error estimation algorithm allows to individuate the subregions where the application of a full-order solver is necessary. Buffoni et al. [12] coupled, iteratively, DD and POD to solve a 2D internal flow. In this work a hybrid method based on DD technique and POD is presented and extended to compressible flow field [13]. A validation of the ROM is showed in the Application section. Furthermore, the Hybrid Method is introduced in a classical optimization loop (see Fig. 1), in place of a FOM model, and an airfoil shape optimization problem is performed with a Genetic Algorithm, in order to assess the method in a design loop. A qualitative comparison of the CPU-time between FOM and ROM in a genetic optimizer is also given, noting that a quantitative analysis will be showed in future works.

2 The Hybrid Method

The aim of this research activity is to develop a computational procedure able to reduce the cpu-time required by a genetic optimization tool and preserving the same level of accuracy of the RANS approach. A Domain Decomposition technique is used to split the whole domain in two subdomains (see Fig. 2): Ω_1 (far field region) and Ω_2 (airfoil region), where different numerical models have been selected to solve the flow fields in these subdomains, with the aim to preserve the accuracy of the aerodynamic solution and, at the same time, save

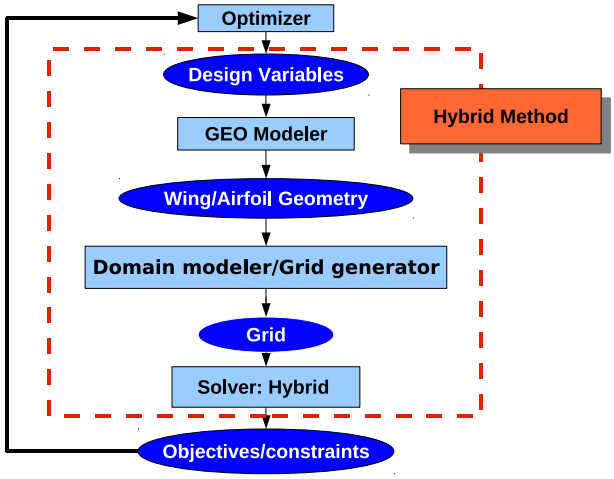


Figure 1: Layout of an Optimization Loop

cpu time. Therefore, a Reduced Order Model (ROM) of the RANS equations, obtained by using the POD technique, is applied in the Ω_1 subdomain. A Full-order model, such as the RANS approach, has been used to solve the flow field in the Ω_2 subdomain ensuring a solution with an high level of accuracy in the region where high non-linearity may occurs, such as strong shock wave and flow separation. A family of airfoils, defined by 10 weighted functions, has been generated and numerically investigated by using a RANS solver on the whole domain. A

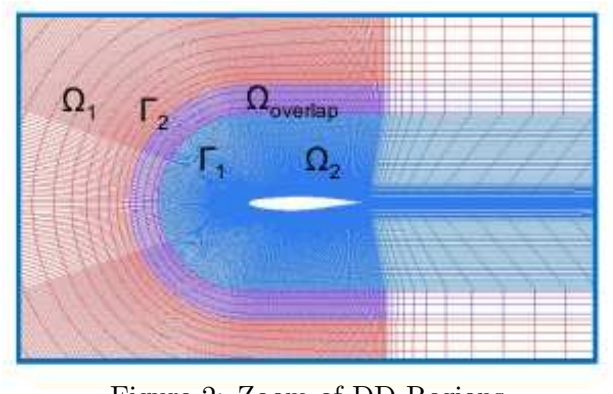


Figure 2: Zoom of DD Regions

POD technique applied to the previous RANS solutions in a partial external subdomain Ω_1, allows to compute the evolution coefficient α_i, and related modes ϕ_i. A kriging technique is then used to set a functional form of the relationship between the evolution coefficients and the design variables defining the generic airfoil

geometry. From this Kriging/POD model, an approximated boundary condition is derived and is imposed on the Ω_2 border line (naming Γ_2 in Fig. 2). The coupling of the POD and full order solver requires special attention because the boundary condition, for the subdomain of the full-order solver, will be an approximation generated by the POD solver and, in general will not correspond to a physical solution. To match the solution of the coupled POD and full-order solver, an overlap region Ω_{ov}, is considered and a Dirichlet-Dirichlet iteration, in the frame of a classical Schwarz method, is applied. The flowchart of the algorithm describing this procedure is shown in Fig. 3. The procedure foreseen the following step:

1. Solve the problem in the Ω_2 region with the RANS model, imposing the Dirichlet bcs on Γ_2;

2. Project the solution in Ω_{ov} on the space of the traces of POD modes ϕ_i;

3. Reconstruct the flow field in Ω_1 with the POD mode;

4. Impose the new bcs on Γ_2 obtained by the trace of the reconstructed flow field in (3);

5. Go to 1, until convergence is attained.

More details on the step 2 will be provided in the following paragraph.

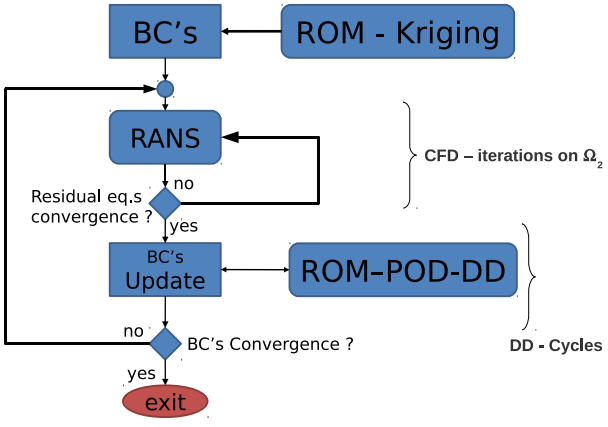


Figure 3: Hybrid Method FlowChart

3 The Numerical Models

In this section, a short description of the techniques used in the procedure are presented. First, the sampling process of the design variable space to conduct numerical simulations that forms the dataset, is achieved by using the Design Of Experiment. Then, a short overview of the POD technique will be provided. The Kriging technique, used to relate design variables and evolution POD coefficients, will be briefly examined. Finally, the Domain Decomposition technique and the main features of the full RANS equations solver used in this work will be described.

3.1 DataBase Design

In order to obtain a reliable ROM, it is mandatory to build a Database of simulations with a sampling of the design variable. Since the CPU-time required of each single RANS simulation is relevant, a Full Factorial design is not applicable for a large number of design variables. Several techniques to reduce the number of sample points are based on the fractional factorial design [14]. These methods ensure the uniformity property, for example, by maximizing the minimum distance among the design points. Here, the design space is constituted by ten geometric variables defining the parametric model of the airfoil. By changing the geometry of a reference airfoil, the design space is uniformly filled by using an Optimum Latin Hypercube that guarantees the uniformity property. This approach has been preferred since gives the chance to select the number of samples. In order to obtain a reliable kriging surface, the number of airfoils has to range between eight and ten times the number of design variables. In this study, a main DataBase (DB_1) consisting of 85 airfoils have been generated by maximizing the minimum distance between the sample points. Two additional datasets with 20 airfoil configurations are also built. One of them, DB_2, is used to build a second ROM model in order to test the size of a DB on the Hybrid Method results. The last one, DB_t, is built to conduct a validation analysis of the two ROM models.

3.2 Proper Orthogonal Decomposition (POD)

In this paper, the Snapshots procedure of Sirovich [3] has been used. The POD methodology allows to have the best representation, in the average sense, of a given data set, $\mathbf{u}^k$, where k is the number of the experiments or numerical simulations. The representation is achieved through a linear approximation of $\mathbf{u}^k$ in a new space defined by a set of optimal basis function, ϕ:

$$u(x,s) = \bar{u} + \sum_{i=1}^{N} \alpha_i(s)\phi_i(x) \tag{1}$$

where $\bar{u}$ is a base solution [10, 15] (i.e. an average), and the scalar α_i are the unknown evolution coefficients (to be defined). The modes ϕ_i are optimal in the sense that are the best approximation, in average, of the entire dataset. The mathematical formulation of this optimization problem is:

$$\max_{\phi} \frac{\langle \|(u,\phi)\|^2 \rangle}{\|\phi\|^2} \tag{2}$$

where $\|\cdot\|$ denotes the modulus and $\|\cdot\|$ the $L^2 - norm$. The problem in Eq.(2) is a calculus of variations problem constrained by $\|\phi\| = 1$. The Sirovich model is particularly useful when a large number of grid points is used, as in these numerical simulations. A modified correlation matrix [3] is given by:

$$R_{ij} = \frac{1}{M} \int_{\Omega} u_i u_j d\Omega \tag{3}$$

where u_i is the $i-th$ snapshot and $i,j = 1,2,\ldots,M$, and M is the total number of snapshots. R is a non-negative definite symmetric matrix and therefore has non-negative eigenvalues and a full set of orthogonal eigenvectors. The eigenvectors of R are then computed solving

$$R\alpha = \lambda\alpha \tag{4}$$

where α_i^k is the $i-th$ element of the eigenvector corresponding to the eigenvalue λ_k. The POD basis function can be calculated projecting the α coefficient on the snapshots of the dataset:

$$\phi^k = \sum_{i-1}^{M} \alpha_i^k [u(x,y) - \bar{u}] \quad K = 1,2,\ldots,M \tag{5}$$

These POD modes are orthogonal and have been normalized. The eigenvalue λ_k is a measure of the related information content associated with α_i. The total information content in the dataset is evaluated by:

$$K = \frac{1}{2} u \cdot u = \frac{1}{2} \sum_i \lambda_i \tag{6}$$

It is possible to define a threshold to truncate the series in Eq.(1) based on the total energy content, in order to reduce the degrees of freedom of the system. This information is available from the following equation:

$$E_j = \frac{\sum_j^p \lambda_j}{\sum_i^M \lambda_i} \tag{7}$$

where $p < M$. If $p = M$ then E is equal to unity. Here, the variables that define the vector states $\mathbf{u}$ are: the velocity components of a 2D field; static pressure; and static temperature, $\mathbf{u} = (u, v, p, T)$. These variables are scaled in order to give a balanced weight for each variable in the correlation matrix computation of Eq. (3).

3.3 Kriging

The ROM, has a dual purpose:

1. to generate the Dirichlet boundary condition on Γ_2 line in order to start and speed up the full-order solver;

2. to update the boundary condition in order to match the solution of the coupled POD and full order solver, as shown in Fig. 2.

Point (1) is achieved by using a Kriging technique that allows to find a relationship between the evolution coefficient α and the design variables. In the frame of computer experiments, the kriging technique has been widely adopted since has the advantage of providing smooth interpolating functions [14]. The response of a deterministic function $Y(x)$ can be modeled as:

$$Y(x) = f(x) + Z(x) \tag{8}$$

where $f(x)$ is a regression model, such as polynomials, and $Z(x)$ is a random function (stochastic

process) that models deviations from the regression model. The random process $Z(x)$ has zero average and covariance given by:

$$E[z(w), z(x)] = \sigma^2 C(\theta, w, x) \tag{9}$$

where σ^2 is the process variance of the response and $C(\theta, w, x)$ is the correlation model with parameter θ which is a measure of the degree of the correlation among the data. A quadratic polynomial model has been selected for the regression model of $f(x)$. A Gaussian correlation has been chosen for the random process $Z(x)$. The accuracy of the kriging technique to predict a boundary condition close to the real value has been estimated by using the following $R^2 - metric$:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} = 1 - \frac{MSE}{Variance} \tag{10}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value and $\bar{y}$ is the average of the actual values. The Mean Square Error (MSE) measures the difference between the estimated and the actual value.

3.4 Domain Decomposition

The Domain Decomposition has been used to split the spatial domain of the problem in two or more blokes (sub-domains). However, the coupling does require special attention to the interface between the approximated flow solution generated via POD (that in this work is used to update the boundary condition) and the full order solver. This section focus on the treatment of the transmission condition between Ω_1 and Ω_2. The transmission condition is approached by using the projecting method. In order to illustrate this method, it is necessary to define φ as the trace of the POD mode ϕ on the overlap region Ω_{ov}. The modes restricted to this region are not orthogonal [15], therefore, to obtain the POD evolution coefficient α it s necessary to define a minimization problem in Ω_{ov}, such as:

$$\alpha = min_{\beta \in \Re^M} \|u_\Omega - (\bar{u} + \sum_{k-1}^{M} \beta_k \varphi_k)\|^2_{\Omega_{ov}} \tag{11}$$

where $\|\cdot\|$ is the norm induced by the canonical L^2 scalar product. The solution of the minimization problem in Eq. (11) is approached by the Least Square method, resulting in the linear problem:

$$\sum_{i=1}^{M} a(\varphi_i, \varphi_j)_{\Omega_{ov}} = (u_{\Omega_{ov}}, \varphi_j)_{\Omega_{ov}} \quad 1 \leq j \leq M \tag{12}$$

Once obtained the RANS solution in the Ω_{ov} region, the optimal set of POD coefficients are computed solving the linear system of equation $P u_{\Omega_{ov}} = \alpha$. P is a Moore-Penrose pseudo-inverse matrix, defined by the POD modes and computed once:

$$P_i = \sum_{j=1}^{M} [(\varphi_i, \varphi_j)_\Omega]^{-1} \varphi_j \tag{13}$$

At this point, the new boundary condition for the RANS solver through the POD expansion equation (1) is obtained. However, care should be taken over P_i and its inversion [16].

3.5 The Flow Solver

The full order model (FOM) is a steady RANS solver that computes the flow solutions on the dataset of airfoils, from which POD/ROM model is derived. The flow solver used for the numerical simulations is the $FLUENT^{TM}$ commercial code. It solves the Reynolds Averaged Navier-Stokes equations on unstructured hybrid grids by means of the Finite Volume Method. A Density Based algorithm is used to solve the compressible flows over the airfoils, where the governing equations of continuity, momentum and energy are solved simultaneously as a set of equation, with a coupled implicit formulation. The ROE flux difference splitting was choosen in order to evaluate the convective flux of the NS equations. At the end, the steady flow fiels was obtained with an implicit pseudo time-marching algorithm. The turbulence model used in the numerical simulation is the SST $k-\omega$ model, with compressible correction.

3.5.1 Computational Grid

A structured C-grid type was used as shown in Fig. 2 . 203 grid points were placed in the

streamwise direction and 130 grid points in the transverse direction, with a first raw spacing ensuring the $y^+ = 1$ requirement. The Ω_1 grid is fixed while the internal one is generated accordingly to the changes in the airfoil geometry. A tool for fusing mesh is used in order to match the two subdomains.

3.5.2 Boundary Conditions

The freestream values for the calculation on the whole domain are fixed and reported in Table 1. In the DD procedure, a velocity Inlet profile is imposed on Γ_2 to start the RANS analysis, in the Ω_2 sub-region.

Table 1: Farfield boundary Condition

Freestream Parameters		
Reynolds	Mach	AoA[*deg*]
$6.5M$	0.73	2

4 Applications and Results

A main database (DB_1) with a population of 85 airfoils have been generated by modifying the RAE 2822 airfoil geometry. These airfoils have been analyzed in a steady compressible flow with a FOM solver with fixed freestream values, showed in Table 1. From this dataset of flowfields, a ROM model was built as described, to solve new flowfields in region Ω_1. The accuracy and efficiency of the Hybrid Model have been validated by performing a flow analisys using the FOM and ROM approaches, on a set of airfoil geometries, named DB_t, not included in DB_1. Furthermore, the HM has been included in an Evolutionary Algorithm based Design Problem, to validate it in a shape design frame. A summary of the steps followed to obtain the results are here reported:

1. RANS solver solutions on whole domain applied to DB_t elements.

2. Hybrid Method on reduced domain Ω_2:

   - ROM of DB_1 tested on DB_t configurations

3. Hybrid Method applied to an Airfoil Shape Design problem

The Hybrid simulations have been stopped when a target value of residual level on the continuity equation has been reached. The level of energy is used as parameter in the truncation of series in Eq. (1): the number of modes retained in ROMs is as much as necessary to reach the 99.5% of E (Eq. 7).

4.1 Validation of DB_1-ROM

This analisys is performed on a new set of airfoils (DB_t), not included in the DB_1. The first results here showed are related to the compuational effort of this approach in terms of CFD iterations. On the left side of Fig. 4 is showed the total amount of CFD iterations needed to:

- achieve the requested level of residual on the continuity equation, starting the RANS solver with the ROM-Krig bc's

- iterate with DD procedure for a fixed number of cycle

In this frame of the research, the DD-cycles number is fixed and based on previous experiences. On the right side of Fig. 4, a cfd-iteration DD-cycles sensitivity is showed: increasing in DD-cycles, a less number of CFD iteration is required in order to achieve the target residual level.

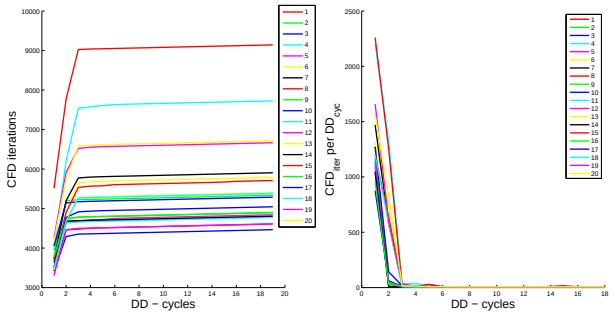


Figure 4: Computational Effort of DD- cycles

In order to evaluate the efficiency of the method, the CPU-time saved are showed. On the histograms of Fig. 5 are showed the cputime saved for each configuration, estimated as:

$$CPU_i[\%] = (1 - \frac{Tot_{It}^{Hyb}}{Tot_{it}^{cfd}} \frac{S_{it}^{hyb}}{S_{it}^{cfd}}) \qquad (14)$$

where S_{it}^{hyb} and S_{it}^{cfd} are the solver cpu-time per iteration. In this case, with the solver settings described in a previous section, the ratio is equal to 0.55, thanks to the reduced DOFs related to the reduced extension of Ω_2. The variable in Eq. (14) Tot_{It}^{Hyb} is equal to $Tot_{It}^{O^{th}DD} + Tot_{It}^{DD}$. Here, $Tot_{It}^{O^{th}DD}$ is the number of iterations to reach the convergence level of the continuity equation (see Fig. 2 and 4). Fig. 5 shows that, in average, the CPU-time saved is the 50%, except for the configuration n. 8, where, as highlighted in Fig. 4, an increased number of CFD iterations has occured to reach the fixed convergence level.

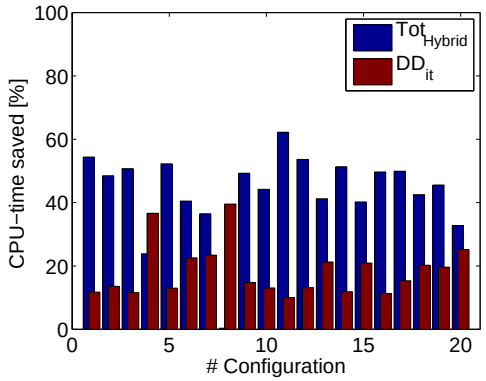


Figure 5: CPU-time saved and CPU of DD-cycle

The accuracy on the aerodynamic coefficients has been evaluated. In particular, the drag coefficient for the DB_t airfoils has been estimated by using the FOM and the Hybrid Method. The picture in the left side of in Fig. 6 has been achieved by building the ROM model using the DB_1, while the picture on the right side is achieved by building the ROM model using the DB_2 dataset.

The black line is obtained when no discrepancies between FOM and Hybrid method exists, i.e. maximum accuracy. The blue circles are the drag coeff. values achieved without using the DD cycles, while the use of this technique provide the red points. It can be noted that the application of this technique strongly improve the solution accuracy. The comparison between the left and right pictures in Fig. 6 shows that the accuracy on the drag coefficient for the two ROM built on DB_1 and DB_2, is comparable, although the second starts the DD-cycles

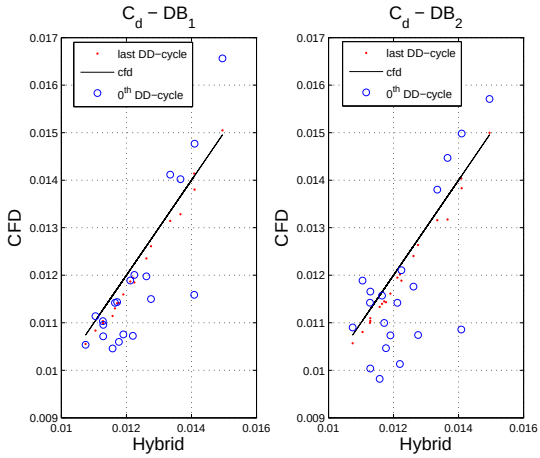


Figure 6: Drag coeff. at end of DD cycles

from worst results in terms of aerodynamic coefficients. Table 2 provides quantitative data on the accuracy of the aerodynamic coefficients in terms of R^2 (Eq. 10) and $\Delta\bar{C}_d$ and $\Delta\bar{C}_l$ (averaged over the 20 configurations of DB_t). The $\Delta\bar{C}_d$ is defined as:

$$\Delta\bar{c}_d = \frac{1}{N}\sum_{i}^{N}(\|\frac{c_{d-FOM(i)} - c_{d-Hybrid(i)}}{c_{d-FOM(i)}}\|) \tag{15}$$

Table 2: Validation Results Summary

	ROM(DB_1)		ROM(DB_2)	
DD-cyc.	0th	20th	0th	20th
$\Delta\bar{C}_l$	0.0027	0.0020	0.0039	0.0024
$\Delta\bar{C}_d$	0.0594	0.0217	0.0787	0.0189
$R^2(C_l)$	0.9991	0.9995	0.9984	0.9995
$R^2(C_d)$	0.4398	0.9541	0.2230	0.9515

It could be noted that the application of the DD technique reduces the Δc_d from 6% to 2% at least.

4.2 Optimization

A simple airfoil design optimization problem is used to test the above descripted hybrid flow-field evaluation method. A genetic algorithm is used to reduce the drag of a given airfoil. No constraints are imposed, with the exception of airfoil thickness, which is scaled to a given value

after the airfoil has been changed using the previously described airfoil modification method. This simple problem is only significant as a design loop validation test. Indeed, a realistic problem would require, at least, a constraint on lift and pitching moment.

Table 3: Optimization Settings

Optimization Parameters	
Obj.function	$\min\{C_d\}$
Constraint	$t_{max} = 12.1\%$
Design Point	See Tab.1
Generation	6
Population size	12
Initial Pop.	*Random*
Crossover	100%
Bit-Mutation	2%

The design variables and the design point are the same that have been used for the Database of FOM simulations (see Table 1). It is worth to note that the number of generation evaluated is small, but it is enough for this simple validation problem. The other optimization parameters are summarized in Table (3). For a detailed description of the genetic algoritm see [17]. A total of 82 configurations was evaluated. The evolution of the objective function is reported in Fig. 7, where a descending trend is appreciable. In particular, the element n. 79 is characterized by the minimum drag.

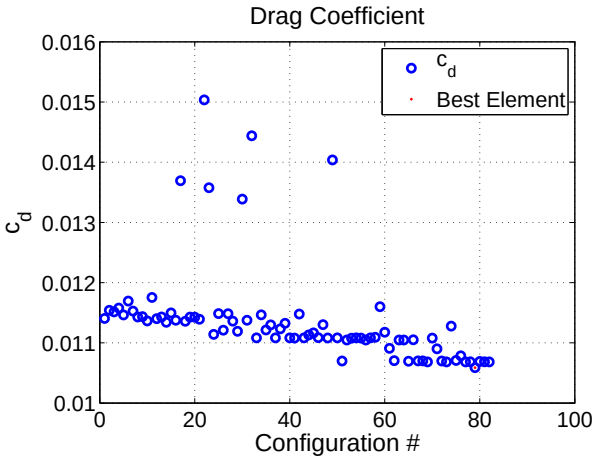


Figure 7: Drag Coeff. evolution

Figure 8 illustrates the evolution of the design variables in the design space. The variable b_4 and b_6 are those characterized by a large change, because they are related to the camber change, yhat is the most significant change in the airfoil shape. Hence, except b_4 and b_6, the weights of the final generation were located around a narrow range of values, upper limited by 0.1. In

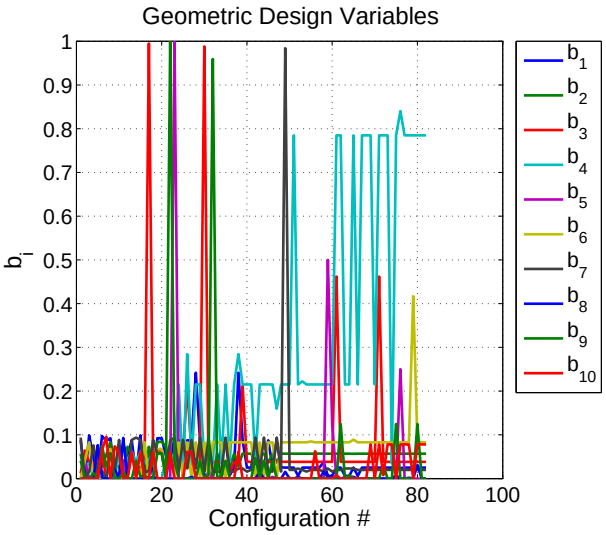


Figure 8: Geometric Design variable evolution

Fig. 9 is illustrated the convergence history of c_d in the Hybrid calculations. Finally, in Fig. 10, the best airfoil shape is compared with the base airfoil shape, the RAE2822. In the same picture, the pressure coefficient distributions for the two airfoils are also showed. It is clear that, lowering the curvature of the base airfoil, the optimizer finds a shock-free solution, at a quite lower value of c_l. In Table 4 the final results

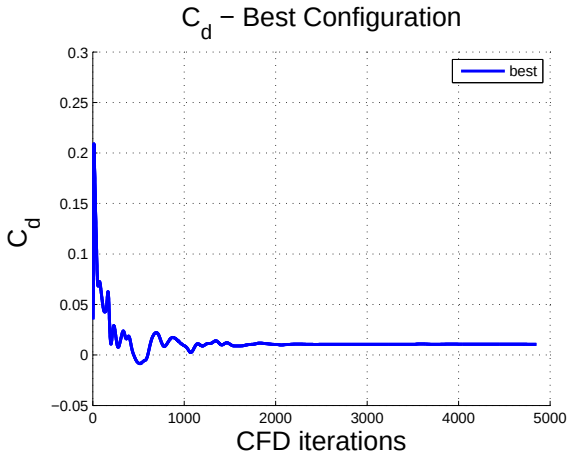


Figure 9: c_d conv. History for best Config.

in terms of c_d are showed for the best airfoil configuration, and compared with a FOM sim-

ulation. The relevant point of this validation

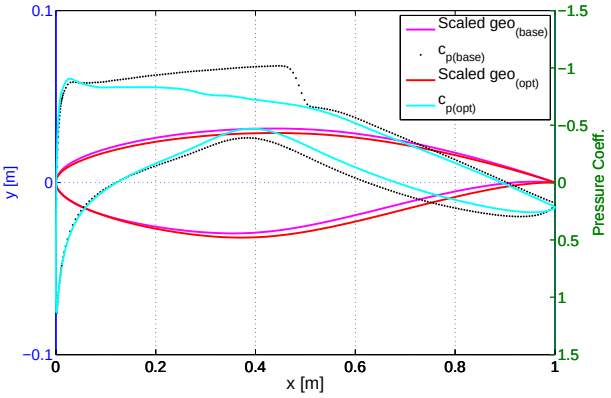


Figure 10: C_p and shapes: base v.s. best(ID 79)

exercise is about the computational effort of the hybrid method in the context of a design loop. As described in the precedent subsection, the cpu-time saved is evaluated as a percentage of the FOM-Full Domain solutions, see Eq. (14), for all the configuration evaluated in this design loop, and showed in Fig. 11.

Table 4: Optimization Results and Comparison

	Base Airfoil	Optim. Airfoil
ROM_{C_d}	0.01153	0.01059
FOM_{C_d}	0.01229	0.01100
ΔC_d	6.1%	3.8%

From this Figure it is possible to note a light but steady trend to the decrease of the CPU-time: this trend is probably related to the shock-free flow solutions, typical of the last generations. In average, the CPU-Time saved for this model, is the 50%. With the aid of the following formula, it's possible to find a *break-even* point:

$$Nc + kcM = cM \tag{16}$$

where N is the evaluation of the objective function evaluated with the FOM, c is a normalized time per simulation, and M is the objective function evaluated with the ROM. The left term is the summation of the time needed to build the DB of the ROM model plus the time of M evaluations of the objective function made with the ROM in the Design loop. The right term is the time needed for M evaluations made using FOM in the optimizer. Figure 12 clarifies the break-even concept. Assuming a normalized value for time, $k = 0.5$ and $c = 1$, with a DB of $N_1 = 85$, the number of simulation M_1 is equal to 170: this number of simulations with ROM are needed to reach the same number of evaluation in a design loop with FOM.

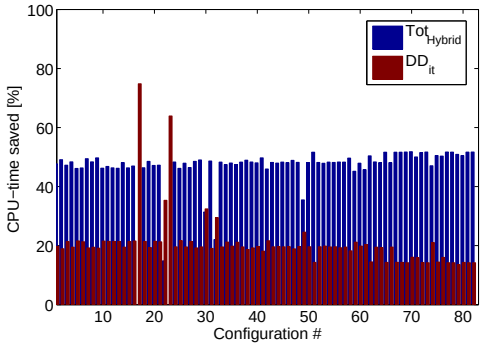


Figure 11: CPU-saved per ements

In order to anticipate the breakeven point, two strategies are possible: to enhance the Hibrid Method efficiency, in order to obtain a lower value of k, or to evaluate a reduced number of configurations in the DB. In the previous subsection, the validation of the ROM based on DB_2 shows that, thanks to the DD-cycle, it is possible to reach the same level of accuracy of the larger DB_1. Indeed the algorithm of the DD-cycles appears to compensate the low quality of the starting reduced subspace, obtained from a number of samples quite low for a given parameter space size. If $N_2 = 20$, M_2 is equal to 40, hence, the breakeven point is reached before, and the gain in CPU-time is more consistent if the number of evaluation increment. In a next work the results of a design loop with a ROM based on DB_2 will be presented, in order to compare the accuracy in terms of the best solution and CPU-time.

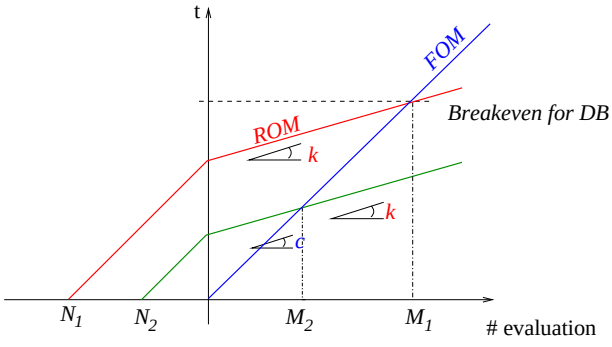


Figure 12: Breakeven point estimation

5 Conclusion

Due to the simple design problem, where constraint on lift and moment coefficient were not considered, this work has to be considered only a preliminary discussion, focused on the feasibility of a design loop with the Hybrid Method with additional, but qualitative comments on the computational feasibility. The design problem has to be further enhanced in future works with engineering consistent desing problems, also together with a quantitative validation of the model from the computational point of view. Clearly, some improvement can be achieved working on various aspect of the Hybrid Method too: the number of FOM simulation in the DB, the number of DD-cycles and so on. These efforts can further improve the efficiency of the method here presented.

Aknowledgment

A special thanks to Emiliano Iuliano (CIRA Fluid. Lab.) for the useful discussions and suggestions.

References

[1] L.Yongsheng, Akira, O., and Meng-Sing, L., "Progress in design optimization using evolutionary algorithms for aerodynamic problems," *Progress in Aerospace Sciences*, Vol. 46, 2010, pp. 199–223.

[2] Lumley, L., "The structure of inhomogeneous turbulent flows," *Athmospheric turbulence and radio wave propagation*, 1967, pp. 166–178.

[3] Sirovich, L., "Turbulence and the dynamics of coherent structures," *Q. Appl.Math*, Vol. 45, 1987, pp. 561–590.

[4] Mulleners, K., Hemireg, A., and Raffel, M., "Investigations of a Trailing Edge stall on 2D airfoils," *14th Int. Symp. on Applic. of Laser Tech. to Fluid Mech*, 2008.

[5] Ravindran, S., "Proper Orthogonal Decomposition in Optimal Controls of fluids," Tech. Rep. TM-1999-209113, NASA, 1999.

[6] Coelho, R., Breitkopf, P., and Knopf-Leinor, C., "Model reduction for multidisciplinary optimization - Application to a 2D wing," *Struct. Multidisc. Optim.*, Vol. 37, 2006, pp. 29–48.

[7] Xiao, M., Breitkopf, P., Coelho, R., Knopf-Leinor, C., Sidorkiewicz, M., and Villon, P., "Model reduction by CPOD and Kriging," *Struct. Multidisc. Optim*, 2010.

[8] LeGresley, P. and Alonso, J., "Airfoil Design Optimization Using Reduced Order Models Based on Proper Orthogonal Decomposition," *Fluid 2000 Conference & Exhibit*, June 19-22 2000.

[9] Quarteroni, A. and Valli, *Domain Decomposition Methods for Partial Derivative Equations*, Oxford Press, 1999.

[10] Lucia, D., King, P., and Beran, P., "Reduced order Modeling of a two dimensional flow with moving shocks," *Computers & Fluids*, Vol. 32, 2003, pp. 917–938.

[11] LeGresley, P. and Alonso, J., "Dynamic Domain Decomposition and Error Correction for Reduced Order Models," *41st AIAA Aerospace Science Meeting & Exhibit*, Reno(NV), USA, January 6-9 2003.

[12] Buffoni, M., Telib, H., and Iollo, A., "Iterative methods for model reduction by domain decomposition," *Computers & Fluids*, Vol. 38, 2009, pp. 1160–1167.

[13] Cinquegrana, D., Donelli, R., and A.Viviani, "A Hybrid Method Based on POD and Domain Decomposition to compute the 2-D Aerodynamics Flow Field," *Congresso dell'Associazione Italiana di Meccanica Teorica e Applicata*, Bologna, Settembre 2011.

[14] Queipo, N., Haftka, R., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P., "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, Vol. 41, 2005, pp. 1–28.

[15] Berkooz, G., Holmes, P., and Lumley, J., "The Proper Orthogonal Decomposition in The analisys of turbulent Flows," *Ann. Rev. Fluid Mech*, Vol. 25, 1993, pp. 539–575.

[16] Everson, R. and Sirovich, L., "Karunen-Love procedure for gappy data," *J. Opt. Soc. Am. A.*, Vol. 12, No. 8, August 1995.

[17] Quagliarella, D. and Vicini, A., "GAs for aerodynamic shape design II: multiobjective optimization and multicriteria design," *Genetic Algorithms for optimization in aeronautics and turbomachinery*, Lecture series 2000-07, von Karman Institute for Fluid Dynamics, May 15-19, 2000 2000.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Interpolation Method for Coupling CFD/CAA for Industrial Applications

F. Cariglino, R. Arina, A. Iob
Department of Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino.

N. Ceresola
Alenia Aeronautica, Corso Marche 41, Torino.

Keywords: *Aeroacoustic, CFD/CAA coupling, Turbulence, Detached-Eddy Simulation.*

1 Abstract

This paper presents an efficient Computational Fluid Dynamics (CFD)/Computational Aero-Acoustic (CAA) coupling strategy for industrial applications, that enables the reduction of the overall computational cost. The proposed method relies on coupling an unsteady CFD solver in the acoustic near field with an integral solver accounting for sound propagation in the far field, using an automatic interpolation of the CFD data on the interface between the flow field and the acoustic far field region.

This procedure is described in the paper and some results of the numerical simulations of the non-punctual monopole source and of a three-elements high-lift wing profile are presented.

2 Introduction

The aerodynamic noise of airplanes at take-off and landing is a critical factor in the future development of aviation. In fact, following the demand of more and more stringent international regulations on aircraft noise in the vicinity of airfields, major improvements need to be achieved. This is one the objectives of the Clean Sky - Green Regional Aircraft initiative launched by the European Community, aimed at demonstrating and validating the technological breakthroughs for a new future generation "green" aircraft. The main target is to reduce the aerodynamic noise by 4 dB. This can only be achieved through the application of reliable, advanced simulation techniques for noise generation and propagation.

Our purpose is to compute the aerodynamic flow field around a body and to predict the far field noise generated by this flow. Several authors [1, 2] have proposed a coupling technique based on the interpolation method, which permits to take out the informations concerning the acoustic sources, that will be used to compute the far field noise. The numerical simulation of the complete problem, including the simulation of the local noise sources and the far field noise radiation faces the difficulty of handling fluctuations covering a very extended range of

length scales and amplitudes. Hybrid methods typically combine CFD techniques to simulate the near field flow which contains the aerodynamic noise sources, and acoustic numerical techniques for the prediction of the far field noise. The idea is to split the physical space into several domains, in which specific physical mechanism are simulated using the most adequate set of equations with the most economic discretization strategy. For these reasons a coupling strategy become an essential issue.

The present paper describes a more straightforward technique based on the coupling of the Navier-Stokes equations, for the aerodynamic generation of sound, and the Linearized Euler equations, for the acoustic propagation in a non-uniform medium. If the region where acoustic propagation occurs is characterized by a uniform steady mean flow, the Linearized Euler equations reduce to the scalar wave equation which can be efficiently solved using the integral formulation of Ffowcs Williams-Hawkings [3]. The assumption of uniform flow in the far field is always valid, and it enables to solve an integral problem with a gain in effiency.

This approach involves the interpolation operator, used to perform the CFD/CAA coupling. It is based on the finite element interpolation theory.

To assess this method, two applications are presented: the first application is the simulation of the sound generated by the non-punctual monopole source and the latter is the simulation of the aerodynamic noise of a three element high-lift wing profile.

The first application allows us to assess the method by comparing the results to a reference test case, for which the analytic solution is known, that is the acoustic monopole generated by a pulsating sphere. The noise source is due to the wall movement, therefore the flow can be computed using Euler equations.

The second application is on a more complex test case. The flow field is computed using the unsteady Navier-Stokes equations, with a Detached Eddy Simulation (DES) [4].

Moreover, a detailed analysis of the influence of the interpolation method, for the reconstruction of the solution on the integration surface, is presented.

3 Hybrid method

In this work an hybrid method is presented. The flow field computation is done using the compressible flow solver UNS3D, developed by Alenia Aeronautica and the sound field computation is done using the Ffowcs-Williams and Hawkings analogy, implemented in the aeroacoustic code FWH, developed by the Politecnico di Torino [5].

3.1 The code UNS3D

The computations have been performed using the code UNS3D. The solution algorithm is based on a finite volume, node centred approach operating on an hybrid unstructured grid. The artificial dissipation model is derived from the nonlinear scheme of Jameson, with no eigenvalue blending. Scalar or matrix dissipation can be chosen.

The Navier-Stokes equations are integrated in time with a second order backward difference scheme and dual time stepping. A five stage Runge-Kutta scheme is used to drive toward zero the residual at each time step. With the use of residual averaging, a local CFL number of 4.9 could be employed in the multistage sub iteration process.

The Weiss and Smith version of low Mach number preconditioning is implemented in the code. A sensor depending on cell Reynolds number was also introduced to avoid applying the preconditioning inside boundary layers. For the computation of the present test case, its application was found to be beneficial in order to reduce numerical dissipation and enhance convergence in low-Mach number pockets.

Matrix dissipation was also found to be beneficial, allowing a strong reduction of the dissipation associated with convective eigenvalues well below the SGS eddy viscosity

levels, hence enabling a better resolution of vortices.

3.1.1 URANS turbulence model

A k-ω turbulence model was employed, that has been developed by Hellsten [6]. The model constants have been calibrated requiring consistent behaviour near boundaries between turbulent and laminar flow, inside shear flows and for zero pressure gradient wall flows. In particular, the calibration have been considered taking into account a variable c_μ, as it is the case if an algebraic stress model (EARSM) is included.

The Wallin-Johansson Explicit Algebraic Stress Mode (WJ-EARSM) [7] is implemented using Hellsten's k-ω as the basis RANS model. The model is an exact solution of the corresponding ARSM in two-dimensional mean flow. In three dimensions there still is a complete, while approximate, solution.

The full anisotropic version of the model is used, i.e. the anisotropic part of the Reynolds stress tensor is directly introduced in the momentum equations, while the isotropic part is taken into account in the form of an effective variable c_μ.

3.1.2 DES turbulence model

The DES methodology essentially consist into using k-ω-EARSM as a subgrid scale model for LES in the regions where the grid resolution is sufficiently fine [4], while URANS is employed in the rest of the domain. The switch between URANS and LES zones is determined using the so-called DES length scale, defined to be the minimum between the characteristic URANS length scale and a length related to the local grid, times an appropriate scaling coefficient:

$$l_{DES} = min\{l_{RANS}, C_{DES}\Delta\} \qquad (1)$$

where

$$l_{RANS} = \frac{\sqrt{k}}{C_\mu \omega} \qquad (2)$$

$$l_{DES} = C_{DES}\Delta \qquad (3)$$

and Δ can be the maximum distance between a cell vertex and the surrounding ones, $\Delta_i = max_j\{l_{ij}\}$. In the case where very high aspect ratio cells are present in the LES region, it may be better to take the square root of the maximum surface between the cell faces, $\Delta_i = \sqrt{max_j\{S_{ij}\}}$. C_{DES} is a constant, depending on the RANS model used and on its particular numerical implementation. In the present case, C_{DES} is derived analytically imposing identity between the eddy viscosities given by the present SGS model and the Smagorinsky model in equilibrium $\left(P^k = D^k, P^\omega = D^\omega\right)$ conditions. It is given by

$$C_{DES} = \frac{C_{SMAG}\left\{F\left(\frac{\gamma_1}{\beta_1 c_\mu}\right)^{0.75} - (1-F)\left(\frac{\gamma_2}{\beta_2 c_\mu}\right)^{0.75}\right\}}{1-F} \qquad (4)$$

where C_{SMAG} is the Smagorinsky constant, γ_1, γ_2 are k-ω constants and F is the Menter blending function, to be used as a "shield" to prevent the LES model switching on inside boundary layers.

The switch between RANS and LES was implemented expressing the kinetic energy dissipation in the k-equation in the following form:

$$\varepsilon = \frac{\rho k^{3/2}}{min\{l_{RANS}, l_{LES}\}} \qquad (5)$$

The modified length scale is also employed in the constitutive relations, to obtain the time scale τ used in the computation of eddy viscosity and also to normalize the vorticity and the strain rate tensor. We have therefore $\mu = c_\mu \rho \tau$, where c_μ is a function of the scalar invariants of $\tau S_{ij}, \tau\Omega_{ij}$, and

$$\tau = \frac{min\{l_{RANS}, l_{DES}\}}{\sqrt{k}} \qquad (6)$$

It is worth to note that by this formulation, in LES regions the ω-equation has no influence, so that SGS model reduces effectively to a one-equation model.

3.2 Coupling method

The method presented here is a simplified overlapping grid technique. It permits to extract the informations about the noise sources on a generic integration surface around the solid (sphere, airfoil, …).
First of all, it is necessary to make a correspondence between the overlapping node of each grid and an element of the other grid. An efficient search by using Alternating Digital Tree (ADT) is performed in order to select candidate elements. Then, a test on the elements is performed in order to select the one which contains the node. The method is based on the finite element interpolation theory.

Following this theory, a function $\boldsymbol{F}$ is approximated as:

$$\boldsymbol{F}(\xi) = \sum_i \boldsymbol{N}_i(\xi)\boldsymbol{F}_i \qquad (7)$$

where ξ represents the coordinates in the parametric space, $\boldsymbol{N}_i$ the shape functions associated to the node i of the element and $\boldsymbol{F}_i$ the value of the function at this node. Using Eq. (7), the following function can be built:

$$\boldsymbol{G}(\xi) = \sum_i \boldsymbol{N}_i(\xi)\boldsymbol{X}_i - \boldsymbol{X}_p \qquad (8)$$

which corresponds to the difference between the interpolation of the coordinates within the element and the coordinates of a given point $\boldsymbol{X}_p$. The parametric coordinates ξ are solution of $\boldsymbol{G}(\boldsymbol{\xi}) = \boldsymbol{0}$, which is solved using the Newton-Raphson method, i.e.

$$\xi_p^{n+1} = \xi_p^n - \boldsymbol{G}_{\xi}^{-1} \cdot \boldsymbol{G} \qquad (9)$$

where $\boldsymbol{G}_{\xi}$ is the Jacobian of the function $\boldsymbol{G}$. Then, the interpolation weights are computed as $\boldsymbol{N}_i(\boldsymbol{\xi}_p)$.

For 2D, triangles and quadrilateral have been taken into account. For 3D application four kinds of elements were taken under consideration: tetrahedra, pyramids, prisms and hexahedra. Suitable shape functions were defined for each of them. Checks on the feasibility of the value of $\boldsymbol{\xi}_p$ are performed on the basis of the knowledge of the valid intervals of variation related to the element.

This procedure is very efficient and cheap but it enables the extraction of the noise sources in a correct way if the integration surface is closed to the solid bodies contained in the computational domain, because the UNS3D code does not propagate correctly the acoustic solution for long distances. In fact if we consider the behavior of the pressure after interpolation along a line in the case of the pulsating sphere, we found that it decays as 1/r for short distances of the integration surface from the sphere (Fig. 1).

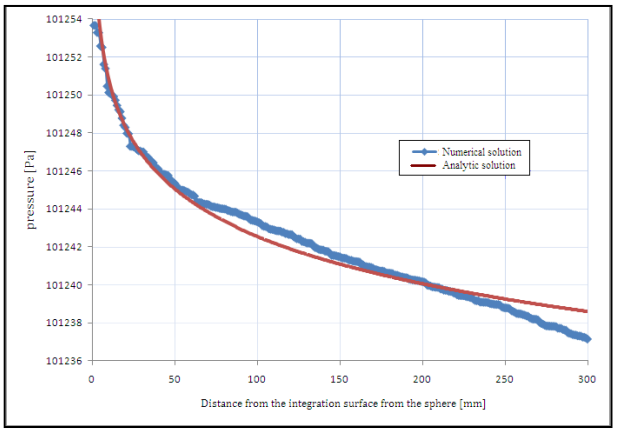


Fig. 1 Interpolated pressure along a line for a pulsating sphere. Numerical solution, Analytic solution.

3.3 The code FWH

To evaluate the acoustic fluctuations in the far field, the three dimensional Ffowcs Williams and Hawkings equation is solved [7].

This equation is an exact rearrangement of the Navier–Stokes equations and allows one to determinate the acoustic signal at distant ob-

server locations if the details of the source region are already known. The idea behind this approach is to extend the definition of all fluid-dynamic quantities to the whole domain, as well as to the regions occupied by a body. To do this, the fluid-dynamic quantities are multiplied by the Heaviside function of argument *f*, *H(f)*, where $f = 0$, defines the region outside of which the solution is desired:

$$H(f) = \begin{cases} 1 & \text{outside the surface} \\ 0 & \text{inside the surface} \end{cases} \qquad (10)$$

The only limitation in the choice of this surface is that it must enclose all the solid bodies contained in the computational domain.

The FW-H formulation can be written in differential form as:

$$\left(\frac{\partial^2}{\partial t^2} - c_0^2 \frac{\partial^2}{\partial x_i \partial x_i}\right)(H(f)\rho') = \\ = -\frac{\partial}{\partial x_i}(F_i \delta(f)) + \\ + \frac{\partial}{\partial t}(Q\delta(f)) \qquad (11)$$

where

$$F_i = (\delta_{ij} p + \rho u_i (u_j - v_j)) \frac{\partial f}{\partial x_j} \qquad (12)$$

$$Q = (\rho_0 v_i + \rho(u_i - v_i)) \frac{\partial f}{\partial x_i} \qquad (13)$$

The Eq. 11 are resolved on an integration surface. A prime denotes a perturbation from the freestream condition denoted by the subscript $(.)_0$. The fluid velocities are u_i, wheras v_i represent the velocities components of the surface *f*. The total density and pressure are given by $\rho = \rho_0 + \rho'$ and $p = p_0 + p'$, respectively.

The FW-H formulation is used for the propagation of the acoustic fluctuations in the far field. It is reasonable to assume that in the far field the mean flow is uniform.

4 Results

4.1 Model validation: non-punctual monopole

To validate the numerical model, the sound emitted by a pulsating sphere was computed.

The value of the parameters used for the CFD simulation are presented in Tab. 4.1.

Parameter	Value
Frequency of pulsation	2000 Hz
Amplitude of pulsation	6.8 m/s
Position of the receiver	1 m
Number of time steps	200
Time step	0.00005 sec
Ambient density	1.225 kg/m^3
Ambient pressure	101325 Pa

Table 4.1 Parameters used to the analysis of the sound produced by the monopole.

In order to reproduce the pulsations of the sphere, we choose to use a velocity boundary condition. In particular we impose a harmonic velocity normal to the sphere wall equal to:

$$u = \bar{u} \sin(\omega t) \qquad (14)$$

where $\bar{u}$ is the amplitude of the pulsation.

The amplitude of the pulsation velocity of the sphere, $\bar{u}$, must be kept sufficiently small to avoid the non-linearities linked to the propagation of the wave. For these reasons, during our computation, we use an amplitude of the pulsation velocity equal to the *1%* respect to the velocity of the freestream, with Mach number equal to *0.2*.

Concerning the time step, since we are using a compressible solver, it has been chosen in order to obtain a CFL number close to unit to correctly represent the acoustic wave propagation. Therefore in our case the time step is equal $5 * 10{-}5\ sec$.

In Fig. 2 is represented the computational domain used for the CFD simulation: it is a non-uniform unstructured grid composed by *255,473* points and *1,465,151* tetrahedral elements.

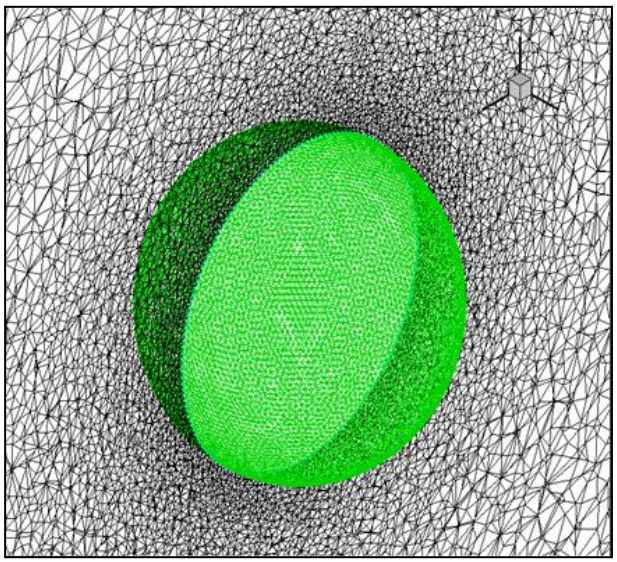


Fig. 2 Computational domain for CFD solution

The choice to use a spherical geometry is useful because all the sources of sound used in this calculation have a spherical symmetry. The computational parameters are reported in Tab. 4.2.

Parameter	Value
Radius of the sphere	0.1 m
Radius of the domain	8.5 m
Number of points	255473
Number of elements	1465151

Table 4.2 Characteristics of the mesh used

The noise source is due to the wall movement, therefore the flow can be computed using Euler equations.

Looking at the computed acoustic results, which are obtained using a integration surface, with a radius equal to *0.125 m*, it is possible to note that the directivity of the acoustic source (Fig. 3) is influenced both by the distance between the acoustic source and the integration surface (Fig. 4) and by the mean flow velocity (Fig.5), in agreement with the analytic behavior. In Fig. 6 is represented the time history of the acoustic pressure in the point showed in Fig. 3, for which theta is equal to 225 degrees.

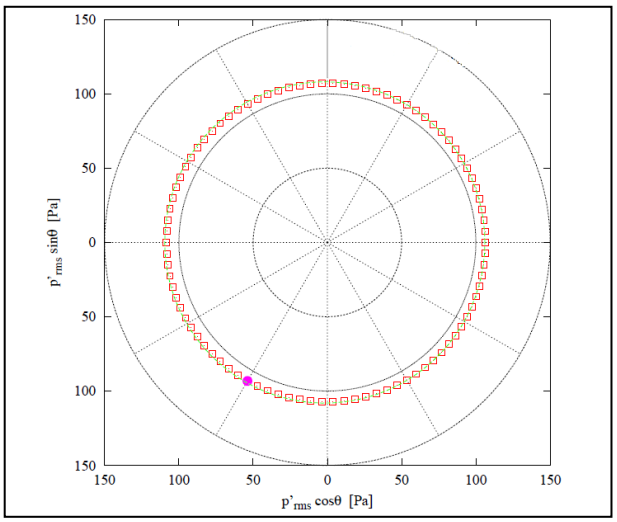


Fig 3 Sound Pressure Level for an acoustic monopole: Mach number of the free stream equal to 0.2, radius of the integration surface equal to 125 mm. Numerical solution, Analytic solution, Point used for the time history plot.

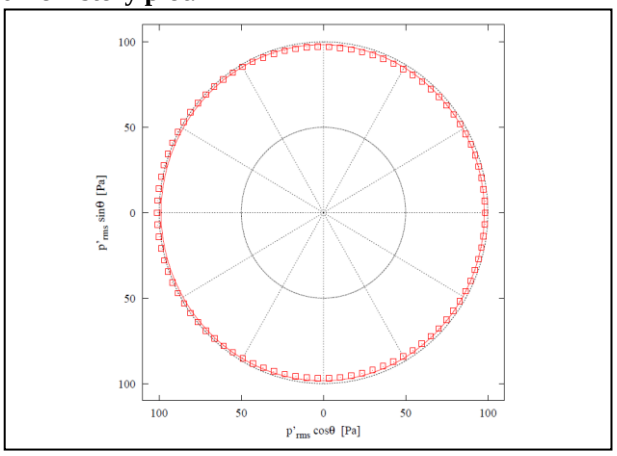


Fig 4 Sound Pressure Level for an acoustic monopole: Mach number of the free stream equal to 0.2, radius of the integration surface equal to 150 mm. Numerical solution, Analytic solution.

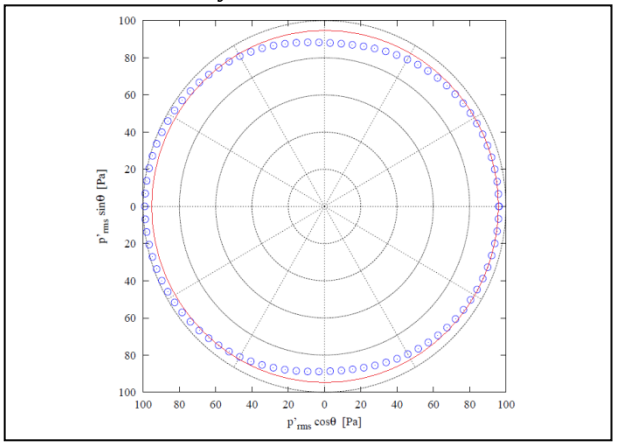


Fig 5 Sound Pressure Level for an acoustic monopole: Mach number of the free stream equal to 0.01, radius of the integration surface equal to 150 mm. Numerical solution, Analytic solution.

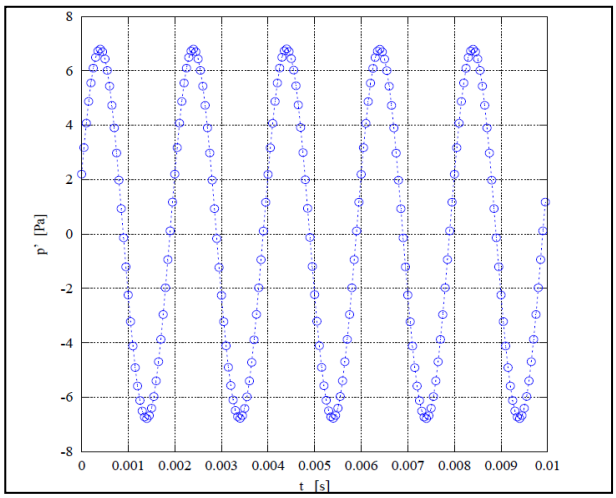


Fig 6 Time history of the acoustic pressure in a point of integration surface with theta equal to 225 degrees.

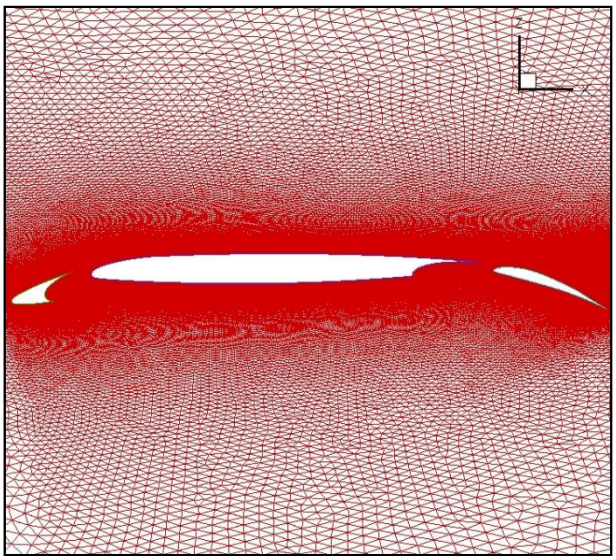


Fig 7 Mesh around the airfoil

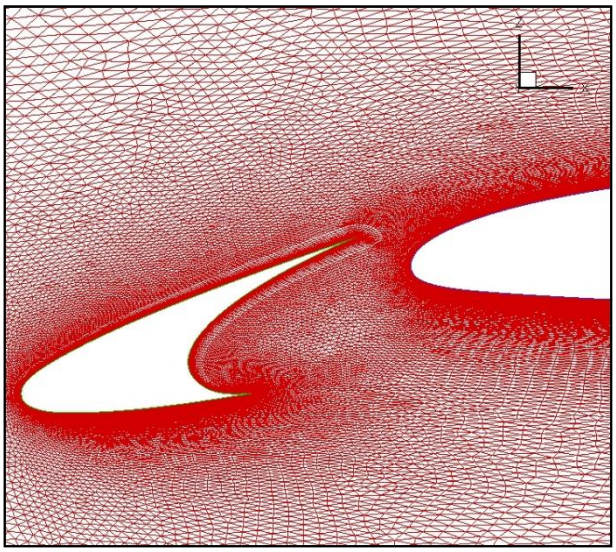


Fig 8 Zoom of the mesh around the slat

4.2 High-lift wing airfoil

The F15 three-element airfoil was used as a more challenging benchmark for the method. The acoustic flowfield around such airfoil has been studied in the DLR project LEISA (Low noise exposing integrated design for start and approach), combining activities in the research areas of high lift system design, flow control and aero-acoustic design methods [8].
The geometry consists of three elements, i.e slat, main wing and flap. The inclination angles are 28,814° for the slat and 38, 296° for the flap A 2D hybrid grid was initially generated, that is composed by hexahedra and prisms. The 3D grid was then built stacking 41 layers of the 2D grid in span-wise direction. The step size in spanwise direction was chosen to be *0.002 c* (chord) with *c=0.6m*, according to the minimal edge size in the slat cove. Thus the resulting span between periodic boundaries is *0.08 c* (Fig 7, Fig 8). The final grid is made by *4,254,078* points and *7,184,400* grid elements.

The parameters used for the CFD simulation are presented in Tab. 4.3.

Parameter	Value
Mach number	0.15
Incidence	6°
Time step	0.00007 sec
Number of time steps	1800
Ambient density	1.225 kg/m^3
Ambient pressure	101325 Pa

Table 4.3 Parameters used for the CFD simulation.

To initialize the flow field, an URANS computation was carried out until the mean flow converged to a constant value, then the DES computation was started. During the initial tran-

sient, a DES simulation was initially conducted for a time interval of 15 time scales, where the time scale t_{scale} is equal to:

$$t_{scale} = \frac{L_{scale}}{U_{scale}} \qquad (15)$$

where the length scale (L_{scale}) is the chord of the airfoil and the velocity scale (U_{scale}) is the freestream velocity. In order to ensure the time-accuracy of the results, the physical time step is taken equal to *0.00007 s*, meaning a sampling frequency of 15 kHz.

At the end of this initial computation, a slow drift of the mean flow features still existed, but the mean flow was considered sufficiently established. The computation was then performed over 1,800 time steps with a total duration of 126 ms, corresponding to 11 time scales.

Fig. 9 and Fig. 10 display the instantaneous y-component of vorticity in the plane (x,z), while in Fig. 10 the instantaneous x-component of vorticity is shown.

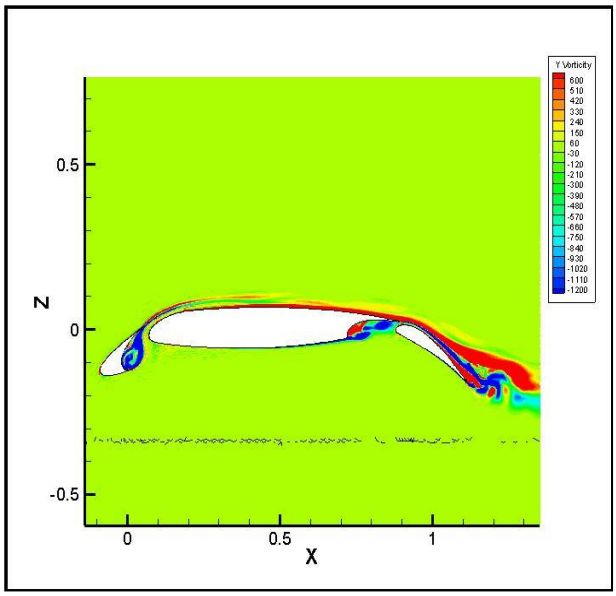


Fig 9 Y-Vorticity around the airfoil

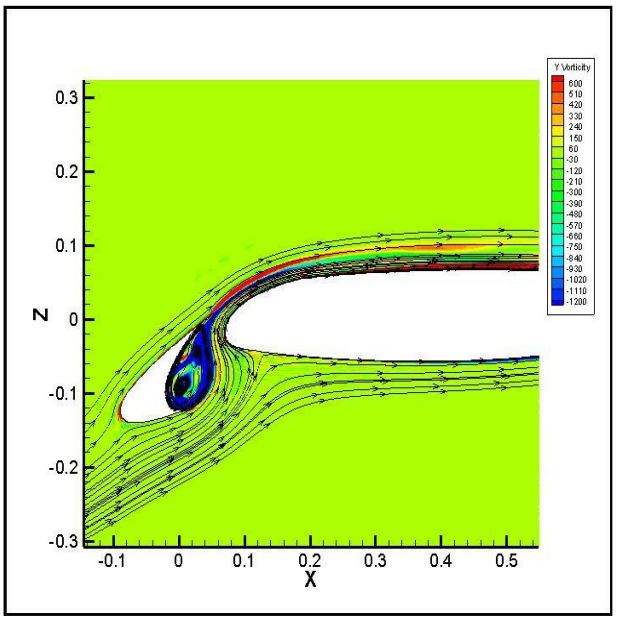


Fig 10 Y-Vorticity around the airfoil

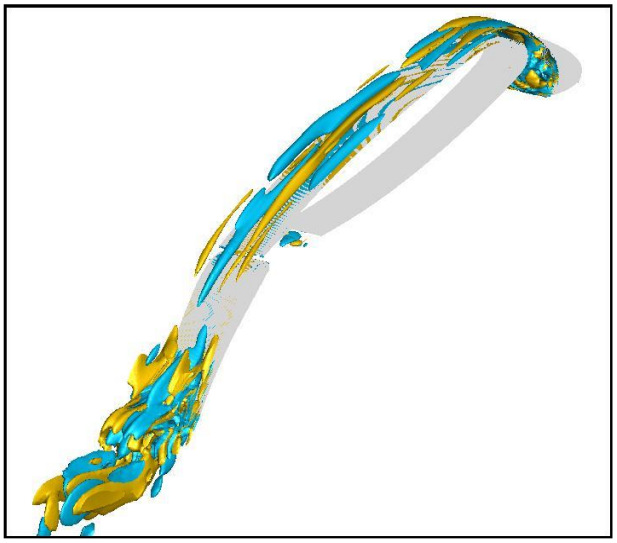


Fig 11 X-vorticity around the airfoil

These results show the presence of a vortex shedding from the trailing edge of the slat. Vortices impact the leading edge of the main airfoil and are carried out by the flow up to the mid-chord of the slat, where a separation region is present and the vortex structure breaks down into a large-scale resolved turbulence. The time-averaged Cp on the profile can be seen in Fig. 12, compared with the DLR experimental values.

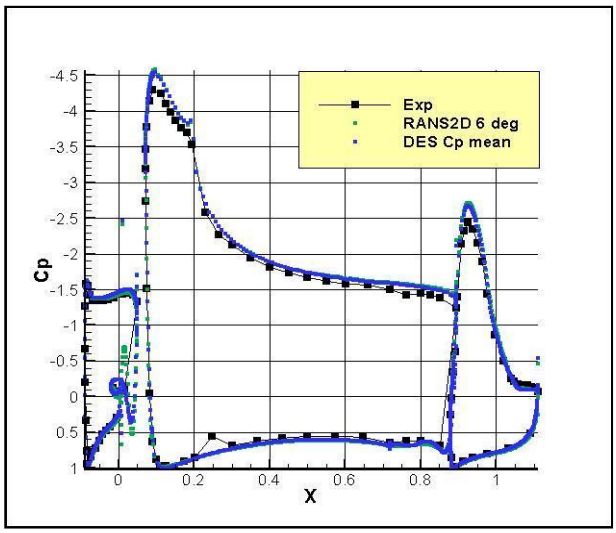


Fig. 12 Time-averaged Cp on the profile

Fig. 13 displays the integration surface, on which it has been possible the extraction of the informations about the noise sources.

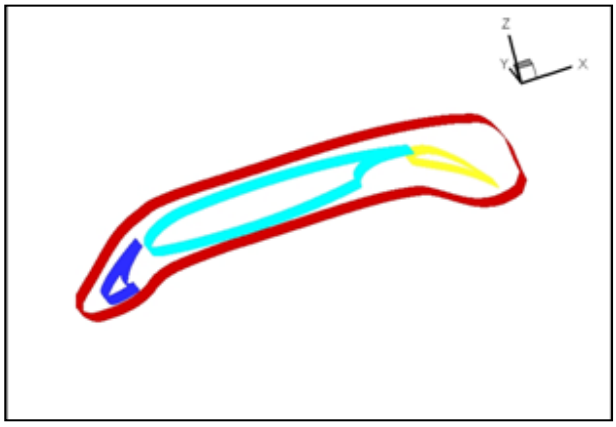


Fig 13 Integration surface

Far-field directivity is computed at r = 1.22 m. The directivity pattern, presented in Fig. 14, shows a dipolar character, with the silence cone in front of the leading edge, due to the finite chord of the airfoil. The main lobes are deflected in forward direction, as result of the convective effect.

The far-field directivities for three different frequencies are presented in Fig. 15, Fig. 16 and Fig. 17. It can be noticed that while for the lowest frequency the pattern is almost dipolar, for the higher frequencies extra lobes appear in the directivity patterns. This phenomenon can be ascribed to the ratio between the wave number and the airfoil cord. For the lower wave numbers, the wavelength is comparable to the chord, so they are less influenced by its presence. For higher wave numbers, multiple diffractions are possible on the walls, causing the extra peaks in the patterns.

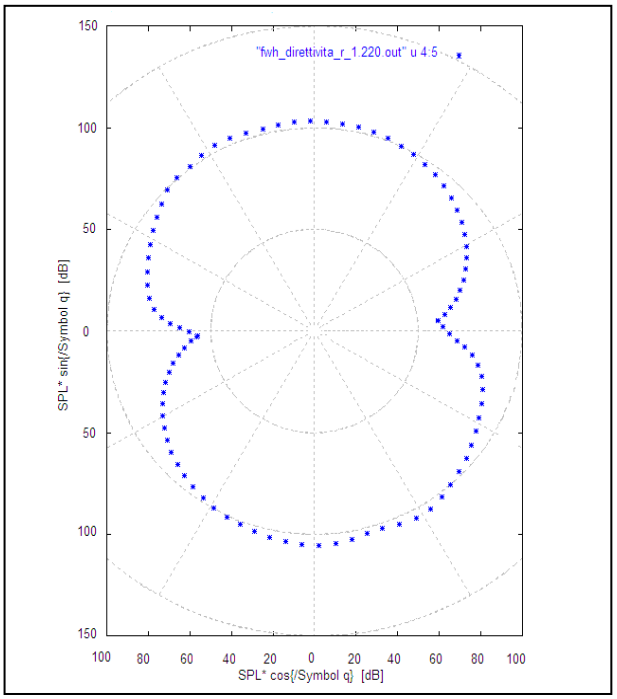


Fig. 14 Far-field directivity (r=1.22 m)

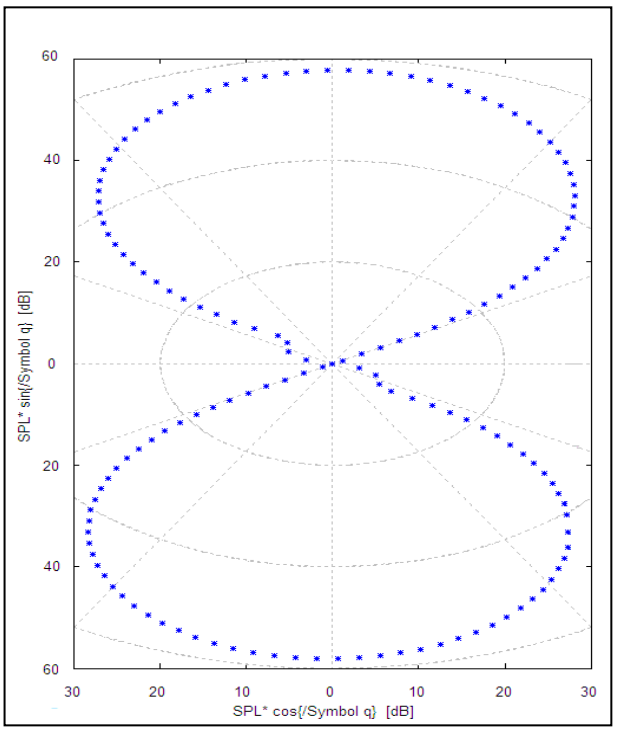


Fig. 15 Far-field directivity f= 15.87 Hz

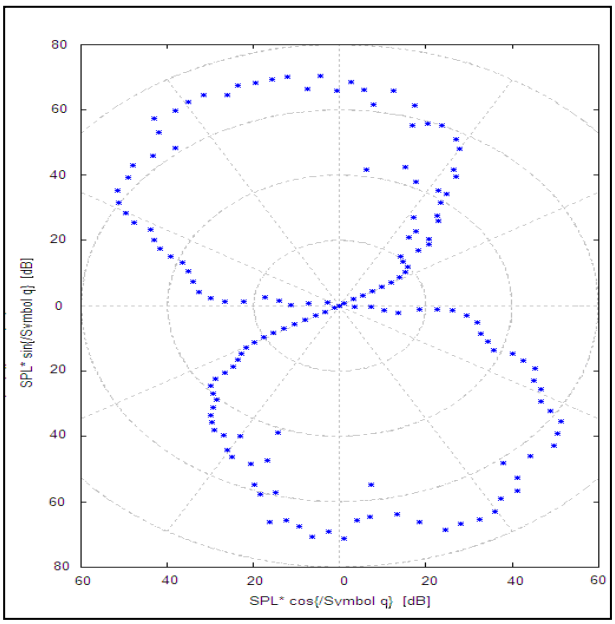


Fig. 16 Far-field directivity f= 4 kHz

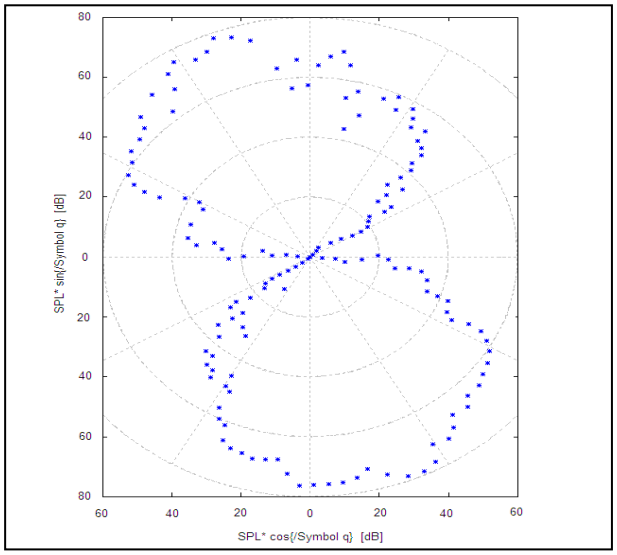


Fig. 17 Far-field directivity f= 15 kHz

5 Conclusions

A CFD/CAA coupling method is presented. This hybrid process combine several numerical methods, each one simulating specific fluid mechanisms (unsteady aerodynamics or acoustics) in a particular domain, by solving the adequate set of equations using the most economic discretization strategy: a compressible unsteady CFD (URANS/ DES) is used to simulate the local unsteady flow around the airfoil configuration and a FW-H integration provides the farfield noise. The critical point of the method is the coupling between the domain of unsteady CFD and the aeroacoustic domain.

The present study uses a straightforward technique based on the interpolation of CFD data, a process which relies on the compressibility of the unsteady CFD and its ability to correctly simulate the local acoustic field.

Acknowledgements

The present work was partly funded by the EC VII Framework Progamme Research Project ATAAC.

The present work has been performed through a close cooperation with Alenia Aeronautica S.p.A and through the use of Confidential Information and Data, property of Alenia Aeronautica S.p.A. which remains the sole owner of all such relevant IP rights.

The results of the present work shall be, therefore, property of Alenia Aeronautica S.p.A.

References

[1] Daniel J. Bodony, George Zagaris, Adam Reichert, Qi Zhang, *"Aeroacoustic: predictions in complex geometries"*, Procedia IUTAM, Volume 1, 2010, pp. 234 – 243.

[2] D. König, S.R. Koh, M. Meinke, W. Schröder, *"Two-step simulation of slat noise"*, Computers & Fluids, Volume 39, 2010, pp. 512–524.

[3] Ffowcs Williams, J., and Hawkings, D., *"Sound Generation by Turbulence and Surfaces in Arbitrary Motion"* , Proceedings of the Royal Society A Mathematical Physical And Engineering Sciences, Vol. 264, No. 1151, 1969, pp. 321–342.

[4] M.Shur, M.Strelets, A. Travin, "*Detached Eddy Si mulation (DES): General Description, Existing Versions, Constants Calibration, and Overview of Applications*", FLOMANIA Deliverable no.D4.41-5, EEC Contract No. G4RD-CT-2001-00613, 2003.

[5] R. Arina, A. Iob, C. Schipani, *"Frequency-Domain Linearized Euler Model for Turbomachinery Noise Radiation Through Engine Exhaust"*, AIAA JOURNAL, Vol. 48, No. 4, April 2010.

[6] A. Hellsten, *"New Advanced k-ω Turbulence Mode For High-Lift Aerodynamics"*, 42nd AIAA Aerospace Sciences Meeting and Exhibit, Reno, 2004.

[7] Wallin, S. & Johansson, A.V., *"An explicit algebraic Reynolds stress model for incompressible and com pressible turbulent flows*", J. Fluid Mech. Vol. 403, pp. 89-132, 2000.

[8] Wild, J., Pott-Pollenske, M., *"An Integrated Design Approach for low Noise exposing high-lift devices"*, AIAA paper 2006-2843, 2006.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Functional Shape Parametrization Approach for Preliminary Optimization of Unconventional Aircraft

Pier Davide Ciampa, Thomas Zill, Till Pfeiffer and Björn Nagel

German Aerospace Center, DLR

Air Transportation Systems, Blohmstr. 18, 21079 Hamburg, Germany

Keywords: *Shape Parametrization, Blended Wing Body, Collaborative Design, CPACS, MSO*

Abstract

The paper introduces a functional shape parametrization approach for the design and the optimization of unconventional aircraft configurations. The proposed design methodology is tailored for the preliminary design stages, and it is based on the use of mixed fidelity simulations. The approach is integrated in the DLR collaborative design framework, which makes use of a centralized model, and distributed analysis modules. The overall design process is applied for the multidisciplinary shape optimization of a Blended Wing Body configuration, including a large number of design variables. Single and multi-objectives optimization problems are formulated, and the optima designs are compared and discussed.

1 Introduction

The expected development of the air transportation system's ecological impact constitutes challenges for the design of next generations' air vehicles. For instance, the ICAO (International Civil Aviation Organization) member states agreed in 2009 on the global goal of an annual two percent fuel efficiency improvement till the year 2050 [1]. Therefore unconventional configurations are currently investigated in order to meet these ambitious design requirements. However, revolutionary air vehicle concepts cannot be assessed using the established predesign methodologies, which primarily rely on statistical data. Nevertheless, the recent advances in computational performance and simulation capabilities of computers provide accessibility to sophisticated physics based models, which can be included in the early stages of the design process. At the same time in order to correctly determine the vehicles' performance and to minimize the risks associated with the development of unconventional designs, multidisciplinary design and optimization (MDO) techniques are necessary to understand the interdisciplinary dependencies and the coupling effects of novel configurations and technologies. Nevertheless, state of the art aircraft predesign approaches are often based on automated, but monolithic design codes which cannot easily be adapted to cope with unconventional designs. In such cases a novel technology is taken into account by the use of technology factors, which are an estimation of the technology's impact on the performance properties of the air vehicle. The

final results depend to large parts on this estimation and may be subject to large uncertainty. Ideally the integration of novel technologies into the overall aircraft design process should make use of physics-based simulation. DLR is developing a collaborative design environment, which allows for the development of MDO-workflows utilizing high-fidelity simulation codes, to account for flexibility in the workflow definition and the straightforward exchangeability of individual analysis modules.

Recently the interest in higher fidelity analysis has shifted from single discipline to multidisciplinary shape optimization (MSO) applications, which commonly involve sizing, topology and shape optimization [2]. In order to cope with the increase in model complexity and computational resources demanded by hi-fi codes, efficient aircraft parametrization techniques are necessary to support a cost effective MDO process [3]. Herein, desirable properties include for instance a small number of design variables, a well-behaved analytical description of the aircraft geometrical shapes, and the capability to provide consistent model information for the disciplinary analysis tools.

In this paper a functional parametrization approach is presented for MSO applications for unconventional air vehicle configurations. The process is tailored for aircraft predesign activities, making use of high fidelity analysis methods at the early stages of the development process. First the employed parametrization and the corresponding functional model generator are described. Subsequently, DLR's collaborative design environment and the multifidelity analysis codes used in the study are briefly introduced. With these building blocks a methodology for the design and optimization of Blended Wing Body [4] configurations is developed.

2 Functional Parametrization

Core of the design optimization process is the shape parametrization of the aircraft and its components. Several parametrization techniques are available [5]. The right selection is dependent on the complexity and the level of detail of the aircraft configuration, as well as on the fidelity of the analysis codes in use, and on the control the designer wish to retain in the optimization formulation. Since the parametrization defines the degrees of freedom of the modeling and the type of allowed shape perturbations, the resulting optimum configuration is highly dependent on this fundamental choice. However a parameterization for MSO applications needs, among others, to be suitable for the integration into an automated process, providing consistent model information among the disciplines as well as analytical sensitivity derivatives [6]. Additionally, a straightforward interface of the parametrization to a geometrical data structure is desired.

Here, the presented approach combines the parameterization of the aircraft's outer shape surfaces and planform, by adopting a series of surface polynomial representation and a spline approach, well suited for representing three dimensional surfaces in the preliminary development stages.

The parametrization for the outer shape of wing-like surfaces is based on the Class-Shape function Transformation (CST) [7], an analytical formulation which is capable of describing complex aeronautical geometries (e.g. airfoils and wing-shaped surfaces) by mathematically well-behaved and smooth functions. The formulation allows for an accurate representation of the overall smooth geometry with a limited number of design parameters (shape function coefficients), and thus desirable for the implementation in large scale optimization problems. The CST approach is based on the definition of the physical geometry as a combination of a class function and a shape function, as follows in Equation 1 for the 3D case of a wing-like aero surface component:

$$\zeta(\psi,\eta) = C_{1.0}^{0.5}(\psi) S(\psi,\eta) \qquad (1)$$

$$S(\psi,\eta) = \sum_{i}^{Nx}\sum_{j}^{Ny} B_{i,j} Sx_i Sy_j \qquad (2)$$

where ζ and ψ are the airfoil coordinates normalized over the airfoil chord, and η is the normalized span. The class function "C(ψ)" defines the shape type, for instance airfoil-like geometries, whereas the shape function "S(ψ, η)" in Equation 2 provides the transformation function from the physical design space of an airfoil to a well behaved analytic function. In this case the shape function is represented by a Bernstein polynomial of order Nx for the chordwise distribution, and Ny for the spanwise distribution, and whose coefficients $B_{i,j}$ control the perturbations of the wing geometry.

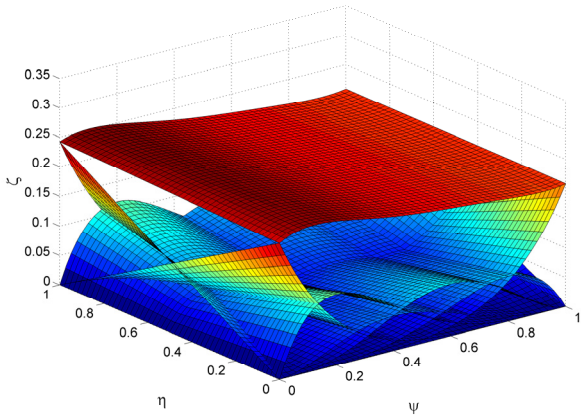


Fig. 1a CST shape surface S(ψ, η)

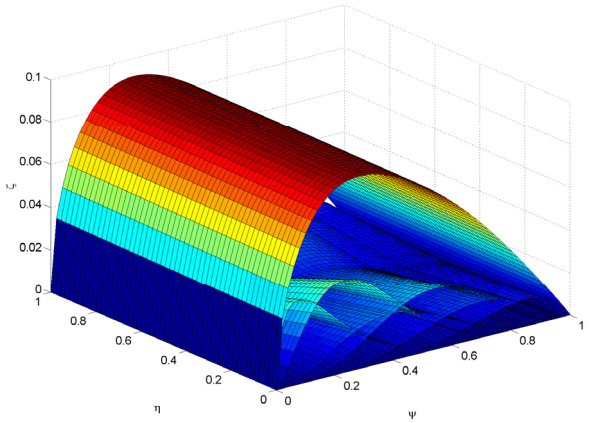


Fig. 2b CST aero surface

Figure 1a shows the shape function S associated with a wing-like segment (and its polynomial components). Figure 1b represents the corresponding aero surface in the physical design space. Multiple CST aero surface segments can be assembled together for complex shapes.

In a previous study [7] the authors implemented the CST concept into a model generator and proved the efficiency of the CST shape coefficients as driving parameters of an MDO problem for a fixed planform aircraft In comparison to this previous study, planform variations are taken into account in this paper, in order to cope with the wider scope of MSO. Therefore, the set of design variables is extended by planform functions, whose coefficients drive the distributions of the aircraft planform parameters such as wing sweep angles, twist angles, and taper ratio. Several types of functions (e.g., Bezier curves, NURBS) are available to control the planform shape, according to the desired complexity of the model, and the scope of the optimization task (global or local shape perturbations). Additional requirements which are taken into consideration in the presented approach are the minimization of the amount of data, which the designer has to initially provide (e.g. number of airfoils, and general planform and structural layout), and the automated enhancement of the model for suitable higher fidelity analysis.

Figure 2 shows an example of planform distribution functions, each with an independent number of control points along the nondimensional span. This allows for the tailoring of the scope of the analysis, and for the integration of the available design knowledge. For instance, the designer may wish to retain a larger number of control points on the twist distribution of the outer wing, without the need to specify the same amount of input parameters for taper ratio, and at the same time predefine a known dihedral distribution, or a typical behavior.

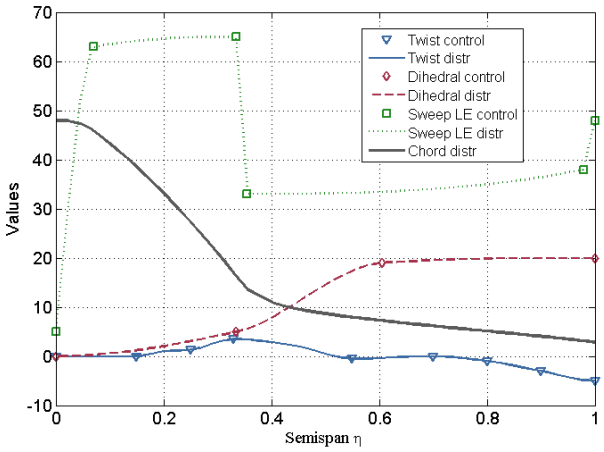


Fig. 2 Planform functions

The planform functions for a wing-like geometry define in space the reference plane separating the lower and upper surface, as shown in Fig. 3, which is fully controlled, and thus deformed, by the coefficients values of the control points. Figure 3 shows the planform produced for a Blended Wing Body configuration with a continuously varying swept leading edge, and nonlinear twist distribution. The corresponding twist, sweep and dihedral distributions are applied along a curvilinear abscissa running from the wing root to the wing tip and placed at an arbitrary chordwise position.

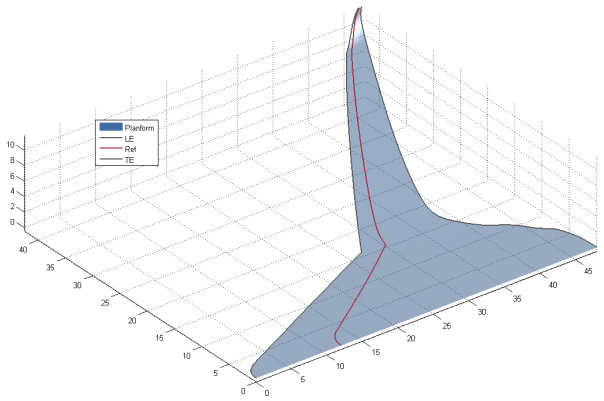


Fig. 3 Reference planform plane

Therefore the CST surfaces, which control the type of airfoils and their properties such as thickness and camber, are placed and deformed along the planform reference plane.

The functional model generator implemented includes a set of knowledge base design rules determine the relations between the design variables in order to define how perturbations of the design variables are propagated along the geometry when changing values of the functions coefficients (e.g. straight lofting between airfoils, or mesh deformation of the initial geometry).

The BWB configuration can be completely described as an assembly of wing-like CST surfaces in spanwise direction, deformed by the planform distributions. The set of the overall function coefficients represents the design variables. The number of input CST segments depends on the number of airfoils the designer wants to specify and control during the optimization process, whereas the output model is enriched by the automated addition of transition airfoils, dependent on the shape surface coefficients and on the gradients of the planform distributions over the reference line.

The overall outer shape is completely described by a unique set of analytical functions, which can be evaluated for different types of discretization, in order to extract computational grids with different levels of complexity and accuracy. Figure 4 shows the BWB shape produced by the functional model generator, and a typical structured paneling grid evaluated on an equivalent spanwise spacing and cosine chordwise spacing of the parametric design space.

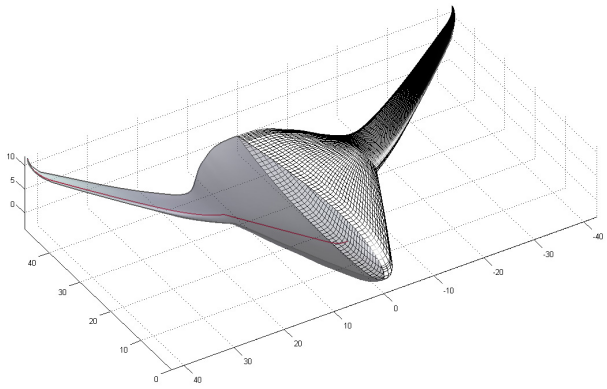


Fig. 4 BWB functional model generator

In this paper the setup of the functional model generator is tailored for the preliminary design stages, hence with the goal to minimize the time to initialize the aircraft modeling, which will be subjected to large variations during optimization. In fact, a fundamental

feature of the chosen parametrization is the capability to produce local planform deformations, but smooth surfaces perturbations. This allows to extensively sweep through the design space, such as for large MSO problems, in order to generate a wide variety of loading-performance cases, but at the same time minimizing the number of geometrically infeasible solutions which may be generated during the optimization process.

3 Collaborative Design Environment

In this study the proposed functional parametrization is integrated into DLR's collaborative design environment. The framework provides a mean for the development of intricate flexible workflows in a distributed computing architecture and consists of three individual building blocks.

The basis for communication is the central data model CPACS (Common Parametric Aircraft Configuration Schema), which is based on XML technology. CPACS holds detailed constructs for the exchange of information between the analysis modules in multidisciplinary workflows. It contains model information as well as data on the process to control parts of the overall analysis workflow, such as necessary input to initialize the individual disciplinary analysis modules, e.g. aerodynamic and structural solvers. CPACS has been under development at DLR since 2005 and it is designed to accommodate data for numerous disciplines involved in the conceptual as well as preliminary design phase and is also seeing application in high-fidelity analysis, e.g. crash simulation of fuselage sections, see Fig. 5.

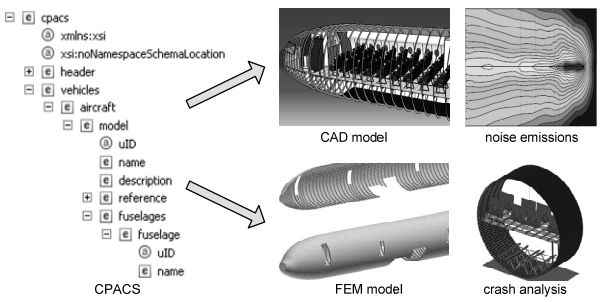


Fig. 5 CPACS concept

The second building block is the toolset of disciplinary analysis modules which exchange information using the common language. The workflow is developed to make use of a decentralized distributed computing architecture, in which the analysis modules are hosted and run on dedicated servers located at the corresponding institute of DLR or at any partners' sites. Whereas only input and output data are made accessible to the user, the disciplinary modules remain on these locally distributed servers during the workflow execution. All analysis modules in the toolset require a CPACS interface. For legacy codes which do not feature this direct interface, a straightforward wrapper component translates between the codes' proprietary format and CPACS.

The third fundamental component is the engineering framework software which manages the workflow control as well as product optimization. Currently the environment uses a commercial engineering framework software solution for these tasks, which has been enhanced by DLR's Chameleon-components [8]. They provide further capabilities for the handling of CPACS within the framework software, such as central program libraries specifically tailored to the central model for unified interpretation of geometric data.

The collaborative design environment is in operational use at all aeronautical institutes of DLR. First applications have been presented in recent works, in which the environment has been used to develop workflows for aero-structural optimization [9] and preliminary aircraft design [10].

4 MDO Analysis Architecture

In this study the functional parametrization module is used for the generation of a centralized geometrical model. The parameterized geometry is transferred and stored in the structured XML data format CPACS.

4.1 Aerodynamics Modules

The aerodynamic analysis is based on the potential multiple lifting line code for non-planar wings LIFTING_LINE [11] chosen for the calculation of the aerodynamic polars. The three dimensional panel code VSAERO is chosen for the computation of the aerodynamics loads which are used for the structural analysis. Friction drag is estimated by the flat plate turbulent boundary layer correlation. All the codes are available through the distributed environment, and interfaced to the CPACS centralized model by the disciplinary experts. Figure 6 shows two types of aerodynamic models extracted from the common CPACS representation, for the BWB configuration under study: the zero thickness panel grid for the port side, and the three dimensional panel model for the starboard side.

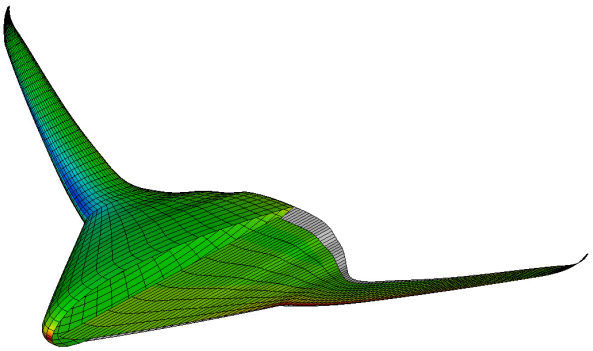


Fig. 6 Aerodynamics models 2D (left side) and 3D (right side)

4.2 Structural Module

The structural model is generated by the DLR in-house application PARA_MAM [12], which is capable of assembling multilayered shell Finite Element (FE) models based on the external geometry (parameterized as previously described), and on the parametric description of the inner structural topology in terms of spars, ribs number, and relative positioning of the structural components. PARA_MAM, which consists in a set of MATLAB macros, fits the inner structures in the reference outer surface, and generates a collection of matrices containing the data required to define the model within the pre-processor of the FE package. For the current approach the multilayered shell FE model generated is described by the ANSYS Parametric Design Language (APDL) format, which can be interpreted by the FE software package ANSYS in batch execution. The assembly of the structural module is fully automated. The model is used for the evaluation of the strength constraints under a series of load cases.

In this study the structural analysis considers only the aerodynamic loading produced from a 2.5g pull-up maneuver, as representative load case for the structural sizing. Hence the corresponding aerodynamics loads over the three dimensional surface elements are mapped and distributed over the structural FE grid.

The sizing infrastructure S_BOT (Sizing roBOT) developed by DLR is used to determine the thickness distribution of the shell elements. The tool's main input file calls the FE software, reads the generated FE model, applies the mapped aerodynamic loads, launches the analysis solver, and post-processes the loading state for each element of the structural components. On the base of the chosen sizing criterion, the new elements thicknesses are calculated, and the dimensions of the skin element layers are updated for all the wing elements within the specified optimization regions. After updating the model, the load path is changed, and the analysis-sizing process is iteratively repeated till convergence criteria are met. The tool explicitly models the primary structure components (i.e., ribs, spars, and skin), whereas the substructures (e.g., stiffeners) are implicitly modeled by equivalent stiffness layers [13].

Figure 7 shows the model generated for the automated structural analysis of a Blended Wing Body configuration as interpreted by the finite element software package. In the figure half of the PARA_MAM generated FE model is displayed without the upper skin surface elements to visualize the inner structural components, whereas the other half of the model shows the aerodynamics loads as output from the aero solver and mapped on the structural surface elements.

In this study the parametric topology of the inner structure (i.e. relative positioning and number of primary structure components) is kept fixed. As for the aerodynamics, all the data needed for the disciplinary model preparation and analysis are extracted from the CPACS structure.

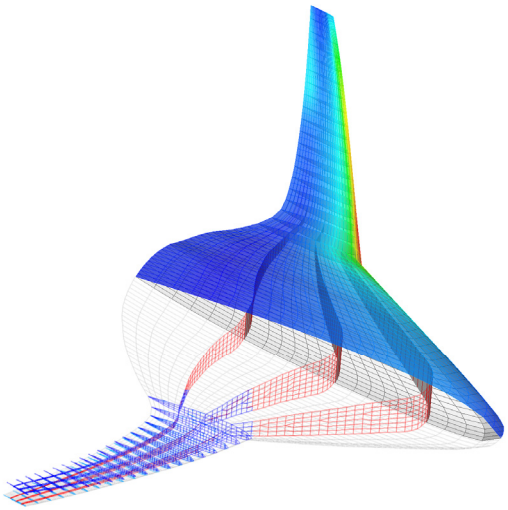


Fig. 7 Structural FE shell model

4.3 Weight and Balance Module

A weight and balance module is implemented for the estimation of the nonstructural masses such as payload, systems, furnishings, engines, and landing gear in order to estimate the Operating Empty Mass (OEM) and evaluate the aircraft's mission performance. Taking the outer shape as reference, additional inner geometrical components are derived, such as fuel tanks, cabin and cargo compartments. Additional components are added such as engines, and landing gears, as shown in Fig. 8. Based on the components dimensions the module provides an estimation of the masses and relative centers of gravities. Here, the number of passengers and are a function of the floor area of the cabin. With the aircraft masses, the aerodynamic performance of the trimmed aircraft, the propulsion systems, as well as the desired design range, the mission fuel is calculated using the established Breguet formulation

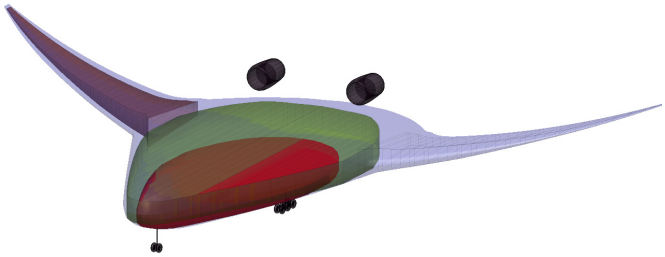


Fig. 8 Weight and balance model

4.4 Reference Configuration

A baseline aircraft is generated in order to check the quality of the functional model generator, and the disciplinary models derived. The baseline geometry is taken from a previous European project known as MOB [14], for which aerodynamics studies are available in literature.

The aerodynamics performance are computed for the reference configuration at the cruise design point Mach 0.85, with angle of attack of 3 degrees. The reference MOB results have been computed by RANS simulations [15], and thus total drag includes induced drag, wave drag, drag due to boundary layer displacements and friction drag. The reference values are compared with the lifting line code, and the three dimensional panel code, which account only for induced drag, and an estimation of the friction drag.

Additionally, an Euler code has been interfaced with CPACS in order to take into account wave drag. The resulting total drag obtained from the Euler computations is within 1% of the reference. Table 1 shows the lift design coefficient Cl and the total drag coefficients Cd calculated by the different tools (including the friction drag estimation), and compared with the reference values.

Analysis Code	Cl	Cd total
Reference (RANS) [13]	0.4136	0.03268
Lifting_Line (2D)	0.3461	0.01553
VSAERO (3D)	0.345	0.01525
Euler (3D)	0.431	0.03210

Table 1 Reference aerodynamics performance

In order to check the sizing process of the structural model with the reference value, aluminum technology has been selected, and the sizing case for a quasi static 2.5g maneuver chosen as representative. Respect to the reference value [3], the sized mass of the primary structures results 14% lower, and the difference is due to a different set up of the optimization regions, and of the design allowables defined.

5 Optimization of BWB

The functional model approach is tested for the optimization of the Blended Wing Body configuration. Table 2 gives the reference transport mission requirements used for this study, which are assumed to be comparable with the class of a large transportation aircraft, such as A380.

Design mission	A380-800	DLR-BWB
Range [km]	13500	14000
Mach	0.85	0.85
Altitude [m]	10000	10000
Payload PAX	525	500

Table 2 BWB TLAR

5.1 Multidisciplinary Design Workflow

Figure 9 shows the design workflow for the test case using the collaborative design environment, and the aforementioned analysis modules. The overall workflow can be described as follows.

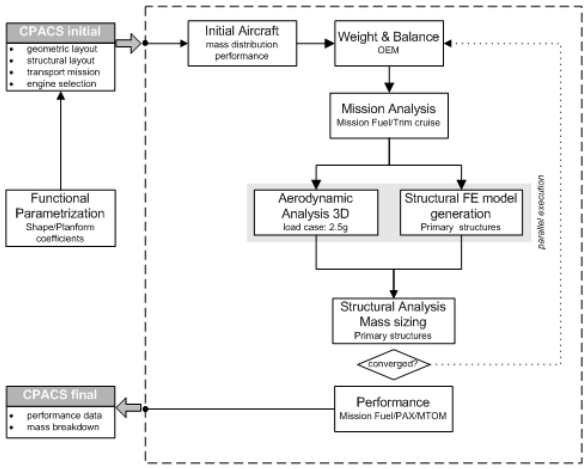


Fig. 9 Design workflow

The geometry is initialized by the functional model generator and transferred to the CPACS format. A first estimation of the structural and nonstructural masses based on the model dimensions is calculated by the weight and balance tool. During this stage, the aerodynamic polar is calculated by the lower fidelity code LIFTING_LINE and used for the calculation of the trimming condition. Subsequently, the mission performance module calculates the required fuel mass for the mission, and this overall preliminary loop is repeated till the global aircraft masses reach convergence. After the preliminary design loop has converged, the aerodynamics loads are calculated by the 3D panel code for the 2.5g maneuver. In parallel the high fidelity FE structural model is generated. The aerodynamic loads, which are used in the structural sizing of the primary structures, are mapped from the aerodynamic grid to the FEM grid. The structural sizing process is run until convergence of the thickness distributions of the main components is reached, satisfying the required strength criteria. The sized structural mass is fed back into the mission calculation, and the initial estimated global aircraft mass is then updated. The mission fuel and the trim conditions are recalculated, producing a new load distribution to be used in the sizing process. The overall design loop is repeated until the calculated sized structural masses have converged. Finally, the masses, the aerodynamic and the mission performance results are collected.

5.2 MSO Methodology

This paper focuses on the feasibility of a functional approach for design optimization with a large number of design variables. Here the design space is limited to 12 design variables for the control of the planform functions (twist, sweep, dihedral, wing span, root chord, and cabin width), and 36 coefficients for the control of the aerodynamic surfaces. To accommodate a rather computationally inexpensive optimization process, a Design of Experiments (DOE) is generated using a space-

filling stratified Latin Hypercube sampling plan, which then allows for the construction of a Kriging metamodel. Table 3 shows the surface coefficients bounding values, and corresponding variations for some of the representative characteristic planform geometrical parameters.

Design Variables	Lower	Upper
$Bi_{,j}$ coefficients	0.05	0.3
Twist at 75% span [deg]	-8	+8
Sweep outer wing [deg]	-40	60
Span [m]	50	80
Root chord [m]	40	60
Cabin width [m]	15	25

Table 3 DOE bounds

Each design point of the DOE has different planform and airfoils distributions. For each sample point the structural sizing and the performance convergence loop are taken into account. Figure 10 shows the Operating Empty Mass, and the sized primary structures mass, against the number of passengers (PAX) which can be fitted in the cabin, for each sampling point. It can be noticed that the growth of OEM is mainly proportional to the number of passengers, and dimensions of the design. Masses of the sized primary structures exhibit a more scattered behavior. In this study for the MSO application, only a part of the FE model is sized, consisting in the wingbox, and part of the central fuselage. Therefore the sized masses largely depend on the loading case, which is dependent on the geometrical parameters like planform, twist and airfoils distributions. The DOE fully spans the design space, and design points with a similar passenger capacity may provide significantly different aircraft shapes, and thus loading, which directly affects the sized masses. For the same set of points, Fig. 11 shows a fuel efficiency index defined as fuel per passenger per 100 km of the mission range. The BWB fuel efficiency tends to increase with higher number of passengers. For the reference mission and the conventional aircraft this value has been determined to be 3.5 l/PAX/100 km.

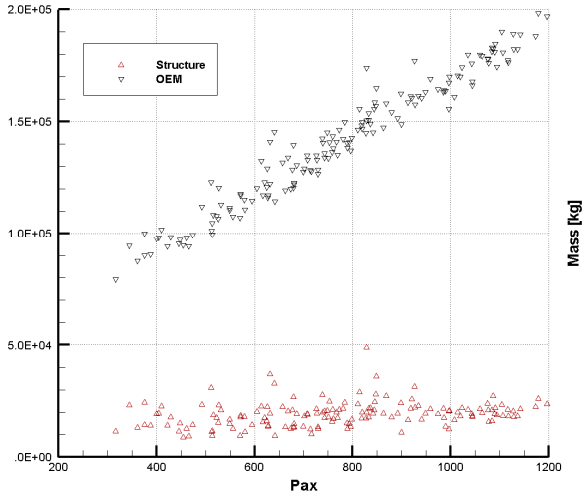


Fig. 10 DOE OEM/Structure Mass vs PAX

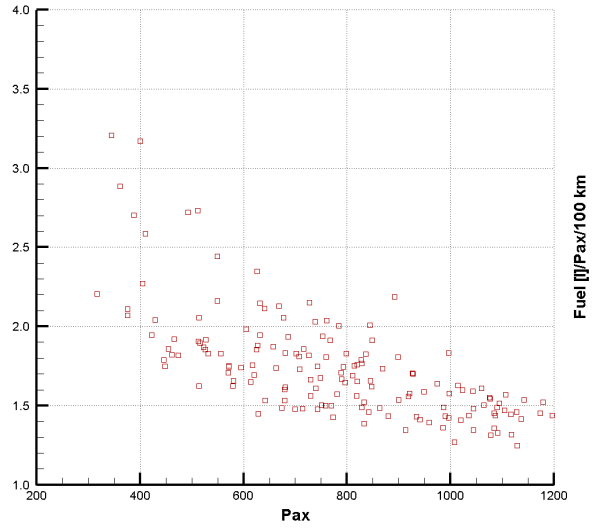


Fig. 11 DOE Fuel efficiency vs Pax

Although the described parametrization includes a compact number of variables, since higher fidelity analysis are used, the computational time for a direct optimization approach would be infeasible to properly fit in the conceptual and preliminary stage, especially when using the black box approach in the collaborative design environment. Hence the DOE simulation results are used to generate a Kriging-metamodel, and a computationally efficient MDO problem is formulated.

5.3 MSO Results

The optimizer drives the planform and the shape function coefficients, as defined for the DOE, to find the optimum design which maximizes the aircraft performance, subjected to the requirements (e.g. the defined mission) and design constrains (e.g. structural failure, stability margin).

A single objective optimization problem is initially formulated in order to minimize the required fuel for the mission defined in Table 2. The constraints of the optimization problem are set on the number of passengers which have to be accommodated in the cabin compartment, and on the volume of the fuel tanks, which have to contain the required mission fuel. For the solution of the optimization problem both, a gradient based and a genetic algorithms are applied, in order to verify the optimum is representing a global minimum.

Figure 12 shows the optimization history of the gradient based algorithm, and the ten best individual designs resulting from the genetic optimization. The optimal solutions for minimum fuel are all within a deviation of 1.9% for both optimizers, with almost identical planform characteristics.

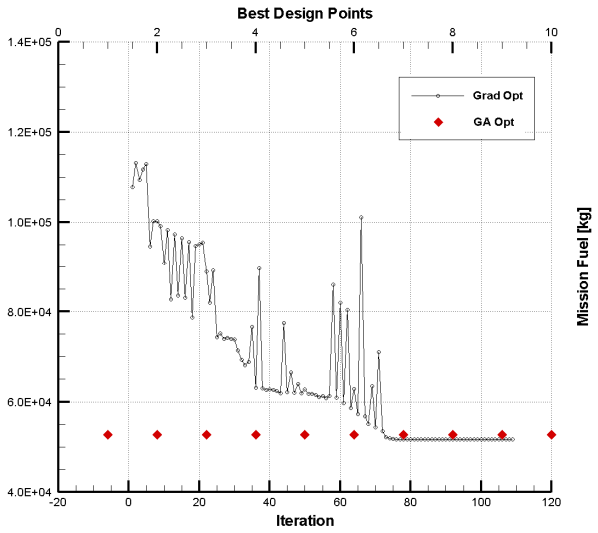


Fig. 12 Single objective optimization

Subsequently, a multiobjective optimization task is formulated, minimizing mission fuel and the primary structural mass using the aforementioned genetic algorithm. The Pareto-optimal designs are extracted and depicted in Figure 13 by the black line. The bottom axis indicates the mission fuel for the Pareto points, whereas the left vertical axis represents the sized structural mass. It can be observed that a decrease in the structural mass has to be traded with an increase in the mission fuel. The effect is caused by the shift of the aerodynamic loading from an ideal distribution to a less efficient one, but with less loading on the structure. However it can also be noticed that the reduction of the structural mass becomes much lower under a certain value. This effect is explained by the presence of active constraints during the structural sizing on the minimum thickness of the primary components. In fact even if the aerodynamic loading would lower the actual stress level on the structure (with decrease of the aerodynamic efficiency), the minimum thickness constraints limit the further decrease of the structural masses. In addition on the right vertical axis, the corresponding values of characteristic planform design variables for each Pareto-optimal design are given. Although the variation of these values is moderate, it is necessary to underline that each Pareto point exhibits different resulting airfoils, and twist distribution, which highly affect the loading and the efficiency of each design.

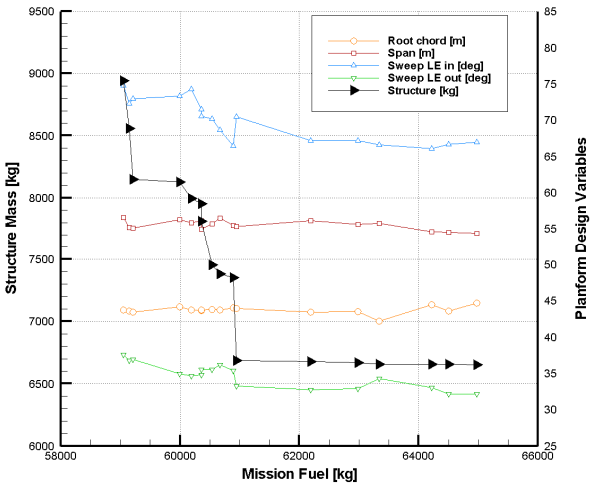


Fig. 13 Pareto front design points and corresponding geometrical parameters

As example two representative design points on the Pareto front are compared with the single optimization results, in Table 4. As explained it can be seen the design 5 has a lower fuel mass due to the higher aerodynamic efficiency, but a higher structural mass, respect to the design 15. Figure 14 shows the local lift distributions over the semi-span for these two design points. Further Table 4 shows the minimum fuel mass is reached by the single objective optimizations, which at the same time has the highest sized structural mass, compared with the Pareto solutions.

	Mission Fuel [kg]	Structure [kg]	L/D
Min Fuel Gradient	51627	9164	-
Min Fuel Genetic	52631	9607	-
Pareto design 5	60192	7995	21
Pareto design 15	64223	6655	19.9

Table 4 Optimization results

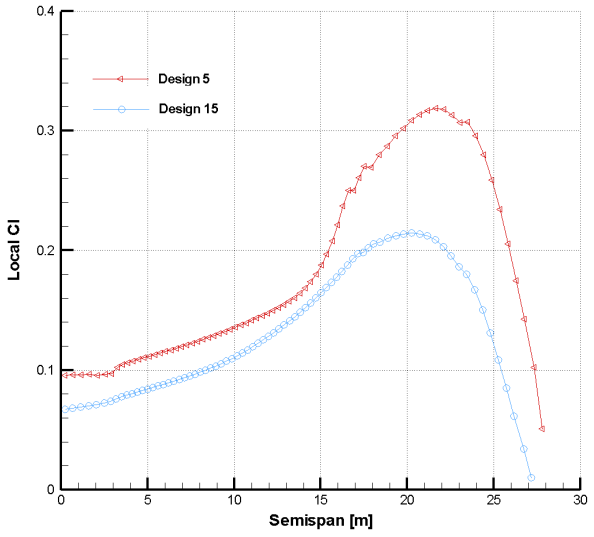


Fig. 14 Pareto points spanwise loading

6 Conclusions and Outlook

This paper presents a functional parametrization approach, suitable for the initialization and the optimization of unconventional configurations. The functional model provides a compact analytical description of the aircraft geometries, suitable for MSO applications at the preliminary design stages. A set of functions are used to describe planform characteristics and airfoils distributions, and the functions' coefficients are used as design variables. The presented approach is integrated into the DLR collaborative design environment, making use of the centralized model CPACS, and a set of distributed and heterogeneous disciplinary tools. The overall design process is tested for a reference BWB design, and results are comparable with available literature. Thus the design approach is tested for the multidisciplinary shape optimization of a Blended Wing Body configuration for a given reference mission comparable with the one for the actual large transportation aircraft. The study focuses on the parametrization methodology, and design framework, and thus a series of simplifications are assumed for the configuration optimized. For instance only aerodynamic load cases are considered for the sizing, and the parametric topology of the structural model is kept constant during the optimization process. Additionally a simplified cabin model is used for the estimation of passenger and payload, and no pressurization is considered for the mass calculations. These factors influence the absolute optimization results, and the global performance of the BWB configuration, but will be tackled in upcoming studies by increasing the number of parameters, and adding complexity to the analysis modules. The study shows the capability of the proposed design approach to drive large scale optimization problems with heterogeneous and mixed fidelity analysis.

7 Authors' Contact

Pier Davide Ciampa
German Aerospace Center, DLR
Blohmstr.18, 21079 Hamburg, Germany
pier-davide.ciampa@dlr.de
+49-40-42878-2727

References

[1] N.N., "ICAO Member States Agree to Historic Agreement on Aviation and Climate Change", *ICAO News Releases PIO 14/10*, International Civil Aviation Organization, Montreal, 2010

[2] Samareh J., "Geometry and Grid/Mesh Generation Issues for CFD and CSM Shape Optimization", *Optimization and Engineering*, Vol. 6, No. 1, 2005, pp. 21-32.

[3] La Rocca G., "Knowledge Based Engineering Techniques to Support Aircraft Design and Optimization," Ph.D. Dissertation, Faculty of Aerospace Engineering, TU Delft, the Netherlands, 2011.

[4] Liebeck, R. H., "Design of the Blended Wing Body subsonic Transport", *Journal of Aircraft,* Vol. 41, No. 1, 2004.

[5] Samareh, J.A., "Survey of Shape Parameterization Techniques for High-Fidelity Multidisciplinary Shape Optimization", *AIAA*, Vol. 39, No. 5, 2001, pp. 877-884.

[6] Kulfan B.M., "Universal Parametric Geometry Representation Method", *Journal of Aircraft*, Vol. 45, No. 1, 2008, pp. 142-158.

[7] Ciampa P.D., Zill T. and Nagel B., "CST Parametrizaion for Unconventional Aircraft Design Optimization," *27th International Congress of the Aeronautical Sciences*, Nice France, 2010.

[8] Bachmann, A., Kunde, M., Litz. M. and Schreiber, A., "Advances in Generalization and Decoupling of Software Parts in a Scientific Simulation Workflow System", *The Fourth International Conference on Advanced Engineering Computing and Applications in Sciences ADVCOMP*, Florence, 2010.

[9] Liersch C. and Hepperle M., "A Unified Approach for Multidisciplinary Preliminary Aircraft Design", *CEAS European Air and Space Conference*, Manchester, 2009.

[10] Zill T., Böhnke D. and Nagel B., "Preliminary Aircraft Design in a Collaborative Multidisciplinary Design Environment", *11th AIAA Aviation Technology, Integration, And Operations (ATIO) Conference*, Virginia Beach, VA, 20-22 September 2011.

[11] Horstmann, K. H.., "Ein Mehrfach-Traglinienverfahren und seine Verwendung für Entwurf und Nachrechnung nichtplanarer Flügelanordnungen" Ph.D. Dissertation, DFVLR FB 87-51, 1987.

[12] Nagel, B., Rose M., Monner H.P. and Heinrich, R., "An Alternative Procedure for FE-Wing Modelling", *Deutscher Luft- und Raumfahrtkongress*, Braunschweig, 2006.

[13] Ciampa, P.D., Nagel B., and van Tooren, M.J.L., "Global Local Optimization of Transportation Aircraft Wings", *51st AIAA Structures, Structural Dynamics, and Material Conference*, AIAA Paper 2010-3098, Orlando, 12-15 April 2010.

[14] Morris A., "MOB: A European Distributed Multi-Disciplinary Design and Optimization Project", *9th AIAA Symposium on Multidisciplinary Analysis and Optimization*, IAA Paper 2002-5444, Atlanta, Georgia, 4-6 September 2002.

[15] Qin, N., Vavalle, A., Le Moigne, A., Laban, M., Hackett, K. and Weinerfelt, P., "Aerodynamic considerations of blended wing body aircraft", *Progress in Aerospace Sciences*, Vol. 40, 2004, pp. 321-343.

CEAS 2011

Safety Systems and Subsystems

Development of integrated functional monitoring and warnings system for Saab 39 Gripen.

C. Birgersdotter and T. Fransson
Saab Aeronautics
S-581 88 Linköping
Sweden

Keywords: *Functional Monitoring, Warnings System*

Abstract

As the requirements of a modern air vehicle changes during development and use, requirements and design of test and recording functions have to be continuously updated. To achieve control over functional monitoring, pilot warnings and health management in a tightly integrated avionics system with several configuration variants and frequent updates, powerful tools are needed, especially when requirements on cost reduction and a small staff are considered. Traditionally the work has been divided among several departments, with own processes and tools, leading to redundant work and inconsistency, despite tremendous inspection efforts.

This paper describes how the workflow to define pilot warnings is integrated in order to reduce development time and reuse data. There are over 500 failure modes defined for the Gripen Aircraft.

The impact of a failure is depending on equipment configuration and thus will the required pilot actions differ between the variants. In complex failure situations it is also important to find the primary fault and filter faults that can be considered as consequences of the primary fault. The presentation will show how primary failures are distinguished from secondaries or consequences.

Experience has shown that the recommended pilot actions often need to be revised after that operational experience has been achieved by the users. As changes in a delivered product are costly, the method of separately loadable databases for flight manuals and warnings information will significantly reduce the cost of an update and also enable an incremental development. This paper will also describe how field loadable databases can be used in aircraft.

1 Introduction

When the Gripen development started in 1982, Test and Monitoring requirements were not the first thing in our focus. The product specification was very detailed on weapon functionality, while embedded test functions was covered in a few sentences. Test and monitoring was by then not neglected, they were just treated like something obvious. The maintenance concept for the predecessor, JA37 Viggen, was a success and thus requirements like "Performance better or equal of that for JA37" could be found. Despite the lack of detailed requirements, experienced personnel

were allowed to develop the monitoring functions from the beginning.

Embedded test, monitoring and recording functions were developed in parallel with tactical functions in the aircraft. With three processors, one MB of code memory and three Mil-STD 1553B data buses, the systems computer was well equipped for the task. As always, the requirements for new functions grew faster than hardware upgrades permitted. Limited computer capacity pushed the development towards highly optimized designs and approaches.

Many tools were developed for various purposes, tools that had great importance for further development but little integration with other tools. To avoid further diversification, common processes were defined for the company.

Today (2011), there are four Gripen versions in service, Single seat 39A and 39C, Dual seat 39B and 39D. Customer specific variants and upgrades exist for both serial aircraft and test aircraft. New variants are under development and new functions are delivered for flight test each month. Processes for systems and software development are now working as intended and roles and responsibilities have been clarified. For monitoring, test and recording, all responsibilities have been allocated to one department with the result that methods and tools have been standardized and are better integrated. Meanwhile, the defense budgets have decreased drastically the last decade and there is no longer room for any major redesign unless cost reduction can be proved in a short term. The issue will then be to improve the performance on legacy systems without expensive updates. See also [1].

2 Terms and Definitions

2.1 Central Functional Monitoring

Central Functional Monitoring provides the aircraft system with the following functionality:
- Functional monitoring of the aircraft system including logic for primary and secondary failure warnings.
- Failure warning generation
- Display of failure warnings on panels and text displays.
- Display of flight manuals.

2.2 Local Functional Monitoring

Local Functional Monitoring at subsystem level handles fault detection. In non degraded modes, all detected failures are passed to central functional monitoring where further processing is done.

3 Problems

3.1 Complexity

A main difference compared to older systems was that in the old federated architectures, all systems had their own failure modes and typically one or a couple of warning lamps connected to that system. A problem with that was that a failure warning in a supply system could cause so many side effects that the real failure warning would be obscured by a myriad of flashing warning lights, see Fig. 1. The problem to solve was how in a complex failure warning pattern pick out the primary fault and display that in a way that would give the best help to the pilot.

Having over 500 distinguishable failure modes, no pilot would keep all required actions in mind. A flight manual mode was developed early in the project and is now for 39C complete

in the meaning that there is one specific flight manual for each failure warning.

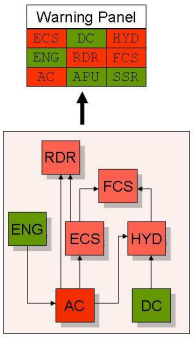


Fig. 1 Conventional Warning System

In an integrated system, there are easier ways to control the warnings displays.

The first effort to be done in 39A was to distinguish between primary faults and secondary faults. A suppression function on the caution and warning panel blocked the lamps for systems secondarily affected and the text info on the head down displays differed between primaries and secondaries for all flight critical failures. In 39C this function was extended to cover also failure warnings in the tactical system that would impact on mission capability.

3.2 Development cycle

Another problem was the time required to develop a new warning, as it required a great effort of integration and verification. To keep information in aircraft consistent with flight manuals and maintenance documentation with paper documentation and manual inspections was extremely time consuming.

It was early decided that we had to use tools to support the development. With all information stored in a database, the same source could be used to generate documentation and data for the avionics computers. In a later step, functionality to support integration between subsystems was added.

3.3 Time to customer

With embedded functions, new releases to customers are very expensive as the entire software has to be certified. By making the warnings system parametric driven, it is now possible to separate the warnings and flight manual data from the resident software and instead having it field loadable in a separate file system.

4 Approach

How do you get a good partitioning of a complex functionality involving several organizations?

A systematic approach to get maintainable software architecture led to the so called Warning Triangle.

In the Warning Triangle there are clearly distinguished layers that interact with clearly defined interfaces.

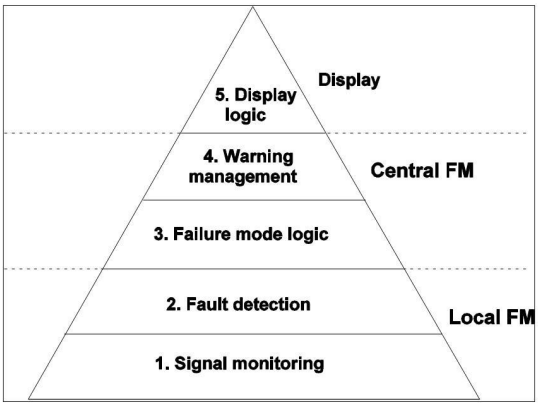


Fig. 2 ”The Warning Triangle”, principle layout for functional monitoring.

Figure 2 illustrates the different layers in functional monitoring:

• Functional level 1, in the bottom is signal monitoring. That is the most basic function that with a very minimum of involved functionality indicates a state that can serve as a warning. Typical examples are lamp indications that are connected to a basic sensor, like fire warning and heat sensor or low fuel warnings. A very important aspect is to monitor unexpected states that are not considered faults, like “gear not out at landing approach”. The signal monitoring is also used to cover the most critical failure warnings to make the failure mode presentation available even in case of a computer fault.

• Functional level 2, Fault detection is distributed and shall be done as close to the source as possible, uses combinations of the monitored signals to detect abnormal conditions or faults.

• Functional level 3, Failure mode logic is a central function that determines what is primary and secondary. It evaluates abnormal conditions or faults, detected by local functions, to see if it shall be considered a system level failure mode. Transients are not considered as failures.

• Functional level 4, Warnings management, is the core of central functional monitoring functions. It determines when failure warnings shall be given for reported failures and when warnings shall be deactivated. It also supplies data to display, I e Flight manuals.

• Functional level 5, Display logic is a special entity as there are several ways to display the information and the mode selection on the text displays are unsynchronized with how the warnings appears and disappears.

The responsibility of Central FM covers layer 3, 4 and 5. Local FM, on equipment or subsystem level covers the bottom two layers.

5 Current Design

5.1 Overview

The Gripen warnings system, according to Fig. 3 consist of

- Two master caution lamps with acknowledgment button.
- Audio warning
- Caution and Warning Panel, CWP
- Displays
- Local monitoring functions
- Central functional monitoring
- FM-database (Field loadable)

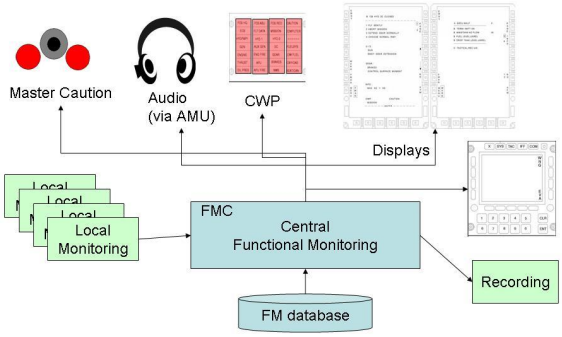


Fig. 3 Overwiev

These components ensure that failure warnings are presented, first with alert (audio and light) to capture attention. A system oriented CWP lamp points out the problem area and the text displays gives further information and decision support.

Redundant displays and distributed logic ensures that no single fault in the central function can prevent failure warnings to be displayed. For more detailed description about the design drivers, see [2].

5.2 Workflow

With shared responsibility between local and central functionality, the subsystem experts will focus on the fault detection and system behavior. The detection capability is implemented as Continuous Built in Test, CBIT, functions in the subsystems. Normally the detection function is limited to one subsystem or line replaceable unit, LRU, for each warning, but when a subsystem has several LRUs interacting, there might also be local failure mode logic, between level 2 and level 3 according to Fig. 2.

The technical manager for the central functions will together with subsystem managers and the test pilot assign a name for the corresponding failure warning. This can sometimes be a bit tricky as the name has to be limited in size and clearly interpretable at a short glance. The failure warning shall primarily identify the affected functionality.

The flight manuals that are shown on the displays are authored by a team including technical manager for central functions, subsystem managers and pilots.

The information regarding failure warnings are continuously entered in a database. When the work is complete a specification is automatically generated from the FM-tool, see Fig. 4. One specification is issued for each affected system. These specifications will then be inspected and finally approved. When all specifications are approved, software for the aircraft is generated from the same source as the specifications.

The generated software is field loadable to the aircraft to enable the flight manuals to be updated at a very short notice.

As the software is generated from the same source as the specifications, no further inspections of the software is necessary but the consistency between code and specification has to be verified. This is done by another tool, which is developed independent of the FM-tool, and is described in section 5.4.

5.3 FM-tool

As all development is tool supported, new and altered failure warnings are easily administered using the by Saab Aeronautics developed FM-tool.

In the FM-tool, several parallel configurations and systems can be handled.

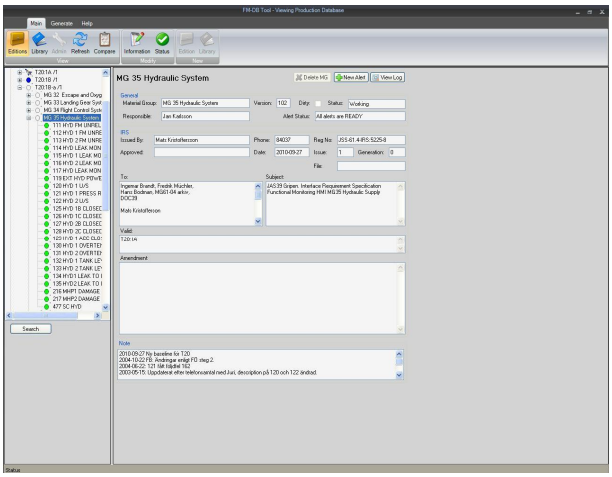


Fig. 4 FM-tool system view

Warnings are defined per system and the system structure is similar to the ATA chapters, see Fig. 5.

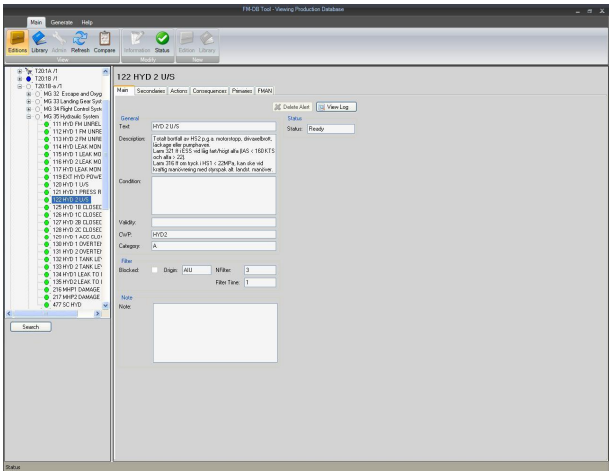


Fig. 5 Main view

For each warning, primaries and/or secondaries are defined. Based on all warnings, complete fault trees are generated and displayed in the tool, see Fig. 6.

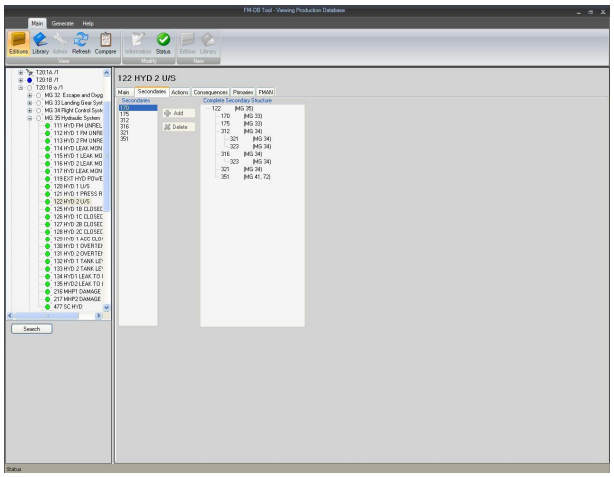


Fig. 6 Fault tree view

Flight manuals consist of a list of recommended pilot actions and a list of consequences. Each failure has its set of actions and consequences but the final display is a synthesis of actions and consequences for the actual fault in combination with secondary failure warnings see Fig. 7 and Fig. 8.

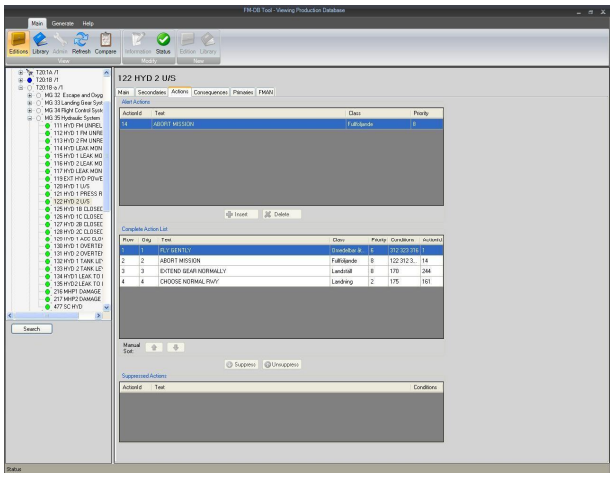


Fig. 7 Action view

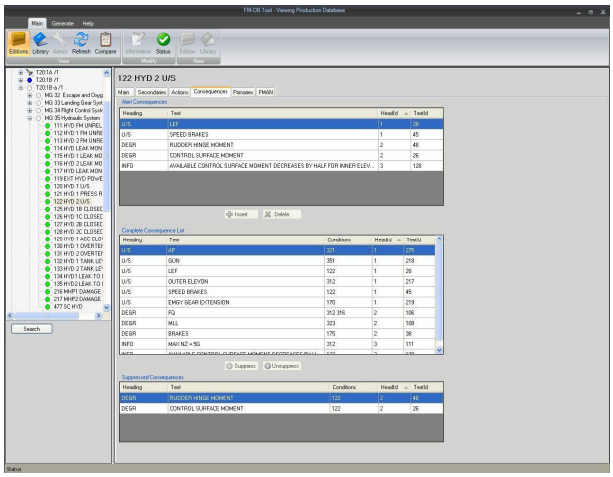


Fig. 8 Consequences view

In some cases the sum of all actions and consequences gives more information than can be displayed on a screen page or it is just too much for the pilot to cope with. There is a priority function in the tool that enables the user to manually sort out the most important information. The result can then be viewed exactly as it is shown in the cockpit, see Fig. 9.

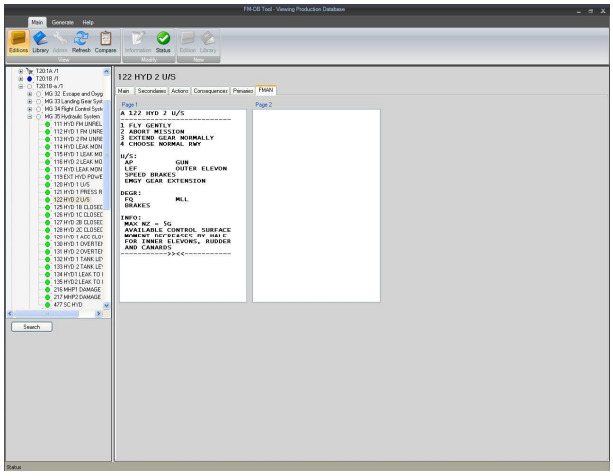


Fig. 9 Flight Manual view

5.4 FM-Verifier

As manual code inspection and testing can be very time consuming, the FM-Verifier was developed to ensure consistency between software and documents.

The FM-Verifier converts specifications to code and compares this code with the field loadable data generated from the FM-tool. To pass the verifier, the data from the two data sets has to be identical.

As one specification is issued for each subsystem, there will be about 30 documents that have to be checked against the code. As only documents affected by the last changes are reissued, the complete set of document dates can vary much in between. It is therefore important to verify that the “old” documents still are valid.

5.5 FM-Viewer

The Flight manual view is useful but it does not give the proper dynamic behavior.

FM-Viewer is developed by Saab Aeronautics to display the behavior of the failure warnings from the pilot point of view.

The tool will at startup ask for a software configuration. Then a field loadable file, exactly the file that is used in aircraft, is loaded. Fig. 10 shows the three displays as they can be seen in the cockpit. In the FM-viewer several warnings can be set in the same time and the displays will show how the failure pattern change as warnings are added or removed.

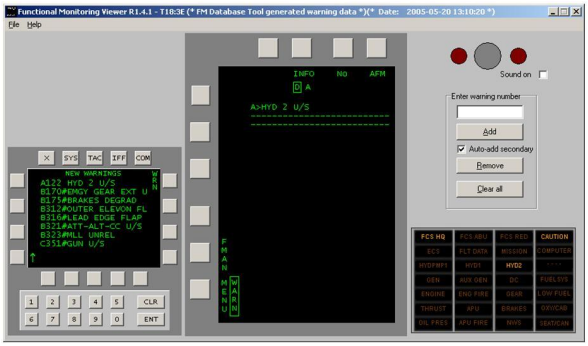


Fig. 10 FM-Viewer

By default only primaries are displayed, as the central display in Fig. 10 shows. The complete list of failure warnings, including all suppressed secondaries, see the left display above, can be manually selected by the pilot.

6 Conclusion

The tool supported process has been a great success. First of all, manual errors in the process that earlier resulted in discrepancies between documentation and displayed information has been eliminated, which has significantly increased the pilot confidence in the warnings system.

The visualization of fault trees has resulted in a better understanding of complex failure modes.

The FM-Viewer has been used as a simple desktop trainer for the pilots. It has also been used by technical staff when analyzing certain fault event.

Another important aspect is that the toolset significantly reduced development time and cost.

Our experience tells us that we are on the right track and we are following this method by introducing similar tools also for recording and test functions.

References

[1] Fransson T., ".Integrating development- and support tools for PHM in Saab 39 Gripen". *IEEE Aerospace Conference 2006.*

[2] Andersson S., Peterson M., "Att skapa uppmärksamhet vid presentation av central funktionsövervakning", Thesis Malardalens Hogskola FLYG/2003/137/10/FS-Av/O

3rd CEAS Air&Space Conference
21st AIDAA Congress

Platform based Realization of Safety Relevant Systems

S. Hesse, R. Reichel and S. Görke
Institute of Aircraft Systems (University of Stuttgart), Germany

Keywords: *Redundancy, Fault Tolerant, Flexible Platform, Simplex minded Functional Design*

Abstract

Since the year 2003 the Institute of Aircraft Systems is developing the technology Flexible Platform. Based on the approach of a clear segregation between system functions (applications) and their fault tolerant realization (Platform Instance), this technology enables a system designer to develop safety relevant systems (like fly-by-wire) by the "simplex minded functional" design of the application (e.g. Flight Control Laws) and the design of the Platform Instance by specialisation of the Flexible Platform. Within the context of this paper, a platform has to be seen as a flexible hardware-, communication- and software architecture of a distributed Embedded System. The key element is a generic Middleware, a software architecture, which comprises all functions for the fault tolerant operation of a Platform Instance. The technology Flexible Platform has shown its suitability and high degree of reuse in automotive applications like SPARC as well as in the aeronautic applications LAPAZ and SAFAR.

1 Introduction

Safety relevant systems such as fly-by-wire systems are very complex. This is due to the high degree of redundancy, functional complexity, heterogeneity, real time requirements and due to the fact, that mechanisms of fault tolerance are sometimes closely linked with the applications. Up today, this results into very application specific realisations. In consequence, each new type of aircraft leads to a widely new development of a fly-by-wire system incl. system- and redundancy-management. Especially the development of this management is very complex and represents a major share of overall development cost.

2 Platform based System Design - Basics

This article presents an approach for the realization of complex, safety relevant systems based on a clear segregation between system functions (applications) and their fault tolerant realization (the so called Platform Instance). System functions will be implemented as applications into the Platform Instance. An API (Application Programmable Interface) is used as

interface between Platform Instance and application. Due to this API, the Platform Instance appears as a virtual simplex system from scope of the application. Therefore redundancy and fault tolerance are almost transparent to the system function. The essential technology within this approach is the Flexible Platform. A platform in this case has to be seen as flexible hardware-, communication- and software architecture of a distributed Embedded System including sensors, actuators and in-/output devices. Besides a flexible hardware network, a key element of the Flexible Platform is the generic Middleware, which comprises the communication of the signal- and network domain, ensurance of the Byzantine Agreement, failure management as well as the operation management of a Platform Instance. The technology Flexible Platform enables development of a new Platform Instance via specialisation - by means of configuration - of the generic Middleware. This is essential to reduce the development effort of new safety relevant systems such as fly-by-wire systems significantly.

3 Flexible Platform

3.1 Safety Related Requirements

The "pure" functionality, like Flight Control or Navigation, is represented by System Functions (SFu). A System Function is characterized by its specific Transfer Function between the in- and output vector of the system (e.g. aircraft) and by its safety related requirements. In the following typical safety related requirements of a System Function will be specified.

SFu(x) is one of several System Functions, which are implemented within a Platform Instance. The state of SFu(x) is characterized by the tuple (SFu(x), z_{FA}).

(SFu(x),z_{FA}=k) : SFu(x) operates correctly degradation level = Deg..

k ϵ {Deg0,Deg1,..}

k ≠ DegN

(SFu(x),z_{FA}= z_{fo}) : SFu(x) doesn't operate correctly and at least one of its safety related output signals has not defaulted to a safe value (i.e. it is „out of control")

$$P(SFu(x),z_{FA}) \neq k \leq P_{f,max} \quad (1)$$

$$P(SFu(x),z_{FA}) \neq z_{fo} \leq P_{fo,max} \quad (2)$$

$$10^{-5}/h \leq P_{f,max} \leq 10^{-11}/h \quad (3)$$

$$10^{-8}/h \leq P_{fo,max} \leq 10^{-11}/h \quad (4)$$

3.2 Platform Elements

Key elements of the Flexible Platform are a flexible hardware network and the generic Middleware.

3.2.1 Flexible Hardware Network

Referring to the safety related requirements (Eq. (3) and Eq. (4)) a platform has to feature a highly flexible hardware network. Thus the Flexible Platform is not constrained to a specific hardware, but comprises of the following general components.

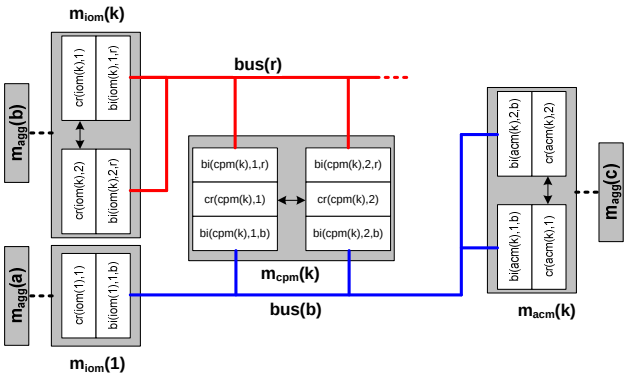


Fig. 1 Hardware components of a Platform Instance

bus(x): Bus System

Every Platform Instance comprises two independent bus systems, which could be internally redundant. Typical bus systems could be FlexRay, AFDX, etc.

m_{cpm}(k): Computing Module

Every m_{cpm}(k) of a Platform Instance is internally redundant (duplex), thus it comprises two lanes: la(1), la(2). Each lane comprises a processing core cr(cpm(k),la(i)) incl. power supply and two independent bus interfaces bi(cpm(k),la(i),bus), one per bus.

m_{iom}(k): Input/Output Module

A m_{iom}(k) is internally simplex or duplex, thus it comprises of just one lane la(1) or two lanes la(1), la(2). Each lane comprises a processing core cr(iom(k),la(i)) including power supply, a bus interfaces bi and interfaces for IO-purposes and the attachment of external Aggregates.

m_{acm}(k): Actuator Control Modul

Each m_{acm}(k) is internally duplex, thus it comprises two lanes: la(1), la(2). Each lane comprises a processing core cr(acm(k),la(i)) including power supply, two independent bus interfaces bi and interfaces for actuator control and safe passivation of the actuator in case of failures.

m_{agg}(k): Aggregates

Aggregates are sensors and "dumb" actuators, which are attached via Input/Output Module to the Platform Instance (not via bus system).

3.2.2 Generic Middleware

The module comprehensive, generic Middleware of the Flexible Platform hides the complexity of the Platform Instance making it appear as a virtual simplex system to the applications.

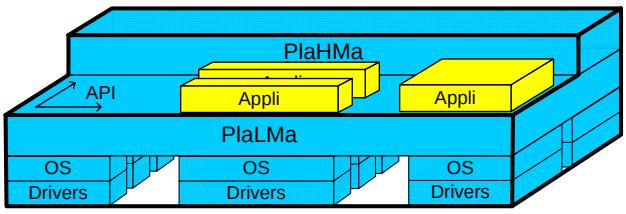


Fig. 2 Generic Middleware

In general a Middleware is an application-independent software (software layer), which manages the interaction between applications in a way, that complexity and infrastructure of an application is transparent to each other [1]. In the context of a redundant platform, a Middleware comprises further topics that are:

- Communication within the platform
- Fault tolerant reduction of redundant data
- Fault detection and assignment of detected faults to granulates
- Reliable Broadcast
- Establishing consensus on all levels
- Performing reconfigurations

This means the Middleware of the Flexible Platform provides not only an abstraction of a complex communication (e.g. between sensors, actuators and application), but also manages the redundant platform core. In the simplest case, a platform instance is running, even without any application signal. That means the Middleware of the Flexible Platform comprises a complex (redundancy) management. Therefore a clear structuring of the Middleware with clear interfaces between the various Middleware elements is required. At first the Middleware is structured into:

Platform-Highlevel-Management (PlaHMa)

- Management of (platform) internal communication
- Reallocation of functions
- Establishing consensus on platform level

Platform-Lowlevel-Management (PlaLMa)

- Performing (platform) internal communication
- Establishing consensus on module level

In the following this article focuses on PlaLMa, especially on the communication of the signal domain.

One challenge for the Flexible Platform is the heterogeneity of the communication, which comprises interfaces like analogue ports and CAN-busses, but also FlexRay and AFDX. So for a generic Middleware a uniform data presentation and data handling of these various interfaces must be assured. This comprises all mechanisms for the detection of faulty

signals/data, assignment of detected faults to granulates as well as reduction of redundant signals into virtual simplex signals and providing these simplex signals including the related status information to the application. For this purpose a layer model for the signal domain was defined.

Within a module this layer model consists of seven layers.

Tab. 1: Layers of Signal Domain

Index	*Layer*
L7	Application Layer
L6	Presentation Layer
L5	Redundancy Layer
L4	Granulation Layer
L3	Unification Layer
L2	Data Link Layer
L1	Physical Layer

- **L1** und **L2:** essentially follow the OSI Model [2].
- **L3**: Standardization of presentation of all heterogeneous signals/data, status and related status information
- **L4**: Assignment of data packages of L3 to source granulates, source granulates could be: a m_{agg} (or part of m_{agg}) or a cluster of several m_{agg}
- **L5**: Reduction of redundant data to virtual simplex data including efficient failure monitoring
- **L6**: Adjustment of data representation for the application
- **L7**: Providing the API for the application

By the expansion of this model over the whole platform, the model can be interrupted between Layer L4 and L5 as well as L5 and L6. This results in the following Layer based platform wide communication (signal domain).

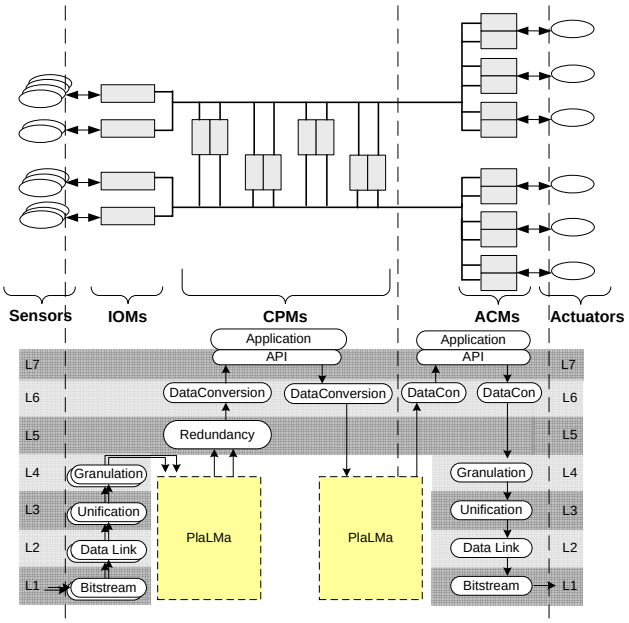


Fig. 3 Layer based platform wide communication (signal domain)

3.3 Platform Mechanisms[1]

In the following some platform mechanisms are discussed in detail to illustrate the operation of the Flexible Platform. This is based on the Platform Instance in Fig. 5.

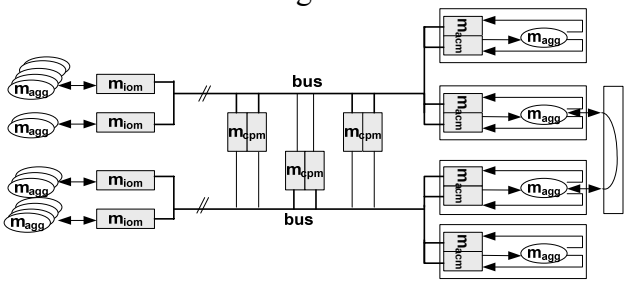


Fig. 4 Simplified Platform Instance

3.3.1 Consensus within the Platform Core

Generally the platform has a hierarchical structure of redundancy: aggregates are redundant, modules may be redundant internally – duplex, triplex or quadruplex modules – as well as redundant to each other. Redundancy on all hierarchies requires a consensus between all replicas. Further on, only consensus at the

[1] Within this section (incl. subsections): $m = m_{cpm}$

highest platform level, namely between the $m_{cpm,}$ will be discussed.

3.3.2 Data Transfer via Frame fd

The transfer of data is performed by so called Duplex Frames fd, consisting of two redundant frames f.

- fd(m(s),bus(x)) is a Duplex Frame with the following characteristics:
  - Frame f(m(s),la(1),bus(x)) is generated by module m(s), Lane 1 and send via bus(x).
  - Frame f(m(s),la(2),bus(x)) is generated by module m(s), Lane 2 and send via the same bus(x).
  - The payload pf of frame f has a time stamp.
  - pf(m(s),la(1),bus(x)) = pf(m(s),la(2),bus(x))

- A received fd is correct, if:
  - Both f are correct concerning time (time domain),
  - Both f are correct concerning code (code domain),
  - Both f are identical.
- Every m(s) sends its fd via both bus systems: bus(b), bus(r).
- A receiving m(r) accepts only correct fd.
- Is m(r) receiving a correct fd via one of the two bus systems, the data transfer was correct.

3.3.3 Failure-Hypothesis

Essential for the design of mechanisms, which ensure the consensus between the non-faulty m within platform core, is the failure hypothesis of the platform core. This hypothesis is defined by the following characteristics:

- A faulty lane la(m(k),la(i))[2] is not able to manipulate the operation of its non faulty opposite lane la(m(k),la(¬i)).
- A faulty bi(m(k),la(i),bus(x))[3] is just able to manipulate bus(x). This manipulation can be of arbitrary kind.
- A faulty bi(m(k),la(i),bus(x)) is not able to manipulate both redundant frames f of an arbitrary fd in a way, that this fd appears correct according to the definition above.
- A faulty module core cr(m(k),la(i))[4] may be faulty in an arbitrary way, but it merely able to manipulate the bus interfaces of its lane bi(m(k),la(i),bus(x)) and bi(m(k),la(i),bus(¬x)) with the following characteristics.
  - Frames f are send with non correct payload pf.
  - No frame is send.

3.3.4 Consensus between all non faulty Modules

In case of an arbitrary faulty granulate the platform structure, bus communication and the defined data transfer assure:

- A faulty module (module with a faulty granulate) is not able to corrupt the cyclic communication between the non faulty modules. Hence, consensus between the non-faulty modules can be assured by a One Round Data Exchange.
- A faulty module core cr(m(s),la(i)) is recognized by all non faulty modules in a consistent way. Thus, all non faulty modules evaluate the same value for the payload of the faulty module.
- For the payload of a sending module m(s) with a faulty bus interface bi(m(s),la(i),bus(x)) all non faulty modules evaluate:
  - The same value and
  - The value of the sending module m(s).

[2] Lane la(i) of module m(k)

[3] Bus interface of module m(k), lane la(i), bus(x)

[4] Core of module m(k), lane la(i)

This definition of the consensus assures, with respect to the failure hypothesis above, that each decision within the platform core is consistent between all non faulty modules.

4 Platform based System Design – Design Process

The technology Flexible Platform bases on the clear segregation of functionality on the one and management (operation-, network-, failure-, redundancy management) on the other hand. By this technology a system designer is able to develop a safety relevant system by

- Designing system function: The system functions SFu(m) are realized as applications and allocated into the Platform Instance, which appears as a virtual simplex system from scope of the application. So by designing the applications you can assume a simplex system (simplex minded functional design).
- Designing Platform Instance: The Platform Instance is arising by the specialization of the Flexible Platform (platform instantiation).

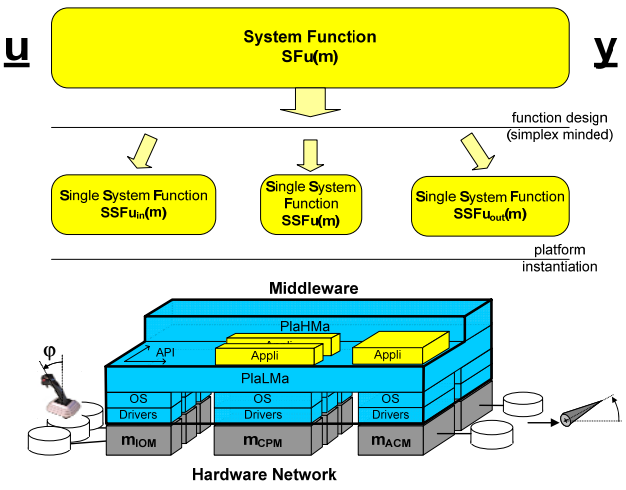


Fig. 5 Design Process

5 Realized Safety Relevant Systems

5.1 SPARC[5]

The goal of the EU- founded project SPARC was to substantially improve traffic safety and efficiency for heavy goods vehicles using intelligent x-by-wire technologies in the power train. To prove the capability of the approach x-by-wire systems basing on the technology Flexible Platform were integrated into two different power trains and a passenger car without any mechanical backup.

$$P(SFu(Steering),z_{FA})\neq k \leq 0{,}5*10^{-8}/h \qquad (5)$$

$$P(SFu(Braking),z_{FA})\neq k \leq 0{,}7*10^{-8}/h \qquad (6)$$

$$P(SFu(Steering),z_{FA})=z_{fo} \leq 0{,}5*10^{-8}/h \qquad (7)$$

$$P(SFu(Braking),z_{FA})=z_{fo} \leq 0{,}7*10^{-8}/h \qquad (8)$$



Fig. 6 SPARC – X-by-Wire Passenger Car



Fig. 7 SPARC – X-by-Wire Power Train

[5] Secure Propulsion using Advanced Redundant Control

Within the project SPARC the Institute of Aircraft Systems developed three fault tolerant, redundant Platform Instances (for two different power trains and a passenger car) basing on the institutes Flexible Platform, implemented 16 applications into one Platform Instance and supported the verification and the driving tests on test track.

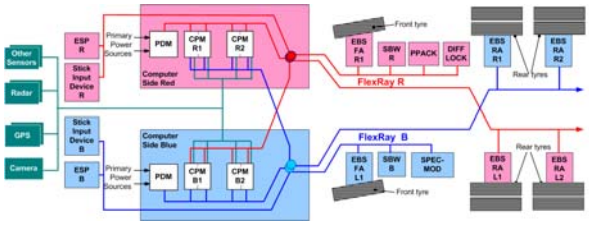


Fig. 8 SPARC – Platform Instance (Power Train)

5.2 LAPAZ[6]

While nowadays aerial work missions are predominantly manually piloted, the LAPAZ project aimed on developing a versatile utility aircraft with increased performance (e.g. flight path accuracy) and enhanced operational scope on the one hand and significantly reduced pilot work load on the other hand. Satisfying these mission requirements requires full control authority of the developed automatic flight control system (AFCS) which renders it safety critical in case of failures.



Fig. 9 LAPAZ - In Flight Evaluation

- During manual flight the pilot controls the aircraft via the existing mechanical control link.
- During automatic flight the AFCS controls the aircraft via actuators, which are attached to the mechanical control link.

$$P(SFu(AFC), z_{FA}) \neq k \leq 0{,}5 * 10^{-6}/h \qquad (9)$$

$$P(SFu(AFC), z_{FA}) = z_{fo} \leq 0{,}5 * 10^{-6}/h \qquad (10)$$

Within the project LAPAZ the Institute of Aircraft Systems developed the fault tolerant, redundant Platform Instance basing on the institutes Flexible Platform, implemented the Auto Flight Control Laws, verified the AFCS via institutes Hardware in the Loop Test Rig and supported the aircraft integration and in flight evaluations.

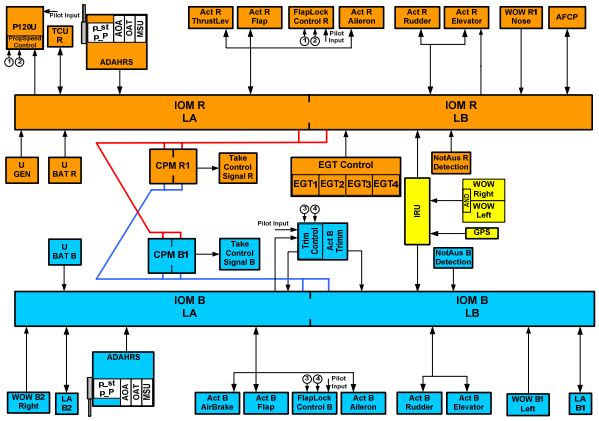


Fig. 10 LAPAZ - Platform Instance

5.3 SAFAR[7]

The goal of the EU- founded project SAFAR is the development of future avionics architectures, which affords an intuitive and safe handling of an aircraft, for general aviation. It aims on a significantly reduction of pilot work

[6] **L**uft**A**rbeits**P**lattform der **A**llgemeinen **Z**ivilluftfahrt

[7] Small Aircraft Future Avionics Architecture

load and an increased safety in all flight phases including start and landing. Therefore the aircraft should provide an Easy-Handling-Characteristic as well as active flight envelope protection. Basis of this functionality is a full fly-by-wire system, which meets the requirements of general aviation concerning mass, energy supply and costs.

$$P\left\{\begin{matrix}\left(SFu(FCL), z_{FA} \neq k\right) \cup \\ \left(SFu(INAV), z_{FA} \neq k\right)\end{matrix}\right\} < 0.5*10^{-7}/h \qquad (11)$$

$$P\left\{\begin{matrix}\left(SFu(FCL), z_{FA} = z_{fo}\right) \cup \\ \left(SFu(INAV), z_{FA} = z_{fo}\right)\end{matrix}\right\} < 0.5*10^{-7}/h \qquad (12)$$



Fig. 11 SAFAR – Flight Demonstrator (Diamond DA42)

Within the project SAFAR the Institute of Aircraft Systems developed the fault tolerant, redundant Platform Instance basing on the institutes Flexible Platform and implemented the Auto Flight Control Laws. At this time the system is verified. Maiden flight and start of in flight evaluation is planned Oct. 2011.

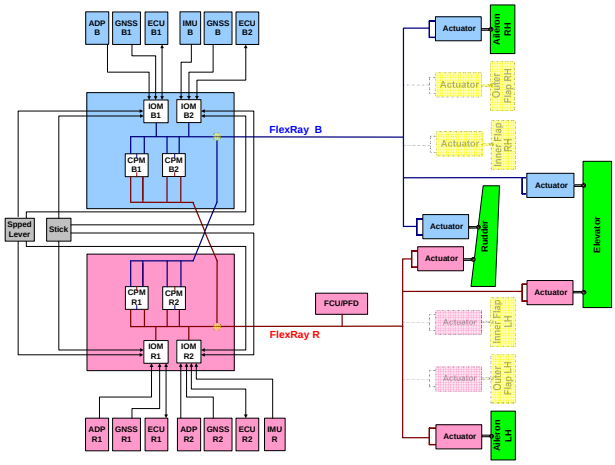


Fig. 12 SAFAR - Platform Instance

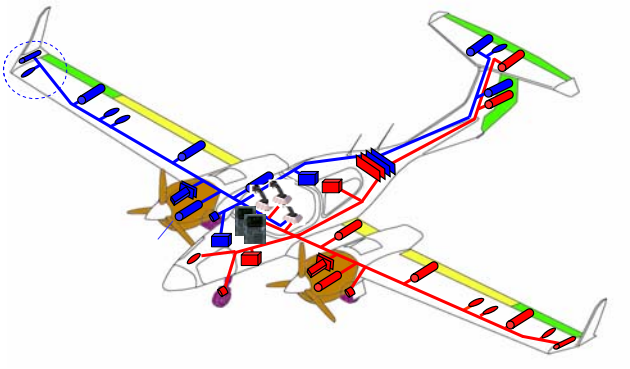


Fig. 13 SAFAR – Fly-by-Wire Aircraft Integration

References

[1] W. Ruh et al: Enterprise Application Integration. Wiley, 2001

[2] Zimmermann, H: OSI Reference Model−The ISO Model of Architecture for Open Systems Interconnection. IEEE Transactions on Communications. (1980) 28-4, ISSN 0090-6778, S. 425-432

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Simulation and Experiments for the Study of a Fault-Tolerant Hydraulic Actuator for Fly-By-Wire Helicopters

G. Di Rito and R. Galatolo
Dipartimento di Ingegneria Aerospaziale – Università di Pisa, Italy

Keywords: *Fly-by-wire, fault-tolerant systems, actuators, simulation, testing*

Abstract

The availability of accurate equipment models plays a key role in the project of fly-by-wire flight control systems, since crucial activities as closed-loop controls' design, performance analysis, integration studies, and the development of monitoring algorithms can be strongly enhanced if a high-fidelity simulation is used. This paper deals with the development and the validation of the model of a hydraulic actuator for fly-by-wire helicopter controls. The model provides a detailed simulation of the main actuator parts, and is capable to simulate the system dynamics with electrical and hydraulic failures. Simulation results are validated via experiments, by performing an extensive test campaign in terms of frequency and time responses. The validation, obtained in both normal and fail-operative conditions at servovalve as well as actuator level, allowed to highlight the importance of some model details for simulating relevant system nonlinearities, such as the actuator limit cycle oscillations.

1 Introduction

The use of modelling and simulation becomes more and more extensive in aircraft systems design. Most activities about closed-loop control design, performance analysis, integration studies are performed well before the first flight or on-ground rigging, thanks to simulation. As a consequence, there is a strong demand for accurate models able to predict the dynamic behaviour of on-board systems as well as the interactions between them. The use of simulation becomes crucial in the project of modern fly-by-wire aircrafts, whose redundant flight control systems (FCSs) are characterized by a high level of sophistication [1] [2]. In this context, many efforts are made for developing practical models of the FCS subsystems, with particular focus on safety-critical equipments, such as the flight control actuators [3] [4].

This paper deals with the development and the validation of the model of a hydraulic actuator for fly-by-wire helicopter controls. The main features of the model are described, and a comparison between simulation and experiments is provided, highlighting and discussing the importance of some model details for achieving a satisfactory accuracy of results.

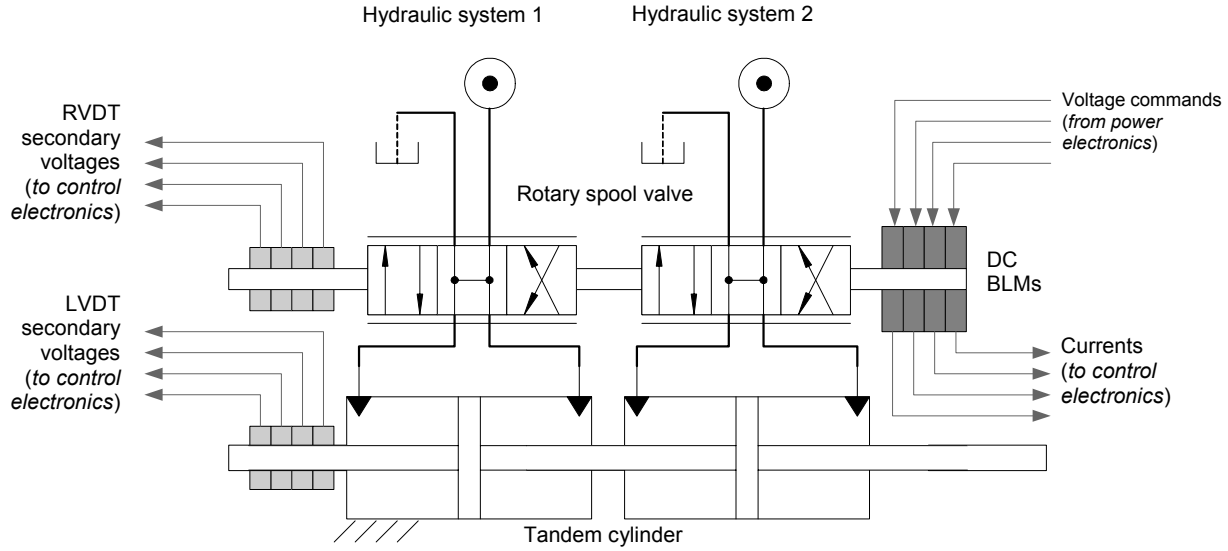


Fig. 1 - Actuator schematics.

2 Actuator model

2.1 System description

As schematically depicted in Fig. 1, the reference actuator is characterised by a complex redundant architecture that allows the system to be fault-tolerant (i.e. to maintain its functionality even after one or more failures). The equipment is essentially composed of

- a tandem hydraulic cylinder
- an eight-way rotary servovalve with unique spool shaft
- a limited-angle brushless motor for the direct-drive servovalve actuation, with four independent coils
- four LVDT (Linear Variable Differential Transformer) transducers for the actuator position sensing
- four RVDT (Rotary Variable Differential Transformer) transducers for the valve rotation sensing

and it is interfaced with two independent hydraulic power systems and a FCC (Flight Control Computer). The FCC is equipped with a quadruple PWM (Pulse Width Modulation) power electronics for driving the servovalve motor coils, and a quadruple FPGA (Field Programmable Gate Array) control electronics for implementing the actuator closed-loop control.

This equipment is representative of a FCS architecture composed of three FCCs, each one able to control one lane of all the four helicopter actuators. A further backup FCC acts as a "stand-by" device, i.e. it is activated on the fourth lane only when all the three main lanes are failed.

It is important to note that, though the quadruple electronics, only three lanes are active in the normal operating condition.

2.2 Main model features

Although the actuator model developed during the research includes the simulation of all system parts, only the model sections that demonstrated to be more relevant for the results' accuracy will be here presented. For details about other model sections (tandem cylinder, flow forces, LVDT and RVDT transducers), refer to [5] [6] [7] [8].

2.2.1 Servovalve motor and spool dynamics

In the reference application, the servovalve spool actuation is obtained by a limited-angle DC brushless motor [9] [10] with four magnet-coil couplings in torque-summing configuration, Fig. 2.

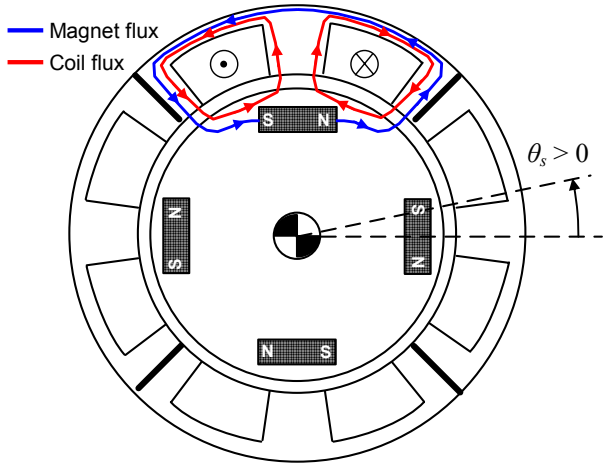


Fig. 2 - Quadruple DC brushless motor schematics.

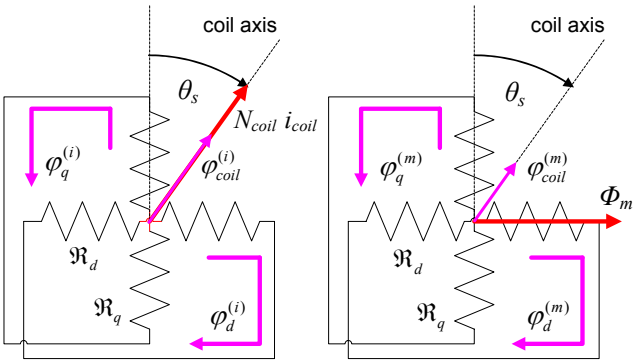


Fig. 3 - Magnetic circuits of a motor coil: (a) current-induced fluxes; (b) magnet-induced fluxes.

The coils are placed in the stator section of the motor with the coil axes that are radial with respect to the rotor, while the four permanent magnets are embedded in the rotor section and mounted with tangential axes.

To describe the servovalve dynamics, the motor torque acting on the spool shaft must be calculated. Let us consider a single magnet-coil coupling and assume that the four magnetic circuits are completely independent. The first step is to derive the magnetic flux linked to the motor coil, which can be expressed as the sum of two contributions: one related to the flux along the magnet direct axis (φ_d) and the other related to the flux along the magnet quadrant axis (φ_q), Eq. (1).

$$\varphi_{coil} = \varphi_d \sin\theta_s + \varphi_q \cos\theta_s \tag{1}$$

where θ_s is the motor shaft rotation.

As shown in Fig. 3, the direct and the quadrant fluxes can be obtained by linearly superimposing the effects of the coil and the permanent magnet, Eqs. (2)-(6).

$$\varphi_d = \varphi_d^{(i)} + \varphi_d^{(m)} \tag{2}$$

$$\varphi_q = \varphi_q^{(i)} + \varphi_q^{(m)} \approx \varphi_q^{(i)} \tag{3}$$

$$\varphi_d^{(m)} = \frac{\Phi_m}{\Re_d} \tag{4}$$

$$\varphi_d^{(i)} = \frac{N_{coil}\, i_{coil}}{\Re_d} \sin\theta_s \tag{5}$$

$$\varphi_q^{(i)} = \frac{N_{coil}\, i_{coil}}{\Re_q} \cos\theta_s \tag{6}$$

where N_{coil} is the number of coil windings, i_{coil} is the coil current, Φ_m is the magnetomotive force of the permanent magnet, while $\Re_d$ and $\Re_q$ are the magnetic reluctance along the direct and the quadrant axis respectively.

By substituting Eqs. (2)-(6) into Eq. (1), the magnetic flux linked to the coil is obtained:

$$\begin{aligned} \varphi_{coil} &= \varphi_{coil}^{(i)} + \varphi_{coil}^{(m)} = \ldots \\ &\ldots = \frac{N_{coil}\, i_{coil}}{\Re_d}\left(\sin^2\theta_s + \frac{\Re_d}{\Re_q}\cos^2\theta_s\right) + \frac{\Phi_m}{\Re_d}\sin\theta_s \end{aligned} \tag{7}$$

The coil flux expression allows to describe the dynamics of the current circulating in the motor coil, by means of Eq. (8),

$$V_{coil} = R_{coil}\, i_{coil} + L_{coil}\frac{\mathrm{d}\, i_{coil}}{\mathrm{d}\, t} + K_e\, \dot{\theta}_s \tag{8}$$

where the coil inductance (L_{coil}) and the coil back-electromotive force coefficient (K_e) are given by Eqs. (9)-(10).

$$L_{coil} = N_{coil}\frac{\partial \varphi_{coil}}{\partial i_{coil}} = \frac{N_{coil}^2}{\Re_d}\left(\sin^2\theta_s + \frac{\Re_d}{\Re_q}\cos^2\theta_s\right) \tag{9}$$

$$\begin{aligned} K_e &= N_{coil}\frac{\partial \varphi_{coil}}{\partial \theta_s} = \ldots \\ &\ldots = \frac{N_{coil}\,\Phi_m}{\Re_d}\cos\theta_s + \frac{N_{coil}^2\, i_{coil}}{\Re_d}\left(1 - \frac{\Re_d}{\Re_q}\right)\sin 2\theta_s \end{aligned} \tag{10}$$

The torque provided by a single coil can be now derived via Eq. (8), multiplying the coil

current to both terms of the equations, so that a power balance equation is obtained.

$$V_{coil}\, i_{coil} = R_{coil}\, i_{coil}^2 + L_{coil}\, i_{coil} \frac{\mathrm{d}\, i_{coil}}{\mathrm{d}\, t} + K_e\, i_{coil}\, \dot{\theta}_s \qquad (11)$$

The term at the right hand of Eq. (11) is the input electrical power, while those at the left hand are the Joule losses, the electrical power stored by the inductance, and the output mechanical power. The torque applied by the active magnet-coil couplings (n_{ac}) is thus given by Eq. (12).

$$T_m = \sum_{j=1}^{n_{ac}} K_{e_j}\, i_{coil_j} \qquad (12)$$

Finally, the momentum equation at the spool shaft dynamics can be written, Eq. (13),

$$J_s\, \ddot{\theta}_s = T_{flow} + T_m + T_{friction} \qquad (13)$$

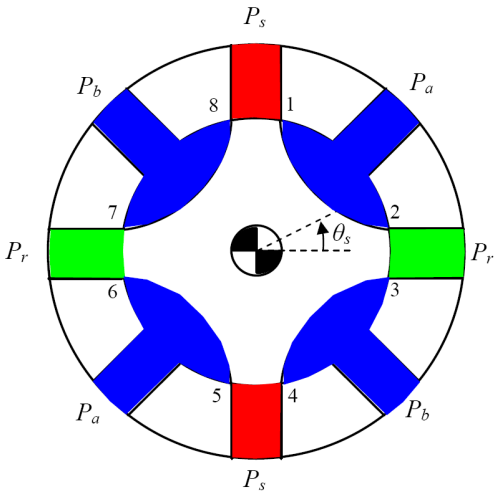
where J_s is the shaft inertia, T_{flow} is the torque related to the flow forces and $T_{friction}$ is the friction torque, modelled with the Stribeck function [11] reported in Eq. (14).

$$T_{friction} = -\left\{ T_C + T_S \exp[-(\dot{\theta}_s / \omega_S)^2] \right\} \mathrm{sgn}(\dot{\theta}_s) \qquad (14)$$

A particular attention has been dedicated to the tuning of the three parameters of the friction model (the Coulomb torque T_C , the Stribeck torque T_S , and the Stribeck velocity ω_S), since, as will be outlined in section 4, it demonstrated to be very important for satisfactorily reproducing the actuator response during low-velocity motion.

2.2.2 Servovalve flow

Another important model feature is the simulation of the flow through the rotary servovalve. Figure 4 schematically shows the orifices' arrangement in one of the two hydraulic units of the valve: the flow rates entering the actuator chambers (Q_a and Q_b) is known if the flows through the eight valve orifices are calculated, Eqs. (15)-(16).



Legend

P_s	Supply pressure
P_r	Return pressure
P_a	Pressure in the load port a (to cylinder chamber a)
P_b	Pressure in the load port b (to cylinder chamber b)

Fig. 4 - Rotary valve schematics.

$$Q_a = Q_1 + Q_5 - Q_2 - Q_6 \qquad (15)$$

$$Q_b = Q_4 + Q_8 - Q_3 - Q_7 \qquad (16)$$

The geometry and the pressure drop of the orifices 1, 2, 3 and 4 are respectively identical to those of the orifices 5, 6, 7 and 8; so we have

$$Q_1 = Q_5 = C_d\, A_a \sqrt{\frac{2|P_s - P_a|}{\rho}}\, \mathrm{sgn}(P_s - P_a) \qquad (17)$$

$$Q_2 = Q_6 = C_d\, A_b \sqrt{\frac{2|P_a - P_r|}{\rho}}\, \mathrm{sgn}(P_a - P_r) \qquad (18)$$

$$Q_3 = Q_7 = C_d\, A_a \sqrt{\frac{2|P_b - P_r|}{\rho}}\, \mathrm{sgn}(P_b - P_r) \qquad (19)$$

$$Q_4 = Q_8 = C_d\, A_b \sqrt{\frac{2|P_s - P_b|}{\rho}}\, \mathrm{sgn}(P_s - P_b) \qquad (20)$$

where ρ is the fluid density, C_d is the orifice discharge coefficient, P_s and P_r are the supply and return pressures of the hydraulic plant, P_a and P_b are the pressure in the actuator's

cylinder, while A_a and A_b are the cross-section areas of the odd and even orifices respectively.

Once defined the orifice port width w (along the longitudinal axis) and the spool radius r_s , the orifice areas are obtained via Eqs. (21)-(22), taking into account geometrical details such as the radial clearance (c_r) and the orifice angular underlap (θ_u), Fig. 5.

$$A_a = \begin{cases} w c_r & \theta_s < -\theta_u \\ w\sqrt{r_s^{\,2}(\theta_s+\theta_u)^2 + c_r^{\,2}} & \theta_s \geq -\theta_u \end{cases} \qquad (21)$$

$$A_b = \begin{cases} w\sqrt{r_s^{\,2}(\theta_s-\theta_u)^2 + c_r^{\,2}} & \theta_s < \theta_u \\ w c_r & \theta_s \geq \theta_u \end{cases} \qquad (22)$$

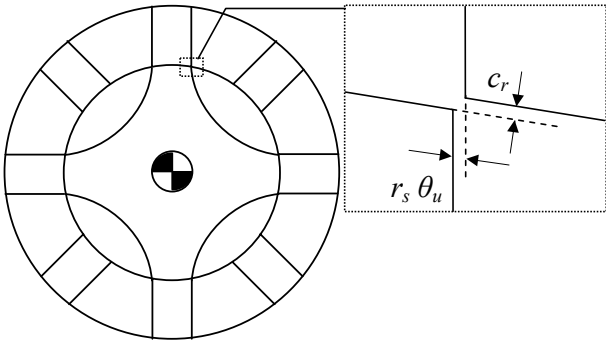


Fig. 5 - Geometrical details of the valve orifices.

It is worth noting that these geometrical details are essential for reproducing basic valve characteristics as the pressure gain or the leakage flow [5].

2.2.3 PWM power electronics

The coils' drive in the servovalve motor is obtained with a three-state PWM technique [9], i.e. the voltage to a coil is provided as a pulse-wave of pre-defined period (T_{pwm}) with three possible amplitudes, obtained by switching the DC voltage supply (V_{supply}).

As depicted in Fig. 6, this result is obtained by connecting the coil to a set of four power switches, in a circuit arrangement that is commonly referred as "H-bridge" [9].

Basically, when the switches S1 and S4 are closed and the others held opened, $V_{AB} = V_{supply}$; while if S2 and S3 are closed and the others opened, the same voltage level is applied, but with opposite sign.

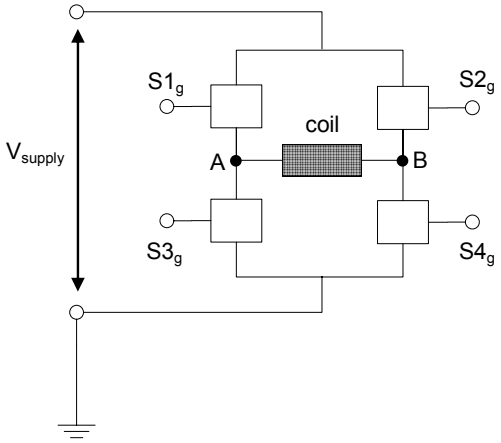


Fig. 6 - H-bridge for a motor coil.

There are six possible combinations of two out of four active switches in the H-bridge, but only three of them are actually used. Table 1 reports all the possible switching logics, and it can be noted that the last two logics must be always inhibited, to avoid a short circuit of the voltage supply.

Table 1 - H-bridge switching logics

State	$S1_g$	$S2_g$	$S3_g$	$S4_g$	V_{AB}
"Zero"	Off	Off	**On**	**On**	0
"Pos"	**On**	Off	Off	**On**	V_{supply}
"Neg"	Off	**On**	**On**	Off	$-V_{supply}$
Not used	**On**	**On**	Off	Off	0
Inhibited	**On**	Off	**On**	Off	-
Inhibited	Off	**On**	Off	**On**	-

Legend: Gate on = switch closed ; Gate off = switch open

The working principle of the PWM drive can be described by means of Eqs. (23)-(26): the H-bridge state is defined by comparing a target voltage signal (V_{pwm}) with a reference saw-tooth signal (V_{ref}) with period T_{pwm} and amplitude V_{supply} . If the target voltage amplitude is lower than the reference voltage signal, the coil voltage is set to zero; otherwise, it is set to V_{supply} or $-V_{supply}$ depending on the target voltage is positive or negative, respectively.

$$V_{ref}(t) = V_{\text{supply}}\, c(t) \qquad (23)$$

$$c(t)=\begin{cases}(t-T_0)/T_{pwm} & \text{if} \quad T_0 \le t < T_0+T_{pwm} \\ 0 & \text{if} \quad t=T_0+T_{pwm}\end{cases} \tag{24}$$

$$T_0(t)=T_{pwm}\,\text{round}(t/T_{pwm}) \tag{25}$$

$$V_{AB}=\begin{cases}V_{\text{supply}} & \text{if} \quad |V_{\text{pwm}}|>V_{\text{ref}} \quad \text{and} \quad V_{\text{pwm}}>0 \\ 0 & \text{if} \quad |V_{\text{pwm}}|\le V_{\text{ref}} \\ -V_{\text{supply}} & \text{if} \quad |V_{\text{pwm}}|>V_{\text{ref}} \quad \text{and} \quad V_{\text{pwm}}<0\end{cases} \tag{26}$$

It must be outlined that, in the actuator model, the simulation of the PWM power electronics has been done by directly implementing Eqs. (23)-(26), without accounting for the electrical behaviour of the switches. This approach has been considered adequate for the purposes of the work.

2.2.4 Closed-loop controls

The actuator control is performed by three closed-loop controls, on the coil current, the servovalve rotation and the rod position respectively.

The position control is proportional, so that the valve demand signal θ_{dem} is obtained from Eq. (27)

$$\theta_{dem}=K_{ac}\,(x_{dem}-x_p) \tag{27}$$

where x_{dem} is the rod position demand and K_{ac} is the actuator closed-loop gain.

The valve rotation control, generating the demand for the current loop (i_{dem}), is composed of a proportional part (with gain K_{vcp}) for tracking performance, and a derivative part (with gain K_{vcd}) for stability augmentation, Eq. (28).

$$i_{dem}=K_{vcp}\,(\theta_{dem}-\theta_s)+K_{vcd}\,(\dot{\theta}_{dem}-\dot{\theta}_s) \tag{28}$$

Finally, the motor current control is again proportional (with gain K_{pwm}), Eq. (29),

$$V_{pwm}=K_{pwm}\,(i_{dem}-i_{coil}) \tag{29}$$

and it generates the target voltage for the PWM coil drive (Eq. (26)).

All the closed-loop controls are digitally processed at 1 MHz sample rate.

3 Test system

3.1 Experimental set-up

The flight actuator testing has been performed with the experimental set-up illustrated in Fig. 7, which is composed of:

- a hydraulic test rig (HTR);
- a hydraulic power plant (HPP);
- a three-lines distributor, for dealing out the hydraulic power provided by the HPP among the users[1];
- a stabilised electrical power supply (SEPS), for the 28 V_{dc} FCC supply;
- a break box, accessing all the signals exchanged between the FCC and the flight actuator;
- a high-frequency acquisition card (8 channels, 14 bit, 2.5 Msamples/sec per channel), for the test data record;
- a current sensing unit (CSU), equipped with Hall-effect sensors for the measurement of the servovalve motor currents;
- two couple of PCs networked via the Matlab-Simulink-xPC Target tools: the first one for the real-time closed-loop pressure control on the three HTR lines[2], and the other for the real-time generation of the actuator commands;
- a PC for the data acquisition plug-in card management in the LabView environment.

The plant also includes a HTR control unit, which monitors the flight actuator stroke, the pressures and the flow rates on the supply and return lines, implements an automatic pressure cut-off in case of any of the monitored quantities exceeds preset limits, and allows to control the load actuator (when used).

[1] I.e. the two lines of the flight actuator tandem cylinder and the one of the load actuator, which, though not used for this work, will be employed for the future research developments.

[2] Specific proportional pressure-reducing valves are included in the distributor manifold.

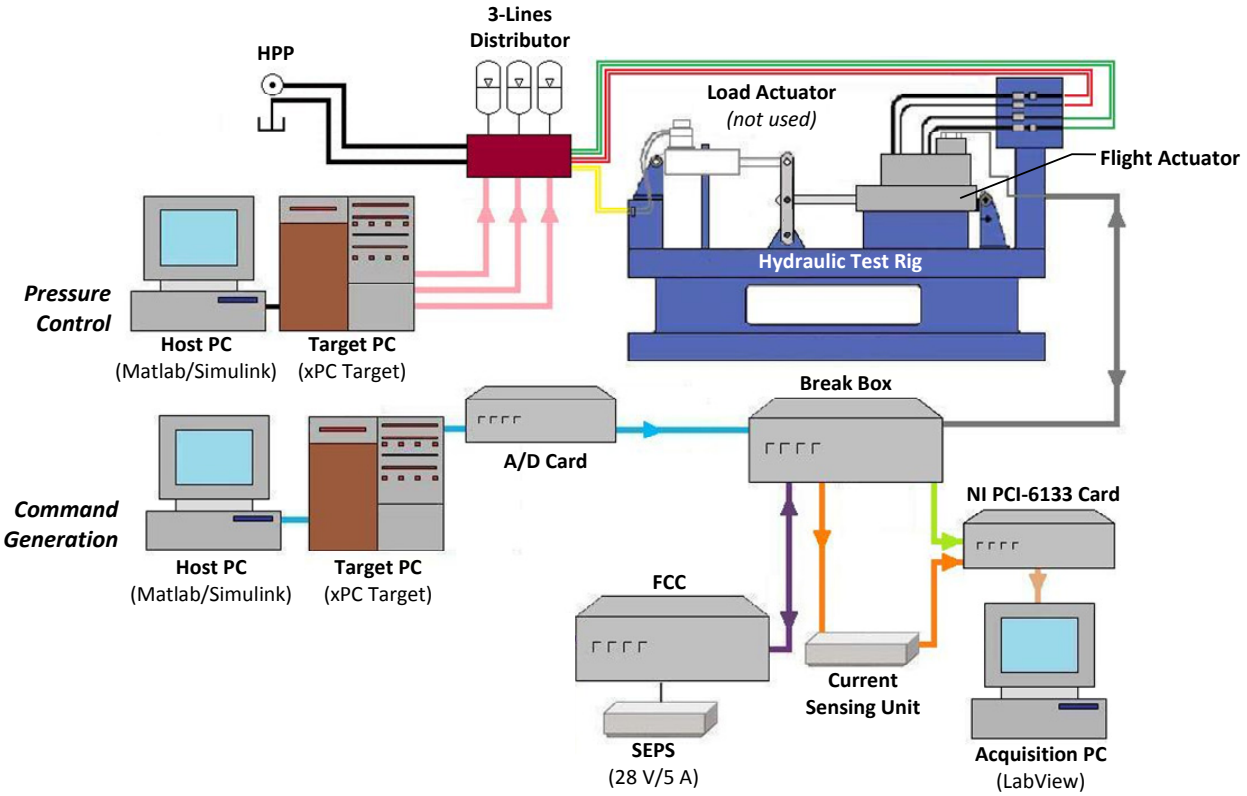


Fig. 7 - Experimental set-up for the actuator testing.

3.2 Test procedure and test plan

Two sessions of experiments have been planned for the flight actuator testing.
In the first one, addressing the characterisation of the actuator servovalve dynamics only, the hydraulic system is depressurised, and the control electronics is configured to exclude the position closed-loop control and to inject the command signal sent by the Command Generation PC (Fig. 7) as a valve rotation demand. In the second test session, the complete actuator dynamics is studied, so the hydraulic system is pressurised, and the control electronics, normally configured, receives the command signal as a position demand.

Before each test session, the actuator operating condition is defined (e.g. the open circuit of a servovalve coil is simulated by physically disconnecting the electrical wiring via the break box; while the hydraulic pressure loss is simulated via the pressure control PC).

During each test, six signals are acquired via the high-frequency acquisition card at 1 MHz sample rate: the command, the current in one of the servovalve coil, the two secondary voltages of one of the servovalve RVDTs, and the two secondary voltages of one of the actuator LVDTs. At the end of each test, the LVDT and the RVDT signals are elaborated in post-processing, by means of a model that simulates the demodulation process operated by the FCC, up to obtain the servovalve rotation and the actuator position signals.

Both step and frequency response tests have been performed, using input amplitudes that are typical of the actuator dynamics during autopilot flight control functions: 25% of the full-scale (FS) command for the servovalve tests, and 10% for the actuator tests.

4 Results and discussion

4.1 Frequency responses

Figure 8 reports the comparison between the simulation results and the experiments with

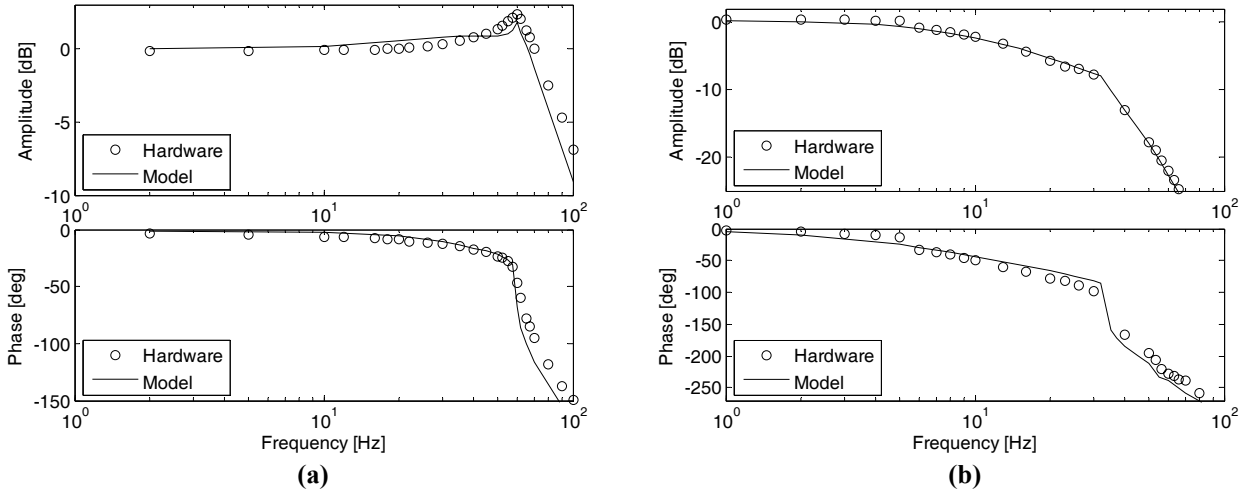


Fig. 8 – Model validation via frequency response (normal condition): (a) Servovalve; (b) Actuator .

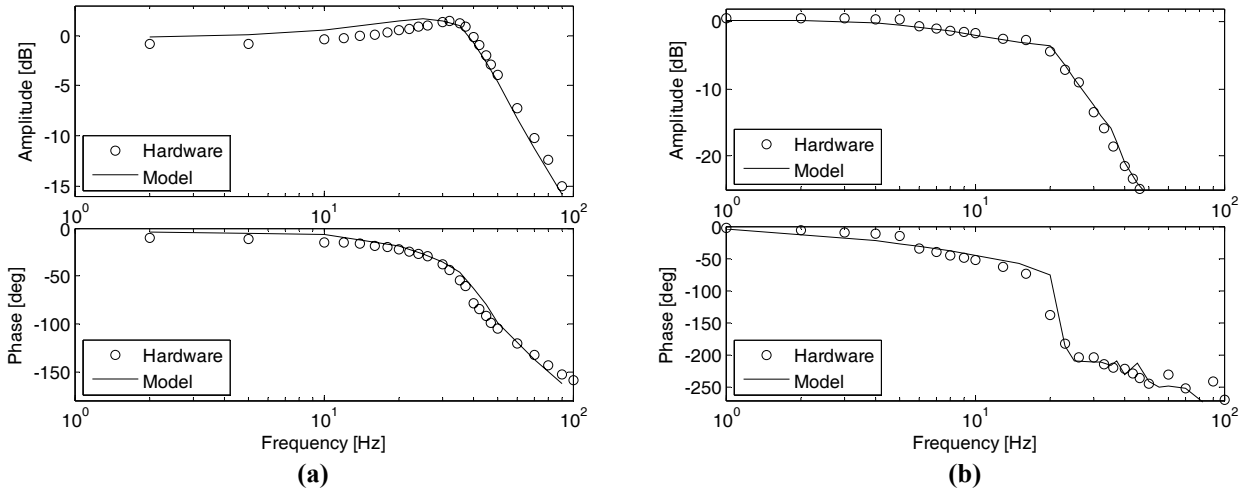
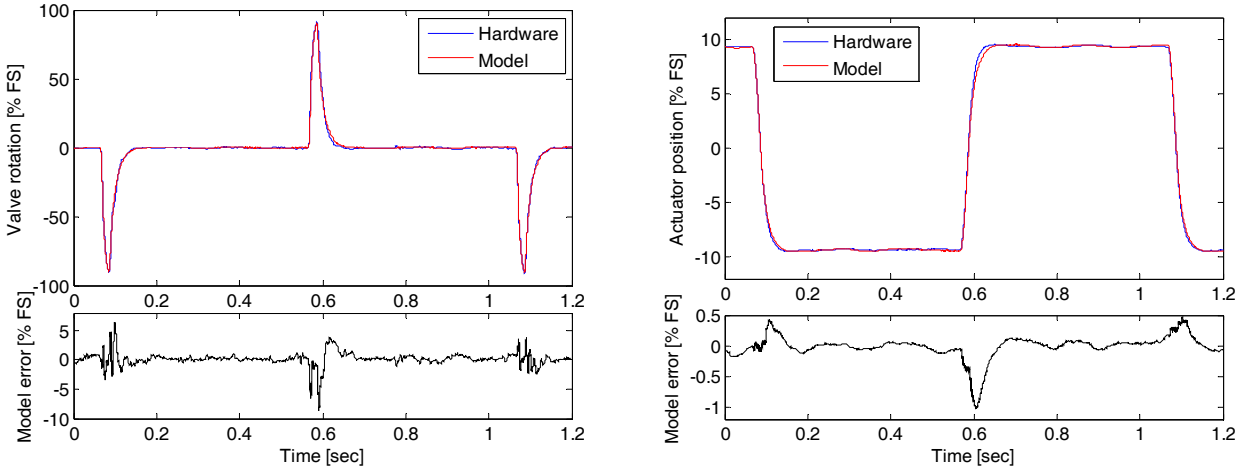


Fig. 9 – Model validation via frequency response (one active coil, no hydraulic failure): (a) Servovalve; (b) Actuator.

reference to the actuator in normal operating condition (with three coils active and both hydraulic systems pressurised).

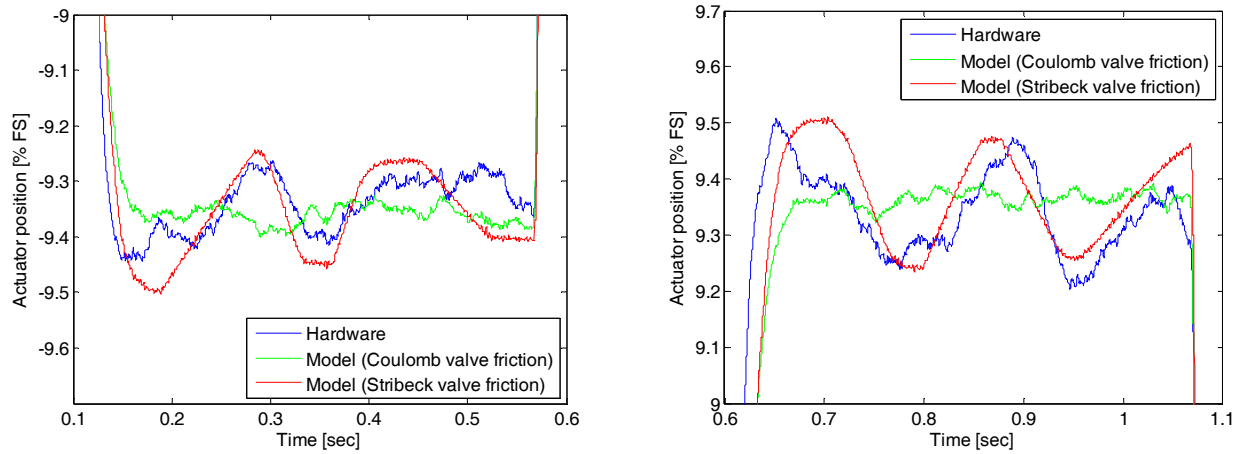
It can be noted that the model is capable of reproducing the hardware response at both servovalve and actuator level. In particular, the model succeeds in reproducing a current saturation phenomenon at 60 Hz, which implies an abrupt phase drop and an amplitude response approaching a -40 dB/decade slope, Fig. 8.a. The current saturation is also observable in the actuator frequency response (Fig. 8.b), but it occurs at lower frequency (40 Hz). This is because the servovalve response depends on the command amplitude, and the valve demand signal increases with frequency when the actuator response attenuates (the position error increases, Eq. (27)).

Figure 9 reports the validation results with the system with only one active coil (fail-operative condition). A good accuracy is once again achieved. In this case, the servovalve bandwidth is lower (Fig. 9.a), and this derives from the diminution of the motor current-to-torque constant, Eq. (12), which can be viewed as a decrease of the valve closed-loop gain. Concerning the actuator response, the current saturation is still present, but it occurs at 20 Hz. This is because, if only one coil is active, higher currents are needed for the servovalve motion, so that the current saturation is achieved at lower frequency.



(a) (b)

Fig. 10 – Model validation via step response (normal condition): (a) Servovalve; (b) Actuator.



(a) (b)

Fig. 11 – Limit cycle oscillations (normal condition): (a) Retraction command; (b) Extraction command.

4.2 Limit cycle oscillations

Since the frequency responses only provide "mean" information about the dynamic behaviour of a system (only the first harmonic signal component is considered), a complete model validation requires that also time responses are analysed, especially for systems exhibiting relevant nonlinearities [12].

Figure 10 shows the results of a step response test performed on the actuator with the system in normal condition. For both servovalve and actuator responses, the model error signal is also reported, so that the simulation accuracy can be quantitatively evaluated. The maximum amplitude of the model error at actuator level is less than 1% FS, which is a very good result in the perspective of using the model for monitoring. The model is less precise at servovalve level (the maximum error amplitude ranges 10% FS), but the accuracy loss is limited to the rapid transient dynamics, so that the model could be used also for the servovalve monitoring, if the monitor latency time is adequately regulated [13] [14] [15].

Another very interesting issue pointed out during the validation activity is that the actuator response is characterised by limit cycle oscillations. This self-sustained motion can generally be related to different system nonlinearities, and it can affect rotorcraft performance and stability [1] [16] as well as the

robustness of the FCC monitoring algorithms [13] [17]. Thanks to the high level of detail of the model, a specific sensitivity analysis has been performed, up to obtain that the model feature most notably impacting the phenomenon is the valve friction.

Figure 11 shows two details of the hardware and model step responses. The actuator limit cycle occurs at about 6 Hz frequency, and with an amplitude of about 0.15% FS. The plots point out that a too simple modelling of the valve friction (a Coulomb model) is not adequate, since no limit cycles would be predicted; while very good results are obtained if a Stribeck model is used.

5 Conclusion

The model of a hydraulic actuator with direct-drive servovalve for fly-by-wire helicopter controls has been developed and validated through experiments, demonstrating the importance of a detailed physical-mathematical modelling for satisfactorily reproducing the dynamics of a complex nonlinear system. The sensitivity of the actuator performance to the electrical failures as well as nonlinear behaviours due to the coil current saturation have been characterised via frequency response. The system time response has been also examined, highlighting that the actuator exhibits notable limit cycle oscillations. The high level of detail of the model allowed to perform a sensitivity analysis on the model features, highlighting that very good results can be achieved only by using a Stribeck valve friction model.

References

[1] Pratt R. W., *Flight control systems: practical issues in design and implementation*, Institution of Engineering and Technology, Stevenage, 2000.

[2] Task Group SCI-026. *Flight control design – Best practices*, RTO Technical Report 29, Canada Communication Group Inc., 2000.

[3] Di Rito G., Denti E., Galatolo R., *Development and experimental validation of real-time executable models of primary fly-by-wire actuators*. Proceedings of the Institution of Mechanical Engineers, Part I, Journal of Systems and Control Engineering, Vol. 222, No. I6, Sep. 2008, pp. 523-542.

[4] Taylor R., Pratt R. W., and Caldwell B. D., *Enhanced system models as a means of preventing aeroservoelasticity effects*, UKACC International Conference on Control, 1996, Vol. 2, pp. 764-769.

[5] Merritt H. E., *Hydraulic Control Systems*, John Wiley & Sons, New York, 1967.

[6] Di Rito, G., Denti, E., and Galatolo, R. *Convalidazione sperimentale del modello di simulazione di un attuatore servoidraulico fault-tolerant per comandi di volo fly-by-wire*. Proceedings of the 18th Aeronautics and Astronautics National Congress (AIDAA), Volterra (Italy), 2005.

[7] *Synchro/Resolver Conversion Handbook*. 4th Edition, Data Device Corporation, 1994.

[8] Khazan A. D., *Transducers and their elements*, Prentice Hall, 1994.

[9] Krishnan R., *Electric motor drives – Modelling, analysis, and control*, Prentice Hall, Upper Saddle River, New Jersey, 2001.

[10] Lyshevski S. E., *Direct-drive flight actuators: new challenging problems in nonlinear analysis and control of servo-systems*, Proceedings of the American Control Conference, Chicago, June 2000, Vol. 2, pp. 1098-1102.

[11] Olsson H., Åström K. J., Canudas de Wit C., Gäfvert M. and Lischinsky P., *Friction models and friction compensation*", European Journal of Control, 1998, Vol. 4, N. 4.

[12] Hanks G. W., Shomberg H. A., Kitto R. A., et al., *Integrated application of active controls (IAAC) technology to an advanced subsonic transport project – Test ACT system validation*. NASA CR-172525, NASA Langley Research Center, 1985.

[13] Bonnice W. F. and Baker W., *Intelligent fault diagnosis and failure management of flight control actuation systems*, NASA Contractor Report 177481, 1988.

[14] Roemer M. J., Kacprzynski G. J., Nwadiogbu E. O. and Bloor G., *Development of diagnostic and prognostic technologies for aerospace health management applications*. Proceedings of the IEEE Aerospace Conference, 2001, Vol. 6, pp. 3139-3147.

[15] Patton J. R., *Fault detection and diagnosis in aerospace systems using analytical redundancy*, Computing & Control Engineering Journal, May 1991.

[16] Wright J. R. and Cooper J. E.. *Introduction to aircraft aeroelasticity and loads*, John Wiley and Sons, New York, 2007.

[17] Webster, F. and Smith, T. D. *Flying qualities flight testing of digital flight control systems*, RTO-AG-300, Flight Test Techniques Series, vol. 21, 2000.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Ultrasonic Guided Waves for Structural Health Monitoring: Simulations and Experiments

E. Monaco, F. Ricci, S. Tancredi, D. Caporrino, L. Lecce
Aerospace Engineering Department, University of Naples "Federico II" 80125 Napoli – Italy

Keywords: *Structural health monitoring, ultrasonic guided waves*

Abstract

Much of Structural Health Monitoring (SHM) research is motivated by the fact that damage tolerant and fail-safe design of aircraft, aerospace and civil structures requires a substantial amount of inspection and defects-monitoring at regular intervals. There is a large inventory of aircraft structures in operation throughout Europe and the world that are undergoing continuous degradation through aging. Moreover, this number is increasing by around 5% every year, resulting in significant negative impact on the economy of many nations. The degradation of defects-critical structures is controlled through careful and expensive regularly scheduled inspections in an effort to reduce their risk of failure. With ever increasing cost of scheduled but often unneeded maintenance, intelligent, real-time monitoring of the condition of these structures is imperative to guarantee their safe and affordable continuing operation. In this paper the development of a reliable structural health monitoring system based on ultrasonic guided waves is considered. Theoretical results are compared with those obtained from an explicit finite element code for their mutual verification, showing excellent agreement.

1 Introduction

The ultrasonic research community has studied guided waves for nondestructive evaluation of plate-like structures for several decades. Guided waves are created by the constructive interference of the bulk waves reflected between the surfaces of the plate; these waves have a number of characteristics that are different from those of the bulk waves. First, they are, in general, multimodal and dispersive; the particle motion (symmetric or extensional and antisymmetric or flexural) and the velocity of each mode depends upon the thickness and material properties of the plate, as well as the frequency of the excitation of the wave. Second, they can propagate a much larger distance than the bulk waves without significant decay in their amplitude. Third, and most important, they are extremely sensitive to the presence of discontinuities in their path, and carry information on certain properties of the flaws as they propagate away from the flaws. Finally, it is relatively easy to generate and record the guided waves using (PZT) actuators and sensors that require very little power, and are therefore suitable for online structural health monitoring. Recently, research on active health monitoring of structural components using guided waves has gained considerable attention due to advances in sensor- and hardware technologies

and the increased usage of composite materials in load-carrying structural components, particularly in the aircraft industry. However, the optimum use of the favorable properties of guided waves in SHM requires a good understanding of their interaction with defects which can be gained through both theoretical simulations and carefully designed experiments. While significant advances in the experimental techniques have been made in recent years, advances in the theoretical studies on the interaction problem in composite structures have been rather limited.

The exact solution of three-dimensional problems consisting of undamaged multilayered, angle-ply laminates of finite thickness and large lateral dimensions subjected to various types of surface loads, has been available for some time [1, 2]. The techniques developed in these papers are difficult if not impossible to extend to realistic structures containing general types of defects. The finite element method (FEM) is a versatile tool to analyze this class of problems and a dynamic finite element code has been developed by NIST for the calculation of acoustic emission (AE) waveforms in isotropic and anisotropic plates [3]. Commercial finite element packages, e.g. LS DYNA, have been used extensively for large-scale simulation of impact problems in a variety of structures. Although the FEM can handle complex geometries and has the capability to accommodate reflections from the lateral boundaries, it is in general computationally much more intensive than analytical methods.

2 Numerical and experimental validation

The general experimental setup for the wave propagation test is shown in Fig. 1. The tests are conducted on a 6061 T6 aluminum panel (0.8 m x 0.7 m x 0.001 m). Four disc shaped piezo patches (ϕ=10 mm and thickness 0.2 mm) are bonded on one surface of the plate. A predefined ultrasonic signal (source) of central frequency at 200 kHz (Fig. 2) is generated by a HP/Agilent 33120A function generator, amplified and sent to one of the PZTs (source). The response is collected by the other PZT sensors (receivers) located at distances 50mm, 100mm and 200mm from the source, and downloaded in an HP DSO7014A 100 MHz 4-channel oscilloscope for the processing.



Fig 1 - The 0.8m x 0.8m x 0.001m 6061 T6 aluminum panel with the four disk shaped PZT's bonded on the surface. The function generator and the oscilloscope are also shown.

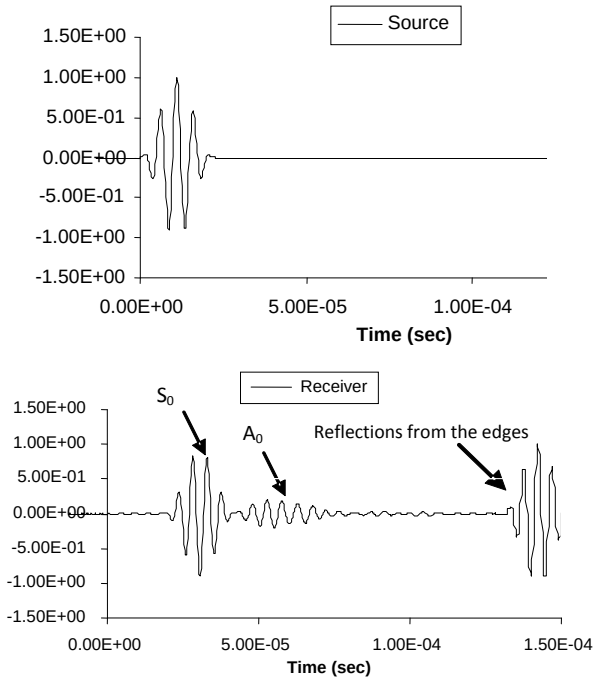


Fig 2 - The 4.5 cycle sinusoidal pulse in a Gaussian envelope used as a source (top) and the ultrasonic signal recorded at a distance of 100 mm from the source (bottom).

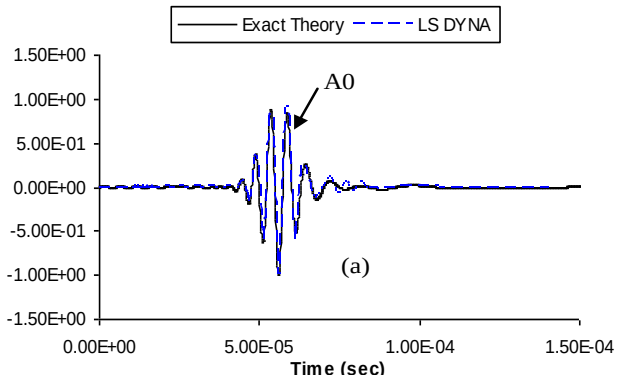
A typical recorded signal collected at a distance of 100 mm from the

source is shown in Fig. 2, which clearly shows the first S_0 and A_0 modes, and the signals reflected by the edge of the plate.

Numerical simulation using LS DYNA and theoretical modeling using exact theory are performed in order to approximately model the PZT excitation in the experiment. Since the behavior of the PZT is complex and involves coupled in-plane and out-of-plane excitation, each case has been modeled first.

The time histories of the vertical (out-of-plane) surface displacement, at a distance 100 mm, due to an out-of-plane and an in-plane point force, are collected and the results are presented in Fig. 3a and Fig. 3b, respectively. Results from the exact theory and the LS DYNA simulation are found to be in perfect agreement for all cases; the results are omitted for brevity in this paper. Notably, in Fig. 3b, a relatively small component of S_0 mode is observed due to the horizontal excitation, as expected.

By considering the experimental results too, a good agreement is found for the in-plane displacement at the receiver caused by an in-plane excitation, for the S_0 modes (Fig. 4), but not for the A_0 modes. Therefore, it can be concluded that the PZT patches used in the experiment, at the frequency of 200 kHz, actually excite and record mostly in-plane motion, which generates stronger S_0 waves than the A_0 modes. Thus, as a first step the A_0 modes were not modeled here.



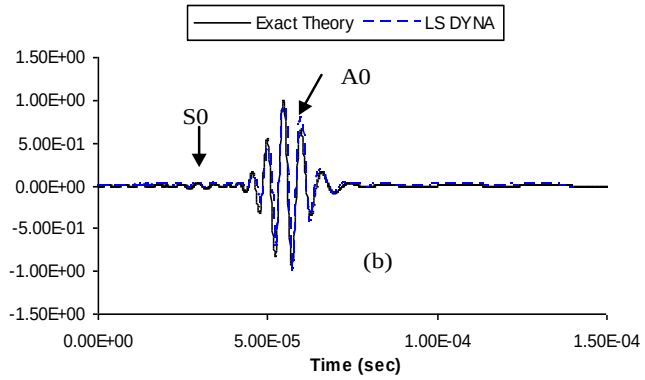


Fig 3 - Time history of out-of-plane surface displacement at a distance of 100 mm due to (a) a out-of-plane point force of Fig. 2 and (b) an in-plane point force of Fig. 2

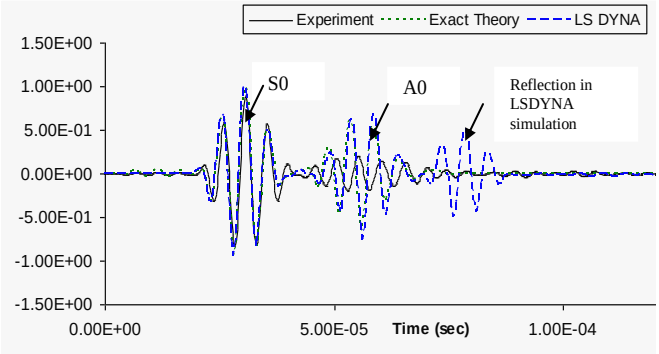


Fig 4 – The time histories of in-plane surface displacement at a distance of 100 mm from an in-plane point force on the surface

To simulate more accurately the action of the PZT patches in the FEM model, a particular load configuration is considered: 8-noded radial point load configuration (Fig. 5a). The piezo sensors are modeled considering an octagon with the vertices corresponding to the nodes of the FEM model (Fig. 5b): the time variation of its area was calculated and then related to the output voltage of the real sensors using the material constants of PZT.

Finally, a calibration is performed in order to find the relationship between the experimental input voltage given to the piezo source and the corresponding in-plane forces actuated by the piezo source itself. To carry out the calibration, finite element analyses are correlated to the experimental signals at the receivers through the piezo-material's electro-mechanical characteristics.

In Fig. 6 the numerical (FEA) and experimental signals received by a piezo patch

located 200mm from the source are reported. The exciting signal is a 4.5 sine cycles with Hanning window and central frequency of 40 kHz. The actual experimental output is 40V (peak excitation. Similarly, in figure 7 the receiver is 100 mm–away from the source and the excitation central frequency is 80 kHz.

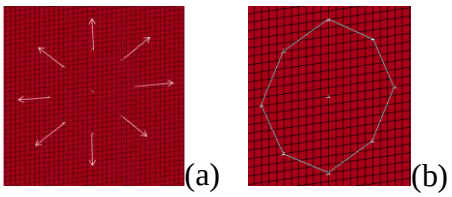


Figure 5 –Piezo modeling: (a) actuator; (b) sensor.

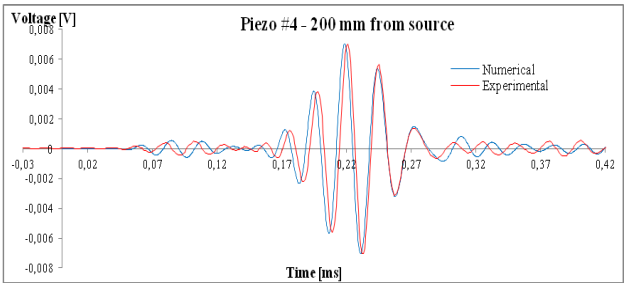


Fig. 6 –Numerical (FEA) and experimental received by a piezo patch located at 200mm from the source The exciting signal is a 4.5 sine cycles with Hanning window and central frequency of 40kHz. The actual experimental output is 40V (peak excitation.

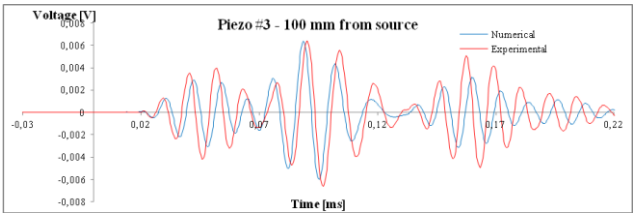


Fig 7 –Numerical (FEA) and experimental received by a piezo patch located at 100mm from the source The exciting signal is a 4.5 sine cycles with Hanning window and central frequency of 80kHz. The actual experimental output is 40V (peak excitation.

Although there are small phase shifts, the agreement of the curves in Figures 6 and 7 is very satisfactory for both S_0 and A_0 modes. No bonding layer effect is considered in the simulations.

3 Damage localization on composite panel

A fabric $[(0°)/(90°)]_5$ CFRP plate is equipped with an array of piezoelectric sensors. The positions are selected by an optimization routine based on the genetic algorithm technique applied to a FE model of the panel for a variety of damage locations and extensions. The FE models are developed taking account of the results discussed in the previous paragraph. The panel, whose dimensions in millimeters are 450x450x3.2, is impacted in a CEAST Fractovis drop weight impact machine at an impact energy of 84J. The actual damaged area is determined by a classical NDT system (Olympus BondMAster). For each pair of piezo patches (source and receiver) a wave propagation path is defined and a damage index calculated, using the signals collected before and after the appearance of the damage. The damage index technique can be found in [4] and it is not reported here for the sake of brevity. In Figure 8 a sketch of the panel with the sensor locations, the impact point and the actual damage extension are depicted. The damage position estimated with damage index technique is found to be very close to the actual damaged area.

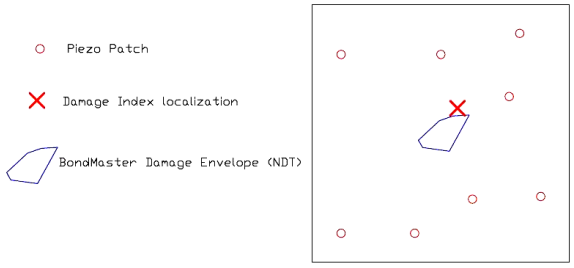


Fig 8 – The time histories of in-plane surface displacement at a distance of 100 mm from an in-plane point force on the surface

4 The beamforming technique

The beamforming is a widely used technique of signal processing to transmit or receive directional signals with arrays of sensors. In this work a linear array of sensors is used to generate an ultrasonic wavefront steered in a specific direction. This effect is achieved by the combination of disruptive and constructive interference of signals generated by the linear

array of sensors sequentially fired with a proper phase shift. The advantage of using a beamforming instead of an omnidirectional transmitter or receiver, is the spatial selectivity and the gain increase due to concentration of most of the energy in a narrower propagating lobe.

Numerical simulations are carried out on an aluminum panel 6061-T6 (1m × 1m × 1mm) with a rivet hole with cracked edges 20 mm wide. The array of sensors consists 9 disk shaped piezo patches (the same as described in the previous section) with a pitch of 12mm. Figure 9 shows the details of the sensors array. Three more piezo patches, acting as receivers, are bonded on the panel at 90°, 60° and 30° degrees with respect of center of the array, at a distance of 250 mm. Preliminary analyses are carried out with LS-DYNA that demonstrates the capability to simulate the beamforming behavior. In Figure 10 some results of steering tests and interaction of S_0 wave with a hole with edge cracks are reported. It is interesting to note that the FE analysis is able to reproduce the mode conversion during the scattering from a crack.

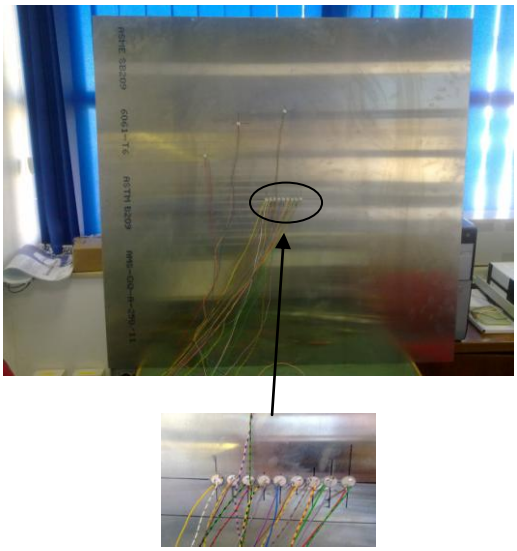


Fig 9 – The aluminum panel and a detail of the sensors array

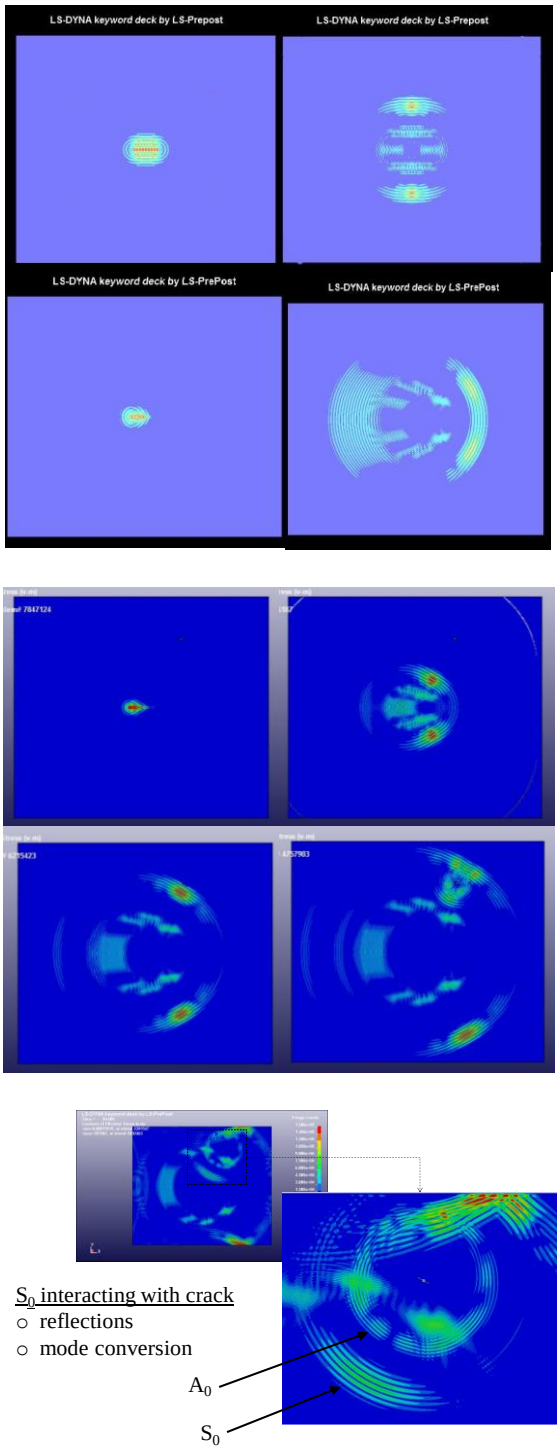


Fig 10 – FE simulations: 90° and 30° steering tests (upper), scattering of a 60° ultrasonic beam from an hole with edge cracks (center), a detail of the mode conversion of an incident S0 beam due to the scattering

Due to the 60° steering achieved with an appropriate phase shift of the array sensors firing, the piezos 6, 7 and 8 (see Fig. 11) receive most of the energy in both intact and cracked

configurations. Moreover, when the crack appears, part of the energy is scattered by the crack edges and the energy level received by those piezos is lower (Fig11).

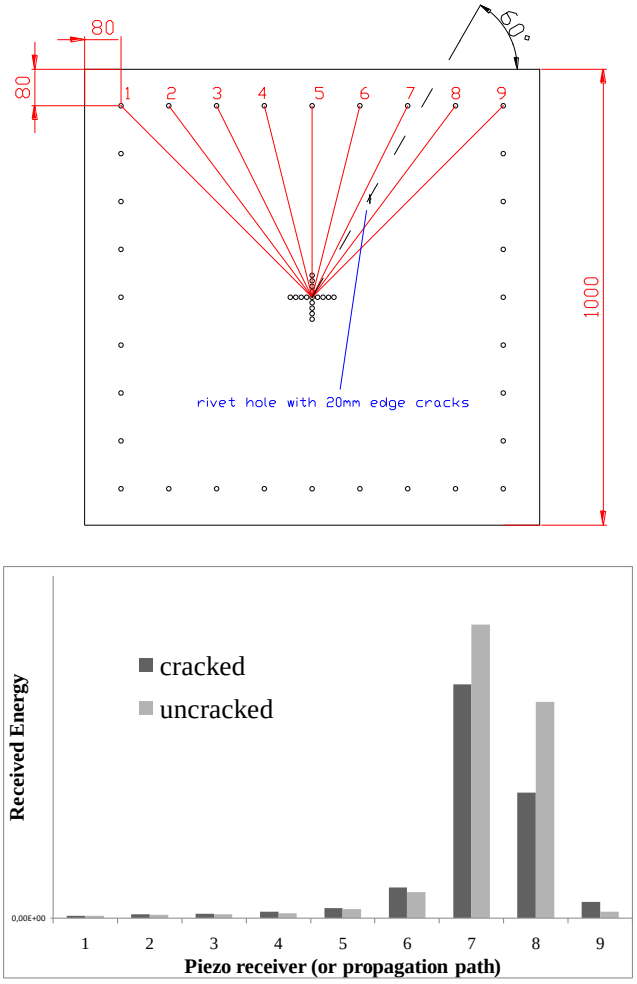


Fig 11 – FE simulations with beam angle at 60°. Most of the energy travels in the direction of the beam. The scattering produced by the appearance of a crack reduces the energy collected by the piezos

Preliminary tests are carried out on the intact aluminum panels, in order to evaluate the capabilities of the experimental setup. The multiple phased firing of the array sensors is performed using the round-robin technique [5], taking advantage of the superposition principle and avoiding experimental device complexities. In Figure 12 the variation of the gain of the phased array transmitter with the propagation direction is shown. The gain is defined as the ratio between the amount of energy propagated in these directions compared to the energy that would be propagated if the antenna were not directional. From Figure 12 the high level of directionality can be demonstrated. Each receiver sensor (R1, R2 and R3) measures a high signal only when the array of sensors transmits in the direction of the receiver. Moreover, when the beam direction is steered from 90° (normal to the sensors array) to 30° the energy collected by all the receivers decreases as the beamwidth increases.

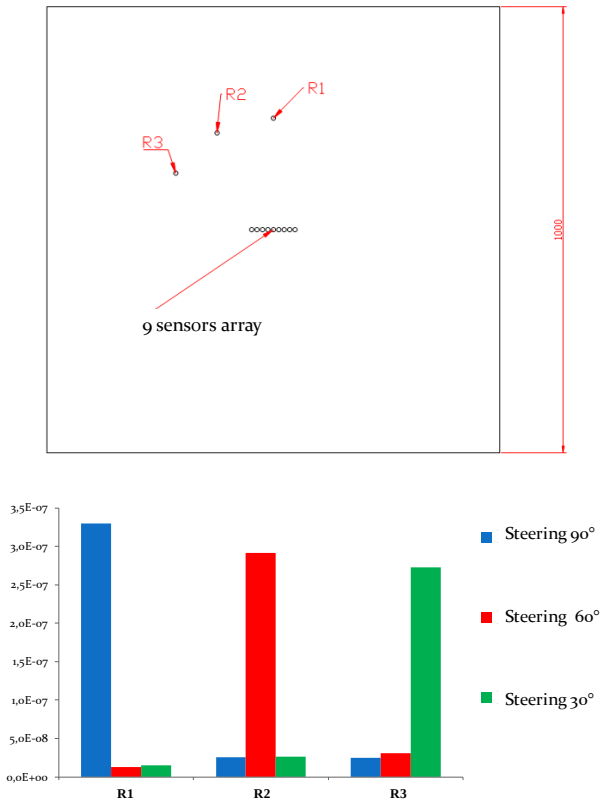


Fig 12 – Experimental evaluation of the beam directionality and gain dependence from the beam direction

5 Conclusions

Since the exact theory is computationally less expensive than the fnite element analysis, it can be effectively used to understand general characteristics of wave motion in simple, undamaged models of structural components. To our knowledge, theoretical solution of the class of wave interaction problems in structures containing hidden defects or other realistic features (e.g., stringers, joints, etc) is not available in the open literature, except for a few highly idealized cases. Explicit F.E.M approach

has the capability to handle models involving complex geometries containing defects and other structural anomalies. Therefore, the two methods substantiated by experimental tests can be combined for efficient and accurate calculation of the waveforms in more realistic structural components with or without damage.

References

[1] Lih, S. S., and Mal, A. K. (1995). ‘‘On the accuracy of approximate plate theories for wave field calculations in composite laminates’’. Wave Motion, 21, 17–34.

[2] Banerjee, S., Prosser, W. H. and Mal, A. K. (2005). “Calculation of the Response of a Composite Plate to Localized Dynamic Surface Loads Using a New Wavenumber Integral Method”. ASME J. of Appl. Mech., 72 (1), 18-24.

[3] Prosser, W. H., Hamstad, M. A., Gary, J., and O’Gallagher, A. (1999). ‘‘Comparison of Finite Element and Plate Theory Methods for Modeling Acoustic Emission Waveforms’’. J. Nondestruct. Eval., 18 (3), pp. 83–90.

[4] Banerjee, S., Ricci, F., Monaco, E., Mal, A. K. (2009). “A wave propagation and vibration-based approach for damage identification in structural components”, Journal of Sound and Vibration, 322 (1-2), 167-183.

[5] Yu, L. and Giurgiutiu V. (2007). “In situ 2-D piezoelectric wafer active sensors arrays for guided wave damage detection”. Ultrasonics,Volume 48, Issue 2, April 2008, Pages 117-134.

CEAS 2011

Safety and Reliability

Adaptation of a Fuzzy Logic Controller to Performance Based Contracting Analysis System

Andre Pozzetti and Cees Bil
RMIT, Australia

Graham Clark
RMIT, Australia

Keywords: *Performance, Contracting, Fuzzy Logic, Defense*

Abstract

In an effort to lower the overall cost of defense ware-fighter acquisitions, both the United States Government and the Australian Commonwealth Government mandate Performance Based Contracting (PBC) methods [1] upon acquisition and through life support contracts. Whilst being relatively new to Australian Defense [1-3], PBC is not a new concept to other industries around the world [4]. Demanding military aerospace acquisition and sustainment requirements however, present the defense industry with a significant challenge. This is inherent in the design of defense PBC and results from each military contract being different [4]. Considering the differing requirements, an awarded prime contractor is presented with a variety of complex management and technical issues, particularly if the prime contractor is also responsible for multiple platforms and corresponding contracts.

In this paper, we aim to adapt a fuzzy logic controller to a data warehouse system for the automation of condition based analysis of performance based contracts, which are specific to multi-platform military, rotary and fixed wing systems.

1 Introduction

PBC, as defined in the Aerospace System Division (ASD) guideline, is 'A product support strategy utilized by Program Managers to achieve measureable war-fighter selected performance outcomes for a Weapon System or Subsystem' [3]. Alternatively, the United States Department of Defense (U.S. DoD) prefers the terminology; Performance Based Logistics (PBL) [4]. U.S. DoD define PBL as 'The purchase of support as an integrated, affordable, performance package designed to optimize system readiness and meet performance goals for a Weapon System through long-term support arrangements with clear lines of authority and responsibility' [4]. The main aim of PBC/PBL as the purchase of outcomes, not products or services [1], [3-4]. The U.S. DoD and ASD define these outcomes as; Availability, Reliability, Maintainability and Supportability.

Outcomes are comprised of integrated Key Performance Indicators, and for Australian Defense contracts, are defined in a negotiation process between the Commonwealth of Australia (CoA) and the prime contractor, during both the tender phase and again through the award contract negotiation phase. This

process more often than not ensures [4] the individuality of each Military Aerospace PBC.

Due to the uniqueness of PBC/PBL contracts, one can assume that there is significant corporate overhead for managing, maintaining and supporting such contracts across multiple platforms. Additionally, simulating a proposed PBC/PBL during a contract negotiation phase would also prove difficult for the reasons above.

Encapsulating the ASD PBC framework within a Knowledge System, coupled with a fuzzy logic controller, could provide Defense contractors with the platform required to alleviate the complexities described above. Thus the goal of this paper is to utilize multi-disciplinary methodologies together in accordance with the ASD PBC Framework, to design and integrate a Knowledge System with a fuzzy logic controller for the automation of condition based analysis of PBC. Once integrated, the system could be used to compute strategy parameter values for genetic algorithm based optimization, which would present prime contractors the capability to undertake predictive analysis, modeling and simulation of various PBCs.

The first half of this paper will provide an overview of PBC, logical data models and system design as developed by Pozzetti et al (2010). The second half will focus on the fuzzy logic controller design, development and adaptation with the PBC Analysis System, with a priority on Aircraft Availability.

2 Performance Based Contracting

For all intents and purposes a PBC is an agreement between the purchaser and supplier and is common practice amongst a number of industries [5-9]. The contract typically specifies desired results, determined through features, services and/or integrated solutions [10], and is usually used to motivate the supplier to meet or even exceed the purchaser's requirements through the use of incentives and disincentives [3-4], [10].

The ASD regard the motivation as the purchase of performance or outcomes, however, the fundamental groundwork to any PBC is through three key inter-related characteristics: *Outcomes, Metrics* and *Payment rules* [3-4].

Through characterizing outcomes with a reliability and maintenance framework, ASD are remaining aligned with Richard and Jacopino [11], and the U.S. DoD [2], [4], [11]. Mandated outcomes for all future ASD Through Life Support contracts are in terms of [4]; System Readiness, Mission Success and Assurance of Supply. Additionally, Jacopino recommends that the continuous monitoring of *Available Aircraft, Mean Time Between Critical Failure* and *Demand Satisfaction Rate*, would effectively and efficiently assure these outcomes [11]. As a result, these three result areas have been adopted by Pozzetti et al (2010) as the design focus for a PBC Analysis System.

2.1 Performance Metrics

For traditional corporate and financial environments, the selection of indicators for the measurement of corporate goals is considered as straightforward, where as not so straightforward for complex military aerospace environments.

ASD PBC Framework emphasizes that performance metrics, not indicators, are lag metrics, and serve the purpose of representing the prime contractor's performance during the previous reporting period. An example demonstrated by Dwyer (2008); accounting information such as the revenue for the previous quarter is often lagged by approximately two months, whilst annual results can be delayed by a far greater margin [12]. Consequently, any corrective action made on the basis of a lag indicator would likely be redundant if the overall performance has changed during the lag period. There is an obvious and significant risk with formulating decisions based on lag indicators, which quite possibly will result in overcorrection, or a complete misunderstanding of the problem [12].

With the example in consideration, ASD PBC Framework incorporates the use of two tiers for performance measurement; Tier 1 being *Performance Metrics* and Tier 2 being *System Health Indicators*.

Tier 1 Performance Metrics, are a representation of mandated Outcomes, are expected to provide the CoA with a consistent measure of the subject Weapon System and its overall usefulness to the defense force. Tier 1 metrics are integrated with payment schemes as they are considered to be lag indicators, and therefore represent past performance [3]. The purpose of incorporating these is primarily to encourage the prime contractor to retain focus on the mandated Outcomes.

Tier 2 System Health Indicators are used to present both the prime contractor and the purchaser with a signal of the background support, process and/or activities behind each of the mandated outcomes [2], [3], [11]. Depending on the desired Outcomes, and as discussed earlier, the indicators are identified and negotiated during the sustainment contract negotiation process [3]. As the purchaser cannot specifically rely on Performance Metrics as the only measure of contractor performance [3], this second tier of performance measures, are considered to be vital. System Health Indicators provide the purchaser and prime contractor with Lead indication of future performance, as they support proactive management of any problematic performance area [3], [12].

2.2 Data Analysis and Design

Having already identified the three mandated Outcomes for the PBC analysis system [pozzetti], the following section illustrates the organization and flow of data in the system.

Available Aircraft can be described as a percentage of aircraft from a fleet available at a set time(s), which are available for use. The most important indicator [13], Available Aircraft KRA is defined as; 'the total number of aircraft available at a prescribed time, or through meeting a schedule daily Flying Program' [3]. It identifies the operable state of an aircraft at the commencement of a mission, and is further defined using the configuration of the aircraft and the operational status [3]. The Aircraft Configuration is defined by the various systems which are installed on the aircraft. The systems are specified through a Mission Critical Item List (MCIL), which details the systems, subsystems and components deemed necessary for a defined list of missions [3]. To provide characterization for the metric, ASD use the following condition codes [3]: Fully Mission Capable (FMC), Partially Mission Capable (PMC), and Not Mission capable (NMC). It must also be noted that the Aircraft Status is dependent on whether those systems installed on/to the aircraft are interpreted to be in an operational, degraded or non-operational state.

2.3 Logical Model

The logical data model, developed by Pozzetti et al, was derived through a combination of system context and data flow analysis [14]. This logical model emphasizes the data required for the three defined [3] outcomes [14], and prepare the requirements for the physical design.

The logical model, Fig. 1, was developed into a corresponding Data Mart, Fig. 2, using the Hybrid OnLine Analytical Processing (HOLAP) technique. This technique combines the best features of both multidimensional and relational database models [14-16]. The Available Aircraft Data Mart is one of the three objective based Data Marts in the PBC Analysis System [14], as shown in Fig. 3.

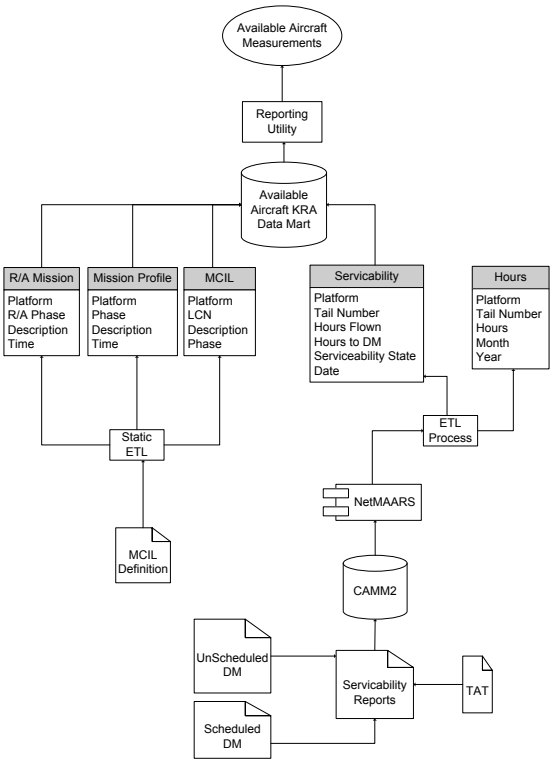


Fig. 1 Available Aircraft Logic Model

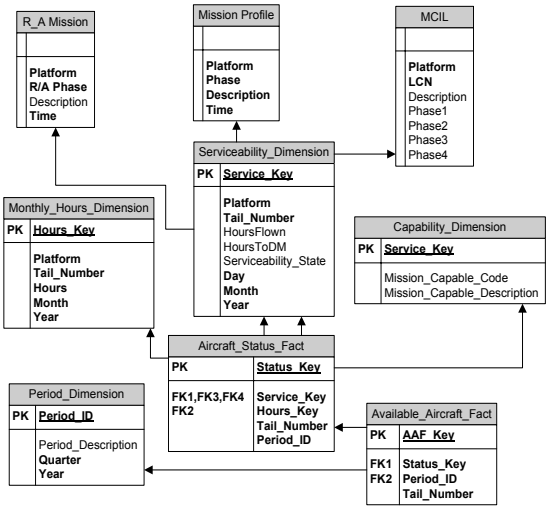


Fig. 2 Available Aircraft Star Schema

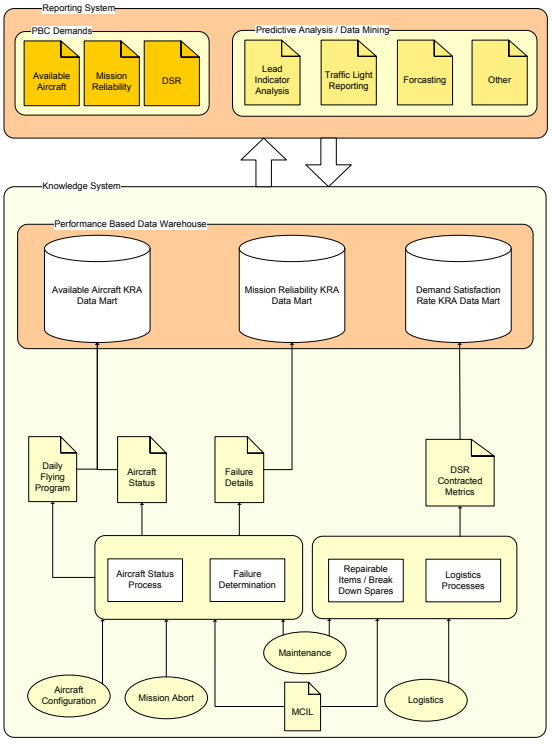


Fig. 3 PBC Analysis High Level System Design

Data Warehouses collect and store integrated sets of historical data from multiple systems, which then become the one prime source for all shared data [17]. A number of common tools are utilized in industry to extract data from data sources, and perform transformations for loading into the target Data Warehouse [17-18]. As proposed by Pozzetti et al [19], the tools may be logically designed in accordance to various contractual requirements to analyze performance.

3 Fuzzy Logic Controller Adaptation

Originally formulated by Zadeh [20], fuzzy logic is considered to be a powerful tool for performing reasoning on problems which may involve a level of uncertainty and/or vagueness [21]. Three main processes are involved with fuzzy logic; 1) fuzzification, 2) rule base reasoning, and 3) defuzzification.

Fuzzification describes the process of converting input values into fuzzy input membership values through defined membership functions. The fuzzy input membership values are then mapped to fuzzy

output membership values through the rule based reasoning process. Finally, the defuzzification process outputs a single, crisp value through defined formulas and fuzzy output membership values.

A fuzzy logic controller is described by Mitsuo as a kind of rule-based system which is built on fuzzy logic and fuzzy set theory [22]. Lee [21], describes a fuzzy logic controller as having four principal components:

1) A knowledge base
2) A fuzzification interface
3) An inference system
4) A defuzzification interface

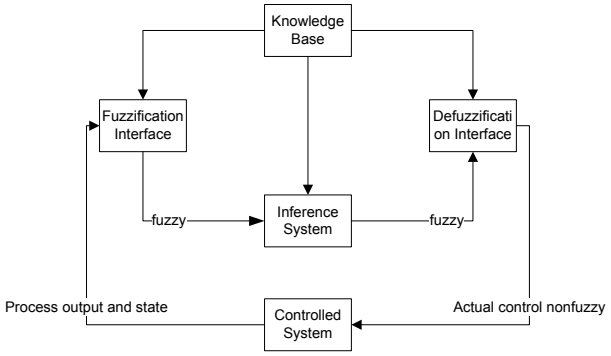


Fig. 4 Fuzzy Logic Controller (Lee [21])

The knowledge base is composed of two components: a database and a rule base. The database contains the definitions of linguistic variables used for the 'if-then' rules. The definitions are made up of membership functions for the fuzzy sets, whilst the inference system provides reasoning based on the knowledge base. The generic structure of a fuzzy logic controller is displayed in Fig. 4.

3.1 Fuzzy Rules

There are two known ways to build a fuzzy rule base; 1) through experience and/or knowledge or, 2) using automatic learning techniques. With consideration to the ASD Available Aircraft scoring process [3] in Fig. 5, the fuzzy rule base is defined by knowledge, or in this specific example, through ASD PBC Framework.

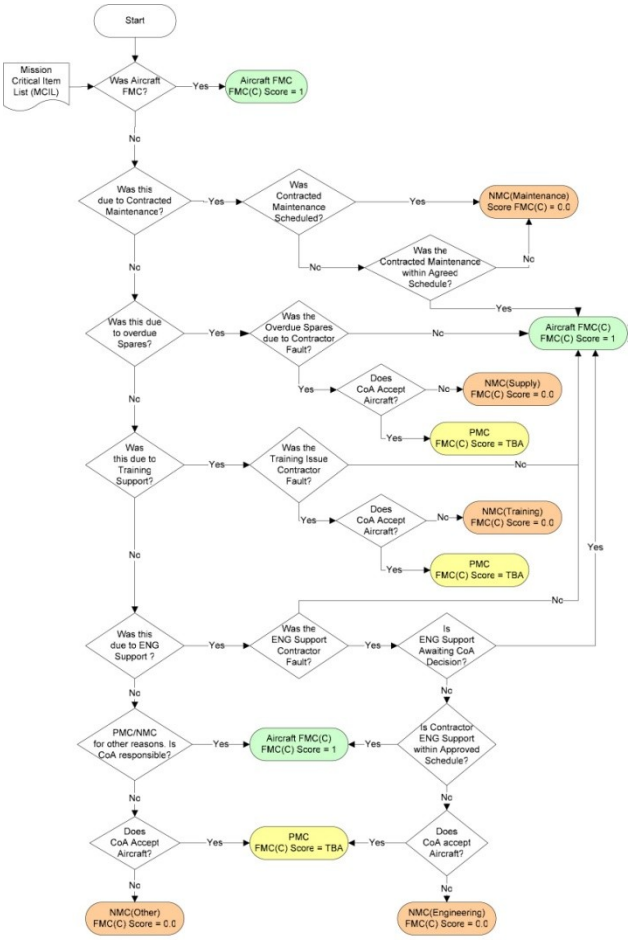


Fig. 5 ASD Scoring Process

The availability of aircraft, by definition is considered as a fuzzy output variable, as it can have three possible outcomes; FMC, PMC or NMC. Not all crisp values for the output variables are defined in Fig. 5, as ASD propose PMC scores to be defined within the program contract [3]. For the purpose of this exercise, we consider the PMC crisp value to be greater than NMC, but less than FMC.

Following Fig. 5, the fuzzy input variables are; Contracted Maintenance, Engineering Support, Overdue Spares, Training Support, Acceptance, Schedule and Commonwealth Responsible.

The linguistic terms for the input, output and membership functions (MF) are defined in Tables 1 and 2 respectively. Pseudo equations, as shown in Table 3, display the fuzzy input membership to the fuzzy output membership mappings in accordance with the process in Fig. 5.

Table 1 Output Linguistic Terms

Linguistic Term for AA Output	Meaning
NMC	Not Mission Capable
PMC	Partially Mission Capable
FMC	Fully Mission Capable

Table 2 Input Linguistic Terms

Linguistic Term for AA input	Meaning
CM	Contracted Maintenance
ES	Engineering Support
OS	Overdue Spare
A	Accepted
NA	Not Accepted
CoAR	Commonwealth Responsible
CF	Contractor Fault
NCF	Not Contractor Fault
WS	Within Schedule
NWS	Not Within Schedule
U	Unscheduled
S	Scheduled

Table 3 Pseudo Fuzzy Rules

$$if\ (CM == U)\&\ (Schedule == WS \rightarrow AA = FMC = 1 \quad (1)$$

$$if\ (CM == U)\&\ (Schedule == NWS) \rightarrow (AA = NMC) = 0 \quad (2)$$

$$if\ (CM == S) \rightarrow (AA = NMC) = 0 \quad (3)$$

$$if\ (ES == NCF) \rightarrow (AA = FMC) = 1 \quad (4)$$

$$if\ (ES == CF)\ \&\ (Schedule = WS \rightarrow AA = FMC = 1 \quad (5)$$

$$if\ (ES == CF)\&\ (Schedule == NWS\&\ Accepted == A \rightarrow AA = PMC \leq 1 > 0 \quad (6)$$

$$if\ (ES == CF)\&(Schedule == NWS\&Accepted == NA \rightarrow AA = NMC = 0 \quad (7)$$

$$if\ (OS == NCF) \rightarrow (AA = FMC) = 1 \quad (8)$$

$$if\ (OS == CF) \rightarrow (AA = NMC) = 0 \quad (9)$$

$$if\ (OS == CF)\&(Accepted = A) \rightarrow (AA = PMC) \leq 1 > 0 \quad (10)$$

$$if\ (TS == CF)\&(Accepted == A) \rightarrow (AA = PMC) \leq 1 > 0 \quad (11)$$

$$if\ (TS == CF)\&(Accepted == NA) \rightarrow (AA = NMC) = 0 \quad (12)$$

$$if\ (TS == NCF) \rightarrow (AA = FMC) = 1 \quad (13)$$

$$if\ (CoAR == True) \rightarrow (AA = FMC) = 1 \quad (14)$$

$$if\ (CoAR == False)\&(Accepted == A \rightarrow AA = PMC \leq 1 > 0 \quad (15)$$

$$if\ (CoAR == False)\&(Accepted == NA) \rightarrow (AA = NMC) = 0 \quad (16)$$

3.2 Membership Functions

Membership Functions (MF) were developed using a MATLAB Fuzzy Inference System (FIS). Gaussian type MFs were applied for each of the input and output variables, which is based on a symmetric Gaussian curve, as defined in Eq. 17. The example of the CM MF details the rule in accordance with the process in Fig. 6, meaning that CM can either be S or US, and is shown in Fig. 6 respectively.

$$f(x) = exp\left(\frac{-0.5(x-c)^2}{\sigma^2}\right) \quad (17)$$

Where c is the mean and σ is the variance.

The pseudo fuzzy rules, developed in Table 3, were also translated into a fuzzy rule base using MATLAB; displayed in the rule definitions Table 4. The rule definitions combined with the corresponding MFs result in the surface area plots shown in Fig. 8 through 13 respectively.

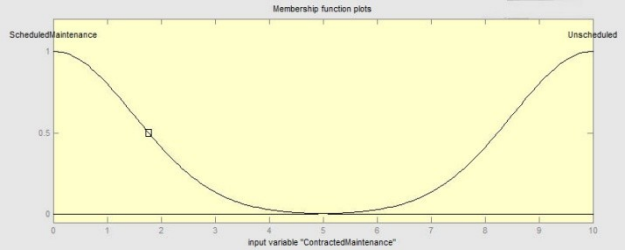


Fig. 6 CM MF

3.3 Adaptation

As discussed in Section 2, the fuzzy logic controller can be adapted to the Extraction, Transformation and Loading (ETL) layer of the PBC Data Warehouse, and in this specific example, the Aircraft Availability Data Mart. information is passed through the inference system, which then produces crisp data for the Availability fact tables. A high level system diagram, representing the adaptation of the fuzzy logic controller is shown in Fig. 7. Basic simulation can be performed using the MATLAB rule viewer, however it must be noted that the limit of this simulation is to confirm anticipated defuzzification results, rather than the simulation of the proposed system as described in Fig. 7.

4 Conclusion

The purpose of this paper was to assess the potentiality for the incorporation of the ASD PBC framework into an analysis system through the use of fuzzy logic. Through the use of MATLABs Fuzzy Inference System toolbox, the translation from contract rules through to MFs are demonstrated with the surface area plots in Fig. 9 through to 13 respectively.

It is worth noting that we also attempted to use multiple MFs for a type of input variable. For example, we assigned each, CM, ES, TS and OS as MFs for an input variable 'DecisionType'. Whilst the defuzzification with this system gave the same results, the surface area plots are visually difficult to interpret, and therefore were excluded from this paper.

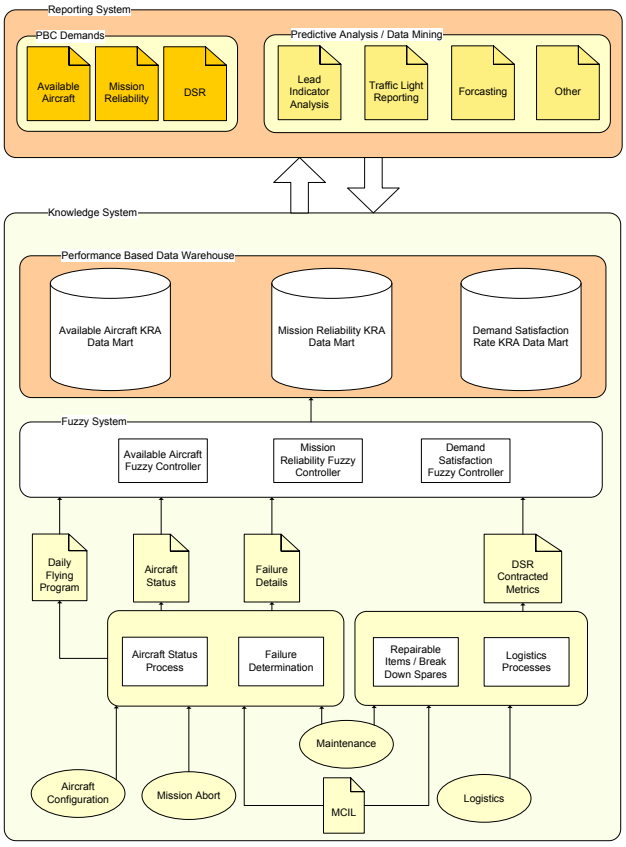


Fig. 7 High Level System Diagram

As a next step, we propose and aim for future experimentation and integrated simulations using each method of MF incorporation. The simulation should involve daily sample data sets over a set time period, incorporating all of the key result areas. The function developed should then be both minimized and optimized according to desired performance goals in order to obtain; 1) requirements for optimal contract performance, and 2) the limit to determine when liquidated damages would apply.

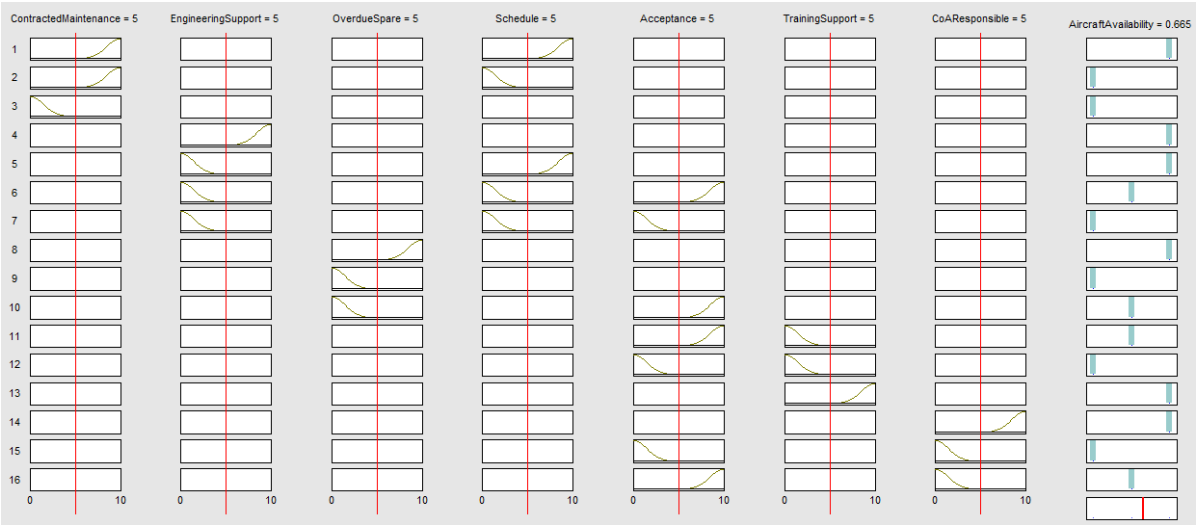


Fig. 8 MATLAB Rule Definitions

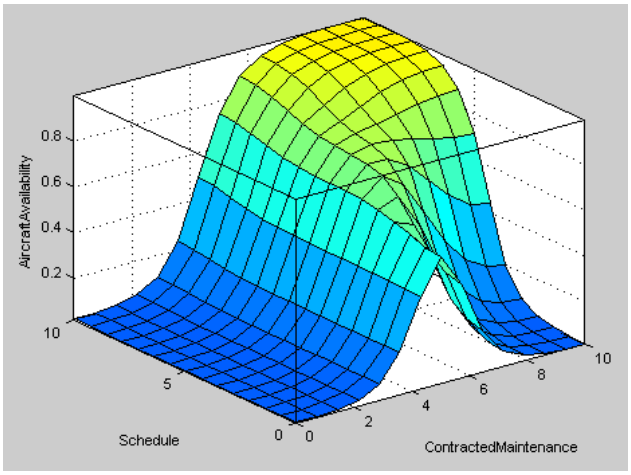


Fig. 9 MF (CM & S)

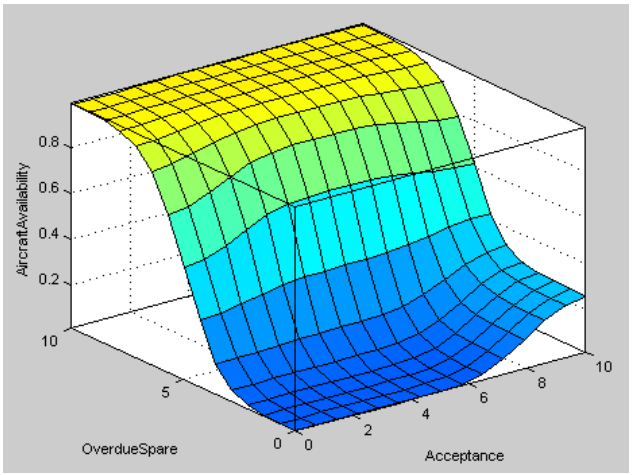


Fig. 11 MF (OS & A)

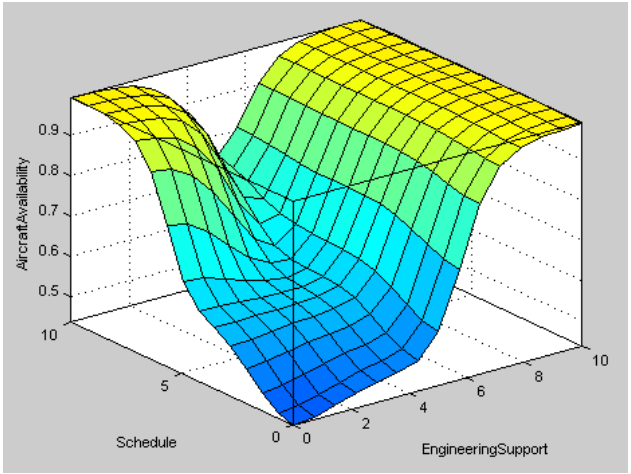


Fig. 10 MF (ES & S)

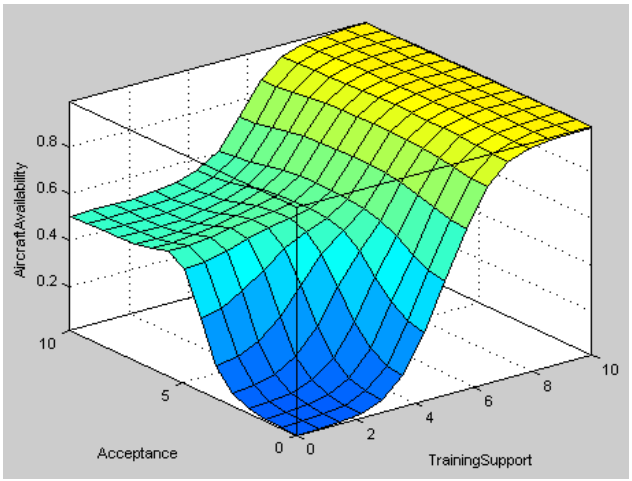


Fig. 12 MF (TS & A)

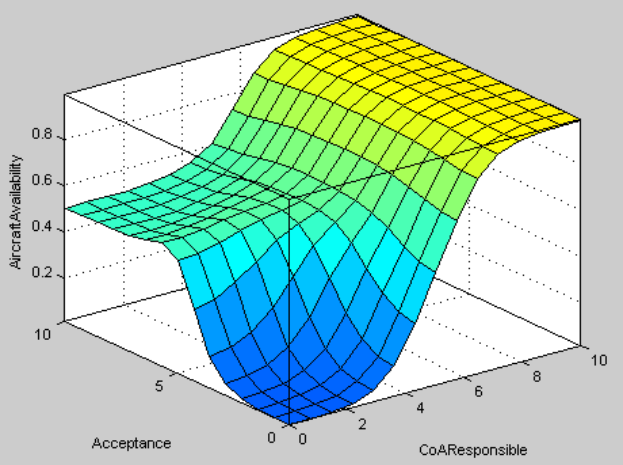


Fig. 13 MF (CoAR & A)

References

[1] DEFENCE MATERIEL ORGANISATION (2003), *The Australian Defence Aerospace Sector Strategic Plan*, in DEFENCE MATERIEL ORGANISATION (ed.). Canberra, Commonwealth of Australia.

[2] SQNLDR DAVID ABRAHAM (2005) Performance based contracting - a panacea for supporting legacy aerspace platforms? IN AUSTRALIAN COMMAND AND STAFF COLLEGE (Ed.). Geddes Papers.

[3] DEFENCE MATERIEL ORGANISATION (2007) Performance Based Contracting Handbook; Guiding Principles and Performance Framework. IN AEROSPACE SYSTEMS DIVISION (Ed.). Canberra, Commonwealth of Australia.

[4] OF DEFENSE (2005) Performance Based Logistics: A Program Manager's Product Support Guide. IN DEPARTMENT OF DEFENSE (Ed.). Defense Acquisition University Press.

[5] MOLENAAR, K., YAKOWENKO, G. & CONSTRUCTION INSTITUTE (2007) *Alternative Project Delivery, Procurement, and Contracting Methods for Highways,* Reston, VA, American Society of Civil Engineers (ASCE) Publications.

[6] TURNBULL, K. F. (2007) U.S. and international approaches to performance measurement for transportation systems. *Transportation Research Board.* Beckman Conference Centre, Irvine, California, Transportation Research Board, 2008.

[7] HOUT, W. (2007) EU development policy and poverty reduction: enhancing effectiveness, Hampshire, England, Ashgate.

[8] ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD) (2009) Managing water for all an OECD perspective on pricing and financing. OECD.

[9] NATIONAL RESEARCH COUNCIL (U.S.) (2004) Intelligent Sustainment and Renewal of Department of Energy Facilities and Infrastructure, Washington, D.C., National Academies Press.

[10] GARRETT, G. A. (2007) *Performance-based contracting,* Vienna, VA, Management Concepts.

[11] JACOPINO, A. (2007) Modelling R&M in Performance Based Contracts - When Does Risk Equal Reward? *Reliability and Maintainability Symposium.* Orlando, FL.

[12] DWYER, K. (2008) What is a KPI. Change Factory. (http://www.changefactory.com.au/articles/business-management/what-is-a-kpi)

[13] GAO (2003) Military Readiness - DOD Needs a Clear and Defined Process for Setting Aircraft Availability Goals in the New Security Environment. DIANE Publishing.

[14] POZZETTI, A., BIL, C. & CLARK, G. (2010) Database Management System for Performance Based System Assessment of Military Multi-Mission Platforms. *SimTect 2010 Simulation Conference and Exhibition (SimTecT 2010).* Brisbane, Australia.

[15] ROB, P. & CORONEL, C. (2007) *Database Systems: Design, Implementation, and Management*, Cengage Learning.

[16] MICROSOFT (2008) KPI How-to Topics (Analysis Services). *SQL Server 2005 Books Online.*http://msdn.microsoft.com/en-us/library/ms174864%28SQL.90%29.aspx

[17] KEYES, J. (2006) Knowledge Management, Business Intelligence, and Content Management - The IT Practitioner's Guide, CRC PR INC.

[18] ADAMSON, C. (2006) *Mastering Data Warehouse Aggregates: Solutions for Star Schema*, Wiley Pub.

[19] POZZETTI, A., BIL, C. & CLARK, G. (2010) Implementation of Performance Based System Assessment of Military Multi-Mission Platforms. *27th Congress of the International Council of the Aeronautical Sciences.* Nice, France.

[20] ZADEH, L. A. (1988) Fuzzy Logic. *IEEE Computer*, 83-89.

[21] LAU, H. C. W., CHAN, T. M. & TSUI, W. T. (2008) Item-Location Assignment Using Fuzzy Logic Guided Genetic Algorithms. *IEE Transactions on Evolutionary Computation,* 12, 765-779.

[22] MITSUO GEN & CHENG, R. (2000) *Genetic Algorithms and Engineering Optimization,* New York, John Wiley and Sons.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

On an Engine Health Management System

K. Efthymiou, P. Georgakakis, N Papakostas, G. Chryssolouris
Laboratory for Manufacturing Systems and Automation
Department of Mechanical Engineering and Aeronautics
University of Patras
26500, Patras, Greece

Keywords: *predictive maintenance, remaining useful life, autoregressive integrated moving average model, symbolic dynamics filtering*

Abstract

Airline operators strive to reduce the overall time and cost of an aircraft's engineering maintenance activities, whilst ensuring at the same time that its safety, reliability and integrity are not be compromised. In order for this to be carried out, some predictive maintenance concepts can be considered, so as for the maintenance costs to be decreased, while the aircraft's reliability is maintained up to the highest levels. This paper proposes a method capable of assessing the engine's deterioration as well as its remaining useful life, before the next major overhaul. The proposed approach employs an ARIMA model for the estimation of the remaining useful life, based on historical flight data, analyzed by symbolic dynamics techniques. It can be argued that this approach may, under certain circumstances, determine the remaining useful life of an aircraft's engine, providing the maintenance engineers with a significant decision support tool.

1 Introduction

Engine related costs contribute a large fraction of the direct operating costs of an aircraft. According to the International Civil Aviation Organization, the maintenance costs range from 10% to 20% of an operator's direct costs, depending on the fleet size, age and usage [1]. The occurrence of unscheduled maintenance can introduce costly delays and cancelations if the problem cannot be rectified in a timely manner. Planned maintenance defines a periodic interval for the performance of preventive maintenance, regardless of the engine's condition. This maintenance approach ameliorates the engine's reliability however, it increases significantly the maintenance costs. Therefore, a more efficient maintenance approach in terms of cost and time that assures the reliability of the engine is required. Condition-based maintenance (CBM) is a maintenance program that recommends maintenance actions based on the engine's status. The CBM utilizes prognostics methods and is considered more efficient without losing its reliability in comparison with the planned maintenance in terms of cost.

One of the two main prediction types of prognostics is the remaining useful life (RUL). The RUL refers to the time currently left before a certain failure criterion is reached. The accurate assessment of the RUL of an engine, while in use, is of critical importance, since it has impacts on planning the maintenance activities,

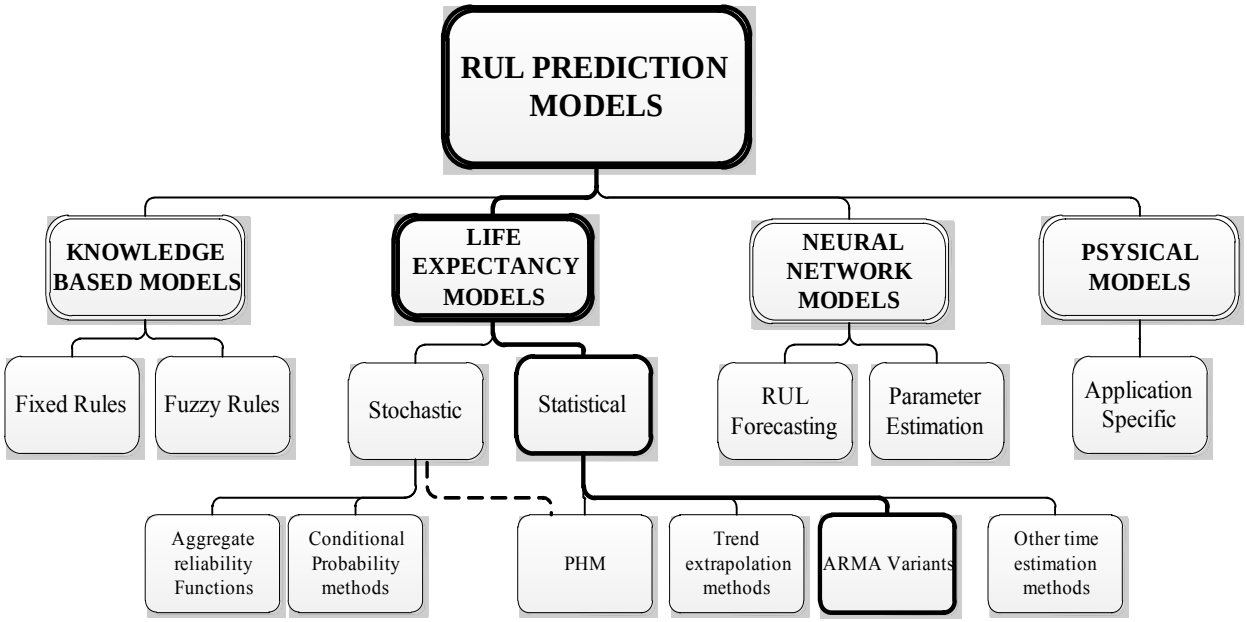


Fig.1 Classification of RUL prediction models

the spare parts' provision, the operational performance and the profitability of the operator. Thus, the RUL estimation of the aircraft's engine may reduce the occurrence of unscheduled maintenance and improve the aircraft's operability without increasing cost, whilst enforcing at the same time, the safety as well as the reliability of the aircraft. This paper proposes a methodology for RUL estimation, based on symbolic dynamic analysis and stochastic regression methods enhanced with an artificial neural networks model, responsible for modeling the engine in its sound condition.

The rest of the paper is organized as follows. Chapter 2 presents a short review of the existing literature on RUL estimation approaches. Chapter 3 provides an overview of the proposed method, describing in brief the three main pillars of the approach, in particular, the aircraft engine modeling with Artificial Neural Networks (ANN), the anomaly measurement with Symbolic Dynamic Filtering (SDF) and the anomaly prediction with ARIMA models. The next three chapters, 4, 5 and 6 describe in detail, the three pillars of the introduced method. The final chapter concludes the basic outputs of this work.

2 Literature Review

During the past years, in the context of the prognostics and CBM, various RUL approaches have been proposed that differ in terms of estimation accuracy, range of application, cost and implementation complexity. These approaches are classified by [18] into four main categories: a. knowledge based, b. physical models, c. artificial neural networks and d. life expectancy. Knowledge based models store past defined failures in a database. In case of a new event, they retrieve the most similar past observation, related to a failure, and they deduce the life expectancy of the asset. The knowledge based models are further classified to expert and fuzzy systems. Artificial neural networks estimate the remaining useful life of the engine directly or indirectly, by training an ANN with past observation data of failure events. Physical models provide an assessment of the RUL, based on a mathematical representation of the physical behaviour of the degradation process. Finally, life expectancy models determine the time left of an asset or an asset's component with respect to the expected risk of deterioration, under known operating conditions. This type of models consists of two main groups, in particular, the stochastic and

statistical models that also include sub-categories. Figure 1, presents a classification scheme of the main model categories adopted by [18].

The physical models are characterized by high accuracy, high implementation cost and complexity. In addition, physics based models are application specific and depend on the type of the asset, in this case, the type of the engine. Therefore, their applicability range is too limited. On the contrary, knowledge based models are more efficient as regards the applicability range, implementation cost and complexity. They can be applied to a broad area of assets, while their cost and complexity are lower in comparison with physical models. However, the accuracy of the RUL estimation is limited, since it is based on similar past cases that may not be so relevant to the observed situation of the existing problem. Artificial Neural Networks and life expectancy models are of medium complexity, with a quite broad of application range, without a high implementation cost [[8], [18]].

The classification of the research work on the RUL estimation is problematic since a lot of models follow hybrid approaches, combining two or more of the aforementioned models. The idea of this hybridization is to combine the strengths of two or more approaches for a better estimation of the RUL. The current paper also follows this hybridization approach, integrating ANN with ARIMA models, while for signal noise reduction the Symbolic Dynamic Filtering (SDF) is utilized. In this way, the health management system for RUL prediction is proposed and it is characterized by medium implementation cost and complexity, a quite broad range of application without losing in accuracy.

The statistical life expectancy models are used in [17], where the performance deterioration of a turbofan engine is studied with the use of forecasting algorithms. In particular, the ARIMA models are developed for the estimation of the RUL of a turbofan engine, for short term predictions. For medium and long term prediction it is the regression analysis used. Specifically, three deterioration models are proposed, for severe, linear and soft deterioration, utilizing the regression analysis. In a similar work [10], a diagnostic and prognostic system for gas turbines is proposed. Firstly, the gas path approach is followed for the diagnosis. Secondly, the prognosis is based on the linear and quadratic regression models that facilitate the estimation of the engine's RUL. ANN models are used in [6], [7] and [9]] in an integrated prognosis process. In a similar approach related to engine health management, the ANN models are used for performing anomaly detection [6]. The proposed system (eSTORM) combines the physical modelling of an engine with an empirical neural net model to provide a hybrid model for reliable engine diagnostics and prognostics. Neural Net models are responsible for enhancing the physical model accuracy and improving the reliability of the diagnosis [7]. The concept of the symbolic dynamic analysis of complex systems for anomaly detection is introduced in [3]. "Anomaly" in a dynamical system, is considered as the deviation from its nominal behaviour. The current study emphasizes on the theoretical background that allows for an early detection of anomalies in a complex system. In [4], a fault detection and isolation methodology in an aircraft's gas turbine engine is proposed. The Fault Detection and Isolation (FDI) methodology is based on the analysis of time series data of available sensors and analytically derived variables in the gas path dynamics. In particular, a novel methodology for FDI, based on symbolic dynamics, statistical pattern recognition and information theory is introduced. The proposed concept seems to be capable of detecting any slowly evolving anomalies apart from the abrupt changes in the nominal behaviour. The validation of the concept is provided in [5], where the SDF based methodology is applied to a simulation test bed of a two-spool turbofan aircraft gas turbine engine.

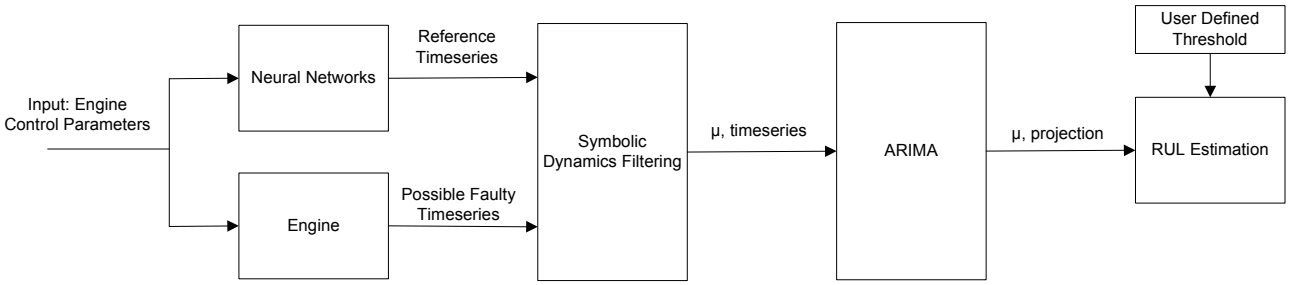


Fig.2 Health Management System

3 Method Overview

The approach presented in this paper is based on three pillars. The first pillar concerns the modelling of the engine in its nominal condition with the use of ANNs. The second pillar is responsible for the engine's anomaly measurement, comparing ANNs nominal model with the engine's behaviour. Within the second pillar, anomaly measurement is performed with the use of the SDF analysis. The third pillar performs the estimation of the engine's RUL with the use of an ARIMA model.

Degradation is considered as the engine's deterioration in performance while it ages. In the present approach, the gas path concepts are followed for the diagnosis of the deterioration employing ANNs. The gas path analysis aims at detecting the gas path faults that have observable effects on the measurable parameters, such as temperature and pressure. The deviation between the measured parameters and the ANN parameters is analyzed by the SDF. The output of this analysis is μ, an anomaly measurement. The behaviour of μ in time is indicative of the engine's deterioration and is captured by the ARIMA models.

ANNs approach for the engine's modelling is selected due to four main reasons. Firstly, ANNs are capable for modelling the complex nonlinear behaviour, such as the behaviour of an aircraft's engine. Secondly, the implementation complexity and cost is not so high in comparison with that of the physical models. Thirdly, ANNs provide an accepted level of accuracy, not as high as the physical models do, but higher than the Knowledge based models. Finally, a significant advantage of ANNs is their generic nature. ANN models are not restricted to a specific type or model of engines, but after a proper training they can be applied to a quite large variety of engines.

Anomaly measurement that is based on the analysis of deviation of an engine's parameters from their nominal conditions depends significantly on the quality of the engine's measurements. The quality of the measurements is affected by the sensor noise, the disturbances, the instrument degradation and the human errors [10]. Concerning the quality of measurement, the SDF presents two major advantages in comparison to similar methods based on ANN and the Principal Component Analysis [4]. Firstly, the SDF is more robust to measurement noise and spurious signal [[3], [4], [5] and [19]. Secondly, the SDF is more adaptable to low resolution sensing, due to the coarse graining in space partitions [3]. The SDF can also be a part of a real time Fault Detection and Isolation process, which provides the opportunity of the same data and methods on deferent processes to be utilized.

The engine's deterioration is captured with the use of ARIMA models because the ARIMA models present three main advantages. Firstly, the ARIMA do not require historical failure data, and in general, a detailed description of the failure mechanism is not necessary [18]. Secondly, the ARIMA models in contrast to the ARMA variants are capable of capturing non stationary deterioration processes. Finally, ARIMA provides the capability of incorporating random shocks, assessing more accurately the stochastic behaviour of deterioration in contrast to the regression models [16].

The proposed health management method is capable of addressing the deterioration of a broad range of engine types since the modeling is based on ANN, is robust to sensors noise due to SDF advantages, and is capable of capturing

non stationary degradation behavior for RUL prediction because it employs ARIMA models.

4 Engine Modeling

In general, Neural Networks (NNs) are parallel distributed information processing systems capable of learning and generalizing from training patterns or data. Their ability to model complex nonlinear systems makes them adequate for the modeling of an aircraft's engine. The level of user knowledge needed to successfully apply NNs is much lower than that in the case of using a more traditional nonlinear statistical method, although some heuristic knowledge of how to select and prepare data, how to select an appropriate NN and how to interpret the result is still required.

4.1 Models of the processing elements

Each neuron is a simple processor, which takes one or more inputs, usually exits from other neurons and produces an output. The processes that a neuron performs are: a multiplication of the inputs by their respective weight, adding up the resulting numbers for all the inputs and the determination of the output, according to this summation and a transfer function. Specifically, the neuron can be viewed as consisting of two parts: a linear and a non-linear part. The first is an integrated function f that combines information, coming from other neurons and the net in general. In the current approach, the function that has been utilized is described in equation 1, where x_j is the output of neuron j of the preceding neurons, w_j its respective weight and θ is the threshold of this neuron. The output of the neuron is a non-linear transfer function $a(f)$, specifically the tansig function (hyperbolic tangent function).

The second part of the neuron is a non-linear transfer function $a(f)$, in this case, the tansig(hyperbolic tangent) function.

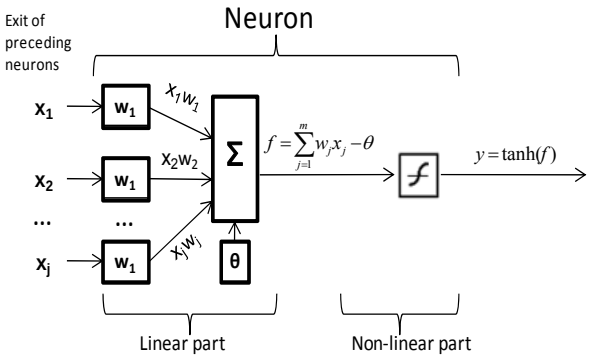


Fig. 3 Neuron Diagram

4.2 Network Structure

Of course, a single neuron is of very limited use, while a number of connected neurons, i.e. a network, can be trained to model non-linear complex behaviors. An NN consists of three main components, namely the input layer, the hidden layers and the output layer. The input layer is fed with data, the hidden layer processes the fed data while the output layer provides the result. Some NNs allow data to be transferred backwards as well as forward, others allow data to be exchanged between the neurons in the same layer or even allow a neuron to send data back to itself. These NN features affect its performance and the applications for which it is suited; hence the structure and its connection geometry are quite significant. The structure of the proposed NN is presented within Figure 3b. The NN includes X inputs, one hidden layer of eight neurons and one output.

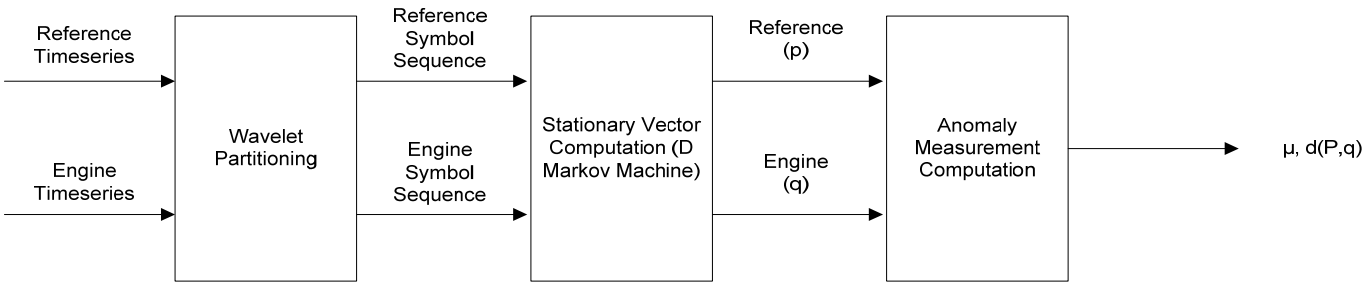


Figure 4: Symbolic Dynamic Filtering Steps

4.3 Training Rules

The training rules are used for the adjustment of the neurons' weights in order for the accuracy of the NN model to be maximized. Two main types of training rules exist, the supervised, and the unsupervised. The unsupervised method trains the network only with data, corresponding to the NN training input data. The NN identifies data inherent classifications. This type of training is usually used for clustering and estimation problems. In the case of supervised learning, the NN are provided by a series of matching input and output examples. Then the NN adjust their weights according to the training algorithm to minimize the deviation of the NN estimated output and the reference output. Supervised learning is preferred in the NN training for the function approximation problem. For the specific reason, this type of training is selected for the current training of the proposed NN.

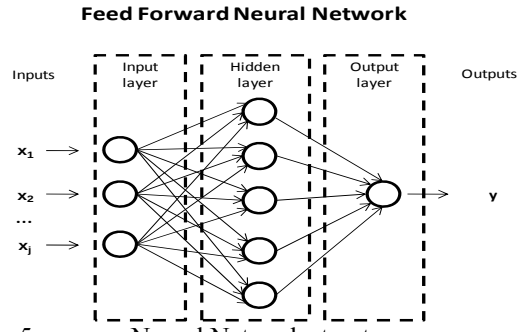


Fig. 5 Neural Network structure

5 Anomaly Detection

In Anomaly Detection, reference time series, produced by NN models and Real Engine time series in small periods (flights) are used as inputs. Time series are first transformed to symbol sequences for each one of which a corresponding stationary vector is produced, then an anomaly measure μ is computed by comparing the vectors. Computing μ for different time periods produces an μ series, used as input for the next step of RUL estimation.

The Symbolic Dynamic Filtering (SDF) is a statistical pattern recognition method for studying dynamical systems with continuous state spaces and continuous time using infinite sequences of abstract symbols. The system's continuous state space, in our case engine, is partitioned into a finite number of regions, each of which is characterized by a symbol. The time series of the engine parameter is transformed into symbols sequence. This specific sequence of symbols is the base for the construction of a probabilistic finite state machine, whose aim is the generation of pattern vectors as a representation of the dynamical system's characteristics. The SDF has been successfully applied to numerous different problems, varying from the description of dry friction oscillator dynamics up to our case engine degradation monitoring. The SDF method for anomaly detection consists of the following steps that are further analyzed in the next sections of the study.

- Wavelet Partitioning
- Stationary Vector Computation
- Anomaly Measurement Computation

5.1 Wavelet Partitioning

The input of this step is two time series, the reference time series produced by the NNs and the engine time series provided by the engine sensors. Of course, both time series take advantage of the same engine parameters (pressure or temperature), which are indicative of the engine's condition. At the end of this step, time series are transformed into symbol sequences as explained below.

The phase space of the dynamical system is partitioned into a finite number of cells. These cells are mutually exclusive and exhaustive so as to obtain a coordinate grid of the whole space. The partitioning is made by maximizing the Shannon entropy, in order for the regions with more information to be finer partitioned and those with sparse information to be coarser partitioned. Each cell is then labeled with a symbol. After that a trajectory of the dynamical

system, starting from an initial state, generates a sequence of symbols as the system changes phases. Let Σ be the alphabet (the set of all symbols that are used) and $|\Sigma|$ alphabet size (i.e. the number of all different symbols). Although this process is subjected to a loss of information it preserves essential robust features (e.g., periodicity and chaotic behavior of an orbit) through an appropriate partitioning of the phase space.

The theory of phase-space partitioning is well developed for one-dimensional mappings, however, the case of two or higher dimensional systems is much more difficult and the techniques proposed in literature [[11] and [12]] are primarily based on symbolic false nearest neighbor approaches. These approaches are not only extremely computation-intensive as the dimension of the system's phase space is growing but also in the presence of noise, a large alphabet of symbols is required as the number of false neighbors required is increased. Thus, a partitioning approach in the wavelet domain is more efficient for noisy data than for high dimensional dynamic systems. In the present approach the data have firstly been converted into the wavelet domain, then the wavelet space has been partitioned by maximizing the Shannon entropy and finally, the symbol sequence of the healthy system has been generated.

5.2 Stationary Vector Computation

In this step, by following the occurrence of each symbol of the incoming symbol sequence, a stationary vector for every sequence is produced containing statistical information of both the reference model and the engine. This is achieved through the D-Markov machine construction.

A D-Markov machine is a probability finite state machine that each of its states corresponds to a possible symbol word of length D that can occur from a symbol sequence of an alphabet set Σ. If N is the number of the states of the machine, it is clear that $N \leq |\Sigma|^D$. In this case, the states of the machine are the words with length D that are produced from the symbol sequence obtained by the system. The choice of $|\Sigma|$ and D depends on a specific application and so is the noise level of the data series as well as the available resources in memory and computation power. The small alphabet could miss information of the signal dynamic and a larger one may be noise sensitive. A larger value of D, although it gives better results, it increases the number of states exponentially requiring a much larger length of data sets and computational power. Example: $\Sigma=\{a,b\}$, $|\Sigma|=2$, $D=2$ *and the symbol sequence is: abaabba...*

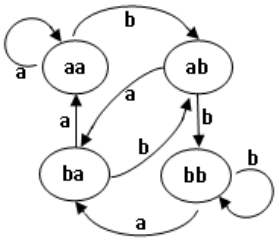


Fig. 6 Markov Machine

ab (the first two symbols of the sequence) and the machine is initially at *ab* state. As the third symbol of the sequence occurs (**a**) the machine transitions to the *ba* state are corresponding to the next produced word. The next words are *aa*, *ab*, *bb* and so on.

This probabilistic finite state machine has Markov properties as the probability of the machine transitioning from state *aa* to state *ab* is the conditional probability of the next symbol in sequence to be *b,* given the two previous were *aa***.** Moreover, the probability for the next state to be *ab* is the conditional probability of the next symbol in sequence to be *b* given the previous one was *a*.

The next step is the construction of a state transition matrix that describes all transition

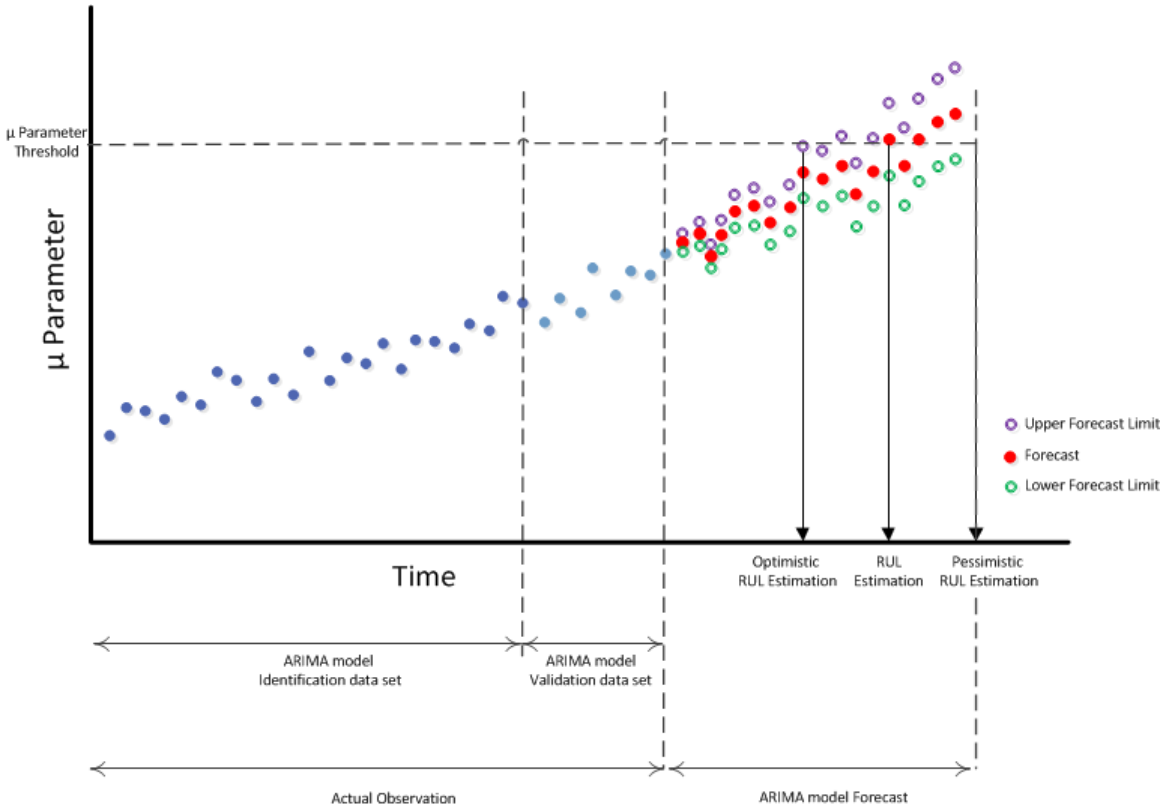


Fig. 7 RUL prediction using ARIMA models

probabilities among the states. Each element π_{ij} of this matrix corresponds to the conditional probability of transitioning to state j given the fact that the machine was in state i. In practice, each of these probabilities is computed by counting the occurrences of each transition for a "sufficiently large" length of symbol series.

Let Π be this transition matrix. Π is non-negative and irreducible. The Perron–Frobenius theorem [13] suggests that there is a unique vector q satisfying $q\Pi=q$. The elements of q are the stationary probabilities of each state of the D-markov machine.

A stop criterion for the length of symbols ℓ to be "sufficiently large" is equation 1, where e is the absolute error of re-computing vector q between the iterations of an increasing x.

$$\ell_{stop} = \text{int}(|\Sigma|^D / e) \ \textit{where} \ 0 < e << 1 \qquad (1)$$

ℓ_{stop} is considered a conservative criterion (even if it is the least conservative for the definition of e) for the length of symbol sequence, required for the parameter identification of the stochastic matrix Π and convergence is usually obtained earlier. It must be noted that in case of $D=1$, because of maximum entropy partitioning the nominal vector q has $|\Sigma|$ equal elements of value $1/|\Sigma|$.

Using the same D-Markov machine and counting the occurrences of each transition possibly a faulty symbol sequence is followed, and a new transition state matrix is constructed. Its unique eigenvector will have elements that correspond to the new stationary probabilities of the machine's states. This state vector p has the same length with q but it is expected to have non-equal elements.

5.3 Anomaly Measure Computation

The outcome of the previous step is the pair of $\boldsymbol{p}$ and $\boldsymbol{q}$ vectors, which are treated as statistical patterns. In the current step, they are compared in order to provide an indication of any anomalies that are deviations between the engine and the reference patterns. The outcome of SDF is the anomaly measure μ, a scalar function to quantify the system's deviation between nominal and test behavior patterns in a specific period. Anomaly measured is defined in

equation 2 as $\mu=d(p,q)$ which is the Euclidean distance between p and q

$$\mu = d(p,q) \qquad (2)$$

The anomaly measure μ is used for an early detection of anomalies as it can easily get high values for small changes in performances even with noisy data. In this study, the progress of μ through a larger timescale is used for the estimation of the Remaining Useful Time of the engine

6 Remaining Useful Life Estimation

The RUL is estimated with the use of ARIMA models. The SDF provides a time series of μ norm, the anomaly measurement of a parameter that is indicative of the engine's degradation. ARIMA models provide a representation of the μ norm future behavior. During the time that the μ norm prediction is greater than a critical threshold, the system is considered being into faulty condition and maintenance actions should take place.

6.1 ARIMA model

In the present approach, the Box and Jenkins ARIMA method is followed. ARIMA is a combination of three parts: an autoregressive (AR) part, a moving average (MA) part and a pert, involving the derivatives of the timeseries (I). The timeseries is derived k times until it is stationary, then a suitable ARMA model is fitted to the resulting series and the estimated forecasts are finally integrated k times [14].

$$x(t) = \sum_{c=1}^{m} a_c x(t-c) - \sum_{c=1}^{n} \beta_c \varepsilon(t-c) \qquad (3)$$

Where:

- x(t): the value of timeseries at time t, in our case, x(t) is the norm of μ at time t
- m: the parameter that denotes the order of the autoregressive part
- n: the parameter that denotes the order of the moving average part
- k: the number that the timeseries has been differentiated

6.2 RUL Forecasting

The ARIMA model determination comprises two main steps: a. model identification, b. model validation. Model identification is responsible for the determination of the ARIMA model parameters, m, n and k, i.e. the order of the autoregressive part, the order of the moving average part and the number that the timeseries must be differentiated. The observed data timeseries is divided into two parts. In order for the ARIMA model parameters to be identified, the first part, 70 % of the observed data, is analyzed. Afterwards, the model is validated with the second part of the observed data timeseries. If the fitting criteria, such as the RSS, are satisfied, then the ARIMA model parameters are kept, otherwise, another set of parameters is selected. This process is followed until the fitting criteria are satisfied. Thus, after the proper identification and validation of the ARIMA model, the model is capable of forecasting the μ norm values. Apart from estimating the values of the μ, the relevant confidence intervals with the upper and lower forecasting limit are also estimated.

7 Conclusion

In the current paper, a method for the estimation of the remaining useful life of a small turboprop engine is proposed. The method is based on the timeseries analysis that is critical for the parameters of the engine's condition. ANNs are proposed for modelling the engine in its sound condition, since they are capable of modelling the behaviour of complex and non linear systems. The SDF provides an assessment of the engine's health condition by analyzing the differences between the reference timeseries, corresponding to the nominal engine's condition and the monitoring engine's timeseries. The ARIMA model follows the SDF analysis, attempts to capture the trend of the engine's health and predict when the engine will reach a critical threshold. Based on the

time left to the critical threshold defined, the RUL is estimated. Thus, the paper proposes a method that fuses the SDF analysis with the ARIMA models in order for the reliability of the Remaining Useful Life estimation to be increased.

References

[1] ICAO, ATMCP-WG/C/2-WP/003, ICAO, (2004) available online from: http://www.icao.int/ [accessed 5 May 2008]

[2] Papakostas N, Papachatzakis P, Xanthakis V, Mourtzis D, Chryssolouris G, (2010) "An approach to operational aircraft maintenance planning, Decision Support Systems" 48 604–612

[3] Ray A (2004), Symbolic dynamic analysis of complex systems for anomaly detection, "Signal Processing", 84, 1115 – 1130.

[4] Gupta S, Ray A, Sarkar S, and Yasar M (2008). Fault detection and isolation in aircraft gas turbine engines. Part 1: underlying concept, *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering,* 222, 307-318

[5] Sarkar S, Yasar M, Gupta S, Ray A, and Mukherjee K (2008). Fault detection and isolation in aircraft gas turbine engines. Part 2: validation on a simulation test bed, *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering,* 222, 319-330

[6] Brotherton T and Johnson T (2001), Anomaly Detection for Advanced Military Aircraft Using Neural Networks, *Aerospace Conference, IEEE Proceedings,* 6, 3113-3123

[7] Brotherton T, Volponi A., Luppold, R, Simon D,(2003) eSTORM: Enhanced Self Tuning On-board Real-time Engine Model, *2003 IEEE Aerospace Conference Proceedings, Big Sky, MT*.

[8] Vachtsevanos G, Lewis F L, Roemer M, Hess and Wu B (2006). *Intelligent fault diagnosis and prognosis for engineering systems*. 1 edition John Wiley & Sons

[9] Marinai L, Singh R, Curnock B, Probert D (2003), Detection and Prediction of the Performance Deterioration of a Turbofan Engine, *Proceedings of the International Gas Turbine Congress 2003, Tokyo*

[10] Li Y G and Nilkitsaranont P, (2009) Gas turbine performance prognostic for condition-based maintenance, *Applied Energy* 86 2152-2161

[11] X. Kantz H and Schreiber T (2004). *Nonlinear time series analysis*, 2nd edition, Cambridge University Press, Cambridge, UK.

[12] Kennel M B and Buhl M (2003) Estimating good discrete partitions from observed data: symbolic false nearest neighbors. *Phys. Rev. Lett.*, 91(8), 84-102.

[13] Gradshteyn I S and Ryzhik I M (2000). *Tables of Integrals, Series, and Products*, 6th edition, CA: Academic Press, San Diego

[14] Box G E P and Jenkins G M (1976) *Time series analysis: Forecasting and control*, 2nd edition San Francisco: Holden-Day

[15] Chryssolouris G (2006) *Manufacturing Systems: Theory and Practice*, 2nd Edition, Springer-Verlag, New York, New York

[16] Efthymiou K, Georgakakis P, Papakostas N, and Chryssolouris G (2011), Small Turboprop Engine Health Management, ICEAF II, 2nd International Conference of Engineering Against Fracture, Mykonos, Greece (2011)

[17] Marinai L, Probert D and Singh R (2004). Prospects for aero gas-turbine diagnostics: a review, *Applied Energy79, 109-126*

[18] Sikorska J Z, Hodkiewicz M and Ma L (2011) Prognostic modeling options for remaining useful life estimation by industry, *Mechanical Systems and Signal Processing 15: 1803-1836*

[19] Rajagopalan V and Ray A (2006) Symbolic time series analysis via wavelet-based partitioning. *Signal Process, 86, 3309-3320*

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

HECTOR: one more step toward the embedded Structural Health Monitoring system

C. Sbarufatti, A. Manes and M. Giglio
Politecnico di Milano, Italy

Keywords: *Structural Health Monitoring, Artificial Neural Network, Helicopter, Fuselage, Crack*

Abstract

The helicopter fuselage crack monitoring through on-board expert sensor network is matter of HECTOR, a project coordinated by the European Defense Agency (EDA) in the framework of the Joint Investment Program on Innovative Concepts and Emerging Technologies (JIP-ICET) which is funded by 10 EDA Member States (Cyprus, France, Germany, Greece, Hungary, Italy, Poland, Slovakia, Slovenia, Spain) and Norway.

The aim of HECTOR is to increase the machine availability by directly monitoring the damage while it is propagating inside the structure, concept at the basis of Structural Health Monitoring (SHM) systems. In this way it would be possible to get a real time knowledge about the damage situation, thus setting a Condition Based Maintenance (CBM).

The current paper is a summary of the advances in the project, covering all the involved engineering areas, from the adopted sensor network technology, to the FEM simulation and coupling with algorithms, thus coming to the definition of a robust methodology for SHM design. The possibilities for result exploitation, as well as the points requiring further advances, are going to be presented.

1 General Introduction

The main purpose of this project is to propose modeling techniques for the analysis of multiple sensors in producing a unified and comprehensive method for identification, monitoring and prognosis of potential damages (like cracks) in the fuselage of a helicopter. So far the structural safety of helicopters is ensured with a thorough fatigue analysis in the design phase and a strict schedule of inspection during life. Only a few attempts to apply reliable methods to monitor directly the on-line damage from vibrations during life have been carried out. In particular for military applications, an integrated and reliable system for monitoring possible damages by interpreting service stresses inside the fuselage and for evaluating the time inspections and remaining life (prognosis) is missing. The motivation for suggesting Structural Health Monitoring SHM is the potential capability to reduce the maintenance cost, which is about 25 per cent of the direct operating cost of the helicopter and can play an important role especially in the case of the ageing helicopters. The purpose of HECTOR is the creation of a real case demonstrator for helicopters that permits, on the basis of a complete knowledge of the stress state on the fuselage (advanced FE model) and the

data collected by a low power sensor network positioned in the most critical areas, to perform the Embedded Structural Health Monitoring and the on-line evaluation of the degradation mechanism, based on advanced damage criteria. This sensor systems could be mounted on new helicopters to define a more reliable, efficient and economic program of inspections, or on ageing airframes to perform a life extension of the machines. The advantage of HECTOR is that once the sensors network is installed and set up, inspection is possible not only without disassembling parts but also in continuum using a real time strategy to detect failures (existence, location, type, extent) and for the prognosis of their effect on the overall reliability of the structure.

The basic premise of most damage detection methods is that damage will alter the stiffness, mass, or energy dissipation properties of a system, which in turn alter the measured dynamic response of the system. Although the basis for damage detection appears intuitive, its actual application poses many significant technical challenges. The most fundamental challenge is the fact that damage is typically a local phenomenon and may not significantly influence the global response of a structure that is normally measured during vibration tests. Another fundamental challenge is that in many situations damage detection must be performed in an unsupervised learning mode. Here, the term unsupervised learning implies that real data from damaged systems are not available. Environmental and operational variations, such as varying temperature, moisture, and loading conditions affecting the dynamic response of the structures will in addition complicate the interpretation of possible damage.

The present paper is the prosecution of the work described in [1], with the aim to describe the steps further in HECTOR project, finally highlighting the most critical points where advances are mostly needed.

2 SHM system method

The SHM method proposed hereafter consists of a combination of simulated data coming from detailed Finite Element Models (FEM) and smart algorithms able to learn the underlying functions that correlate a certain damage sensitive parameter to the proper damage characterization. One could notice that it would be possible to set up a damage detection system just based on algorithms, able to distinguish a healthy case from a critical one just relying on some assumed thresholds. However, given the algorithm capability in solving this crucial step with a sufficient reliability, it would be very difficult to physically appreciate the damage characteristics that induced the modification on the acquired sensor data, thus generating an alarm. As a matter of fact, it is crucial to consider that, in order to be able to set up an adaptive maintenance schedule, the damage should be characterized not only in its existence, but also in its type, position and extent at least.

For this reason, the introduction of FEM, as a source of simulated reference data, should help the algorithms in associating a certain physical meaning to the generated alarm. Given the installation of a network of sensors inside the monitored region of the fuselage, a wide database of many different damages, in many different positions and with various levels of criticality can be modeled, thus generating a low cost basic information, upon which to train and help the smart algorithms in interpreting real sensor data.

Thus, the output of this FEM based system would allow another family of reasoning algorithms, namely the prognostic unit, to estimate the Residual Useful Life (RUL) of the monitored component or region inside the fuselage. Obviously the statistical propagation of the uncertainties relative to the diagnostic level inside the prognostic unit has to be carefully considered, in order to keep the reliability of the whole system at the desired level.

2.1 Choice of damage sensitive parameter

In the most general terms, damage to a structure can be defined as changes introduced into a system that adversely affects its current or

future performance. Implicit in this definition is the concept that damage is not meaningful without a comparison between two different states of the system, one of which is assumed to represent the initial, and often undamaged, state. Here it is focused on the damage identification in structural and mechanical systems and is limited to changes to the material and/or geometric properties of these systems, including changes to the system connectivity. As an example, a crack that forms in a mechanical part produces a change in geometry that alters the stiffness characteristics of that part. Depending on the size and location of the crack and the loads applied to the system, the adverse effects of this damage can be either immediate or may take some time before they alter the system's performance. In terms of length scales, all damage begins at the material level and then under appropriate loading scenarios progresses to component and system level damage at various rates. In terms of time scales, damage can accumulate incrementally over long periods of time such as that associated with fatigue or corrosion damage accumulation. Damage can also result from scheduled discrete events such as aircraft landings and from unscheduled discrete events such as enemy fire on a military vehicle.

In particular, considering HECTOR's framework, the objective is to set up a diagnostic / prognostic system based upon strain field modifications due to the occurring damage, focusing the attention on fatigue type damages on metallic structures (cracks). Even thinking about bullet impacts in military applications, the assumption is that the evolution of the damage would be however a crack generated at the occurring notches. In [2] was demonstrated that the strain field sensitivity to crack damage can be exploited for structure monitoring; it was also confirmed the possibility to accurately model with Finite Elements the static and fatigue dynamic behavior of a stiffened structure. However, the main issue was the strain sensitivity also to the applied load. In fact, the helicopter fuselage is experiencing a load that is absolutely varying in time and, even worse, not easily predictable, due to the environmental influences. The approach used hereafter consists in normalizing each sensor of a confined region with respect to the average value measured among all the sensors in the same region [3]. In fact, the normalizing factor should be robust to the presence of the damage but only sensible to the applied load, thus allowing filtering the load effect, without significantly altering the sensitivity to the crack. Assuming *N* strain sensors measuring exactly in the same direction, the expression of the normalized output for the *k-th* sensor is reported in Eq. (1).

$$\varepsilon_k^{norm} = \frac{\varepsilon_k}{\sum_{i=1}^{N} \varepsilon_i \Big/ N} \qquad (1)$$

Then, focusing the attention on performances, it is important to clarify the intention of the on-board SHM. In fact, classical Non Destructive Technologies (NDT) can detect cracks at their very early stage of propagation, often detecting anomalies in the length order of 1mm or less. On the other hand, the on-board detection system is expected to be designed for a target crack length around 60mm. The difference lays in the associated maintenance procedure. If from one hand classical NDT technologies are manually used to scan the structure at fixed time intervals during the component / region life, HECTOR SHM is designed in order to be permanently installed and continuously operating to monitor in real time the structure. Thus, considering the load carrying capabilities of the considered region (regulations require the ability to carry the design loads when 2-bay long cracks are present), being able to detect automatically a 60mm crack with an high level of reliability would allow to increase the machine availability, thus reducing the operating cost of the machine.

2.2 The noise definition

The key point for the success of the SHM methodology described above is the correct definition of noise affecting the SHM system. In

fact, being the algorithms based upon a deterministic knowledge coming from FE simulations, each source of deviation from FEM must be included inside the noise characterization.

The following are the main reasons for real data dispersion with respect to FEM prediction on the damage sensitive parameter measures:

- Manufacturing process
- Material non-uniformities
- Non Linearity (other than damage)
- Environmental influences
- Sensor layout definition
- Unpredictable loading condition
- Non considered damage parameters (i.e. crack angles inside the current work)

It is also important to consider that each one of the reported factors, entering in the measure as a source of perturbation, must be characterized in its extent as well as in its shape. In fact, most of them can assume a Gaussian distribution, as demonstrated in [2], where some panels were tested in their damage tolerant capabilities. The mean of all the considered specimens was perfectly consistent with the FEM prediction, though a certain level of deviation was present, due to panel manufacturing, material non-uniformities, environmental influences and uncertainty in sensor positioning / alignment. As a matter of fact, all these factors can be considered Gaussian in their distribution around FE prediction.

On the other hand, considering the influence of crack angle on FE strain estimation, one could notice that, given the angle variation remains in the ±5° range, the error in damage sensitivity stays below 2%, though the shape of the error distribution is not Gaussian anymore.

Finally, the influences of random load on the damage detection parameter are depicted inside the next paragraph.

2.3 The problem of unpredictable load

As previously introduced in paragraph 2.1, strain field is sensitive in both damage existence and external applied load. Thus, the parameter described in Eq. (1) was introduced as a mean to filter out the load effect, but without reducing the capability in describing the damage sensitivity. Given a network of 15 sensors is considered, all equally-directed, the strain field sensitivity to a growing crack damage is reported in Fig. 1 for one single sensor, normalized with respect to the mean value calculated for the whole sensor network. In particular, the continuous line was obtained when a constant 30kN load was applied to the structure, while the dotted distributions are for random loads ranging from 3 to 30kN.

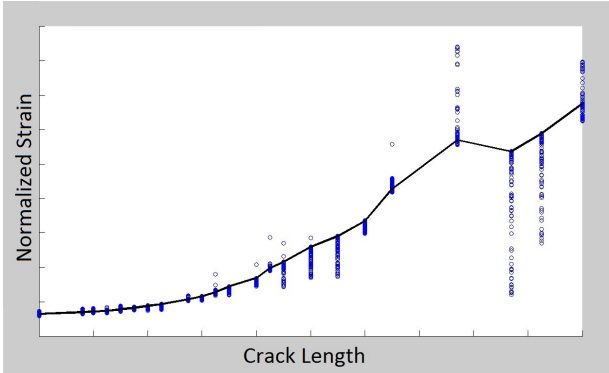


Figure 1: Strain sensitivity to growing crack damage.

As can be easily noticed, normalizing with respect to something that is “less sensitive” to damage and “more sensitive” to load, one could reduce the load effectiveness while highlighting the damage weight, at least for relatively small damages.

3 The Finite Element Model

In order to demonstrate the SHM methodology, it has been decided to realize a structure which is similar to the real helicopter modeled in [1], eliminating some particulars, such as panel curvature, not required for the purposes of the project. The demonstration panel, constituted by a skin stiffened with four stringers, is shown in Fig. 2. In particular, it is possible to notice how the strain field is modified because of the presence of a skin crack

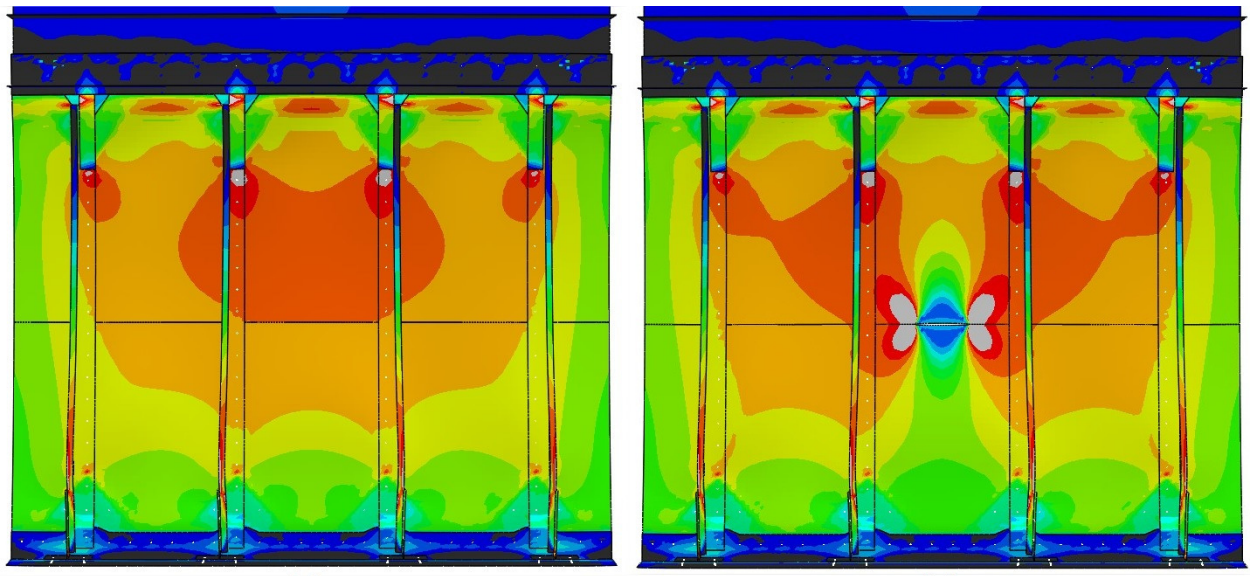


Figure 2: FE model of the demonstrator panel (a) in healthy condition and (b) with crack on a skin bay

located in the middle of the central bay. The aim is to create a database of strain maps generated by different damages, each one with different characteristic parameter. In particular, two damages were considered, respectively consisting of a typical skin crack and a stringer failure (with crack starting from the position where the stringer was broken). For each case, the two parameters of crack length and position were made vary, thus calculating more than one thousand cases.

The generated database will be first used in order to design the sensor network, thus choosing the sensor number and positions. However, given the symmetrical and repetitive structure, it has been decided to limit the sensor locations along 4 paths along the stringers and 3 paths in the middle of each bay, leaving the sensor number parameter free but equal inside the stringer / bay family (all the stringers / bays have respectively the same number of sensors applied).

The key point for the success of the whole SHM methodology is the experimental validation of the FE model, which should be able to well describe not only the strain field as a function of applied load, but also the strain field dependence on the considered damage. However, it has been demonstrated that the damage model follows very well the real experimental data, at least sufficiently far away from the boundary regions.

4 The algorithm structure

The SHM algorithm is mainly constituted by two parts, namely diagnostic and prognostic units. The diagnostic part has been structured as follow [4]:

- *Anomaly detection*: the algorithm must be able to recognize that damage is occurring inside the structure, without any concern about the damage type, position or extent.
- *Damage characterization*: the type of damage is identified, distinguishing, for the current case, between simple skin crack or stringer failure.
- *Localization and quantification*: given the damage type is recognized, an algorithm is optimized to estimate its position and extent, thus setting the basis for the prognostic unit inference.

Some details are presented in the next paragraph, highlighting the expected sensor network performances as well as the design methodology. As a matter of fact, FE data are relatively low cost information, upon which it is possible to design the SHM hardware and software.

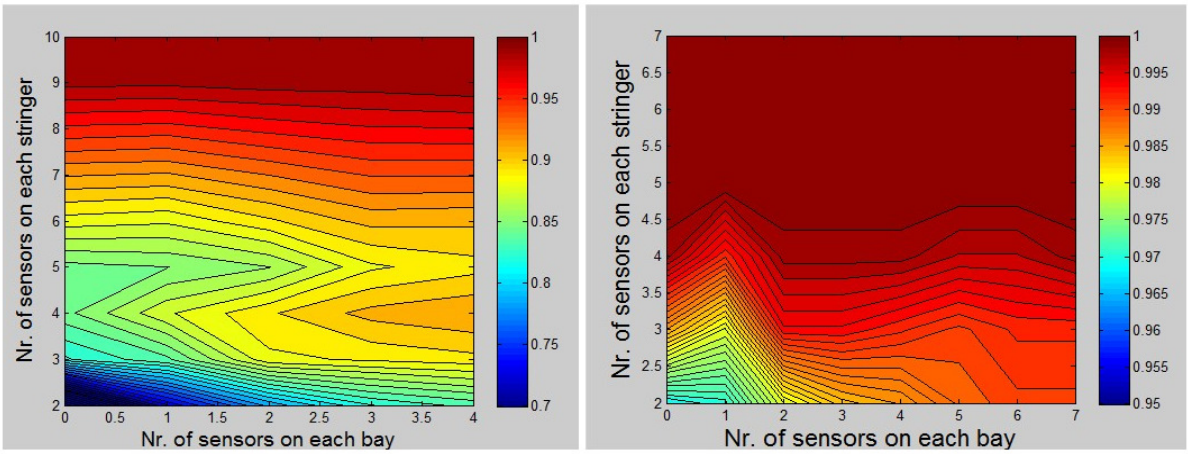


Figure 3: PCA as a function of sensor number on each bay / stringer (a) for a 60mm skin crack (b) for a stringer failure plus 60mm crack.

4.1 Anomaly Detection

This is the first step of the diagnostic algorithm, based upon Artificial Neural Network (ANN) theory [5, 6]. ANNs appear as particularly suited for HECTOR application. The reason lays firstly in the fact that they can be trained with artificial data, relative to the damages included inside the database introduced above. Moreover, once they have been trained off-line, their simple structure is particularly attractive for on-board application, requiring a minimum computation time, thus allowing to perform a real time structural monitoring with low power consumption.

In particular, considering the Anomaly Detection step, the MATLAB pattern recognition tool was identified as the best candidate for HECTOR purposes. It received as input the FE simulated strain pattern during the training phase, setting as output a binary number indicating whether the structure is healthy or damaged. In particular, a certain amount of noise was introduced for the undamaged case in order to produce a population of probable healthy states, to be distinguished by all the available damaged cases. The so trained ANN could now be used to process other simulated and noised strain patterns, thus evaluating the algorithm performances as a function of the sensor number on each stringer or bay.

Results are reported in Fig. 3, where two sensitivity plots for network performance as a function of sensor number and positions are reported, respectively when (a) skin crack damages or (b) stringer failures are considered. In particular, performances are evaluated through the Probability of Correct Assumptions (PCA) parameter, calculated as in Eq. (2).

$$PCA = \frac{Correct \quad Decisions}{Total \quad Analized \quad Cases} \qquad (2)$$

To be noticed that PCA takes into account the occurring of false alarms (the structure is healthy but an alarm is provided) as well as missed detections (damage exists but it is not identified). It is appreciable that stringer locations are more effective than bay positions for the optimization of network efficiency. Moreover, as expected, the bottleneck for the required sensor number is the necessity to detect skin cracks, while stringer failures are easier to be detected, thus requiring fewer sensors.

4.2 Localizing and quantifying damage

Given the alarm was generated at previous diagnostic steps, the damage must be characterized in terms of position and extent. This step is crucial as the following prognostic algorithm will be based upon the results produced here, thus propagating the level of uncertainties inside the estimation of Residual

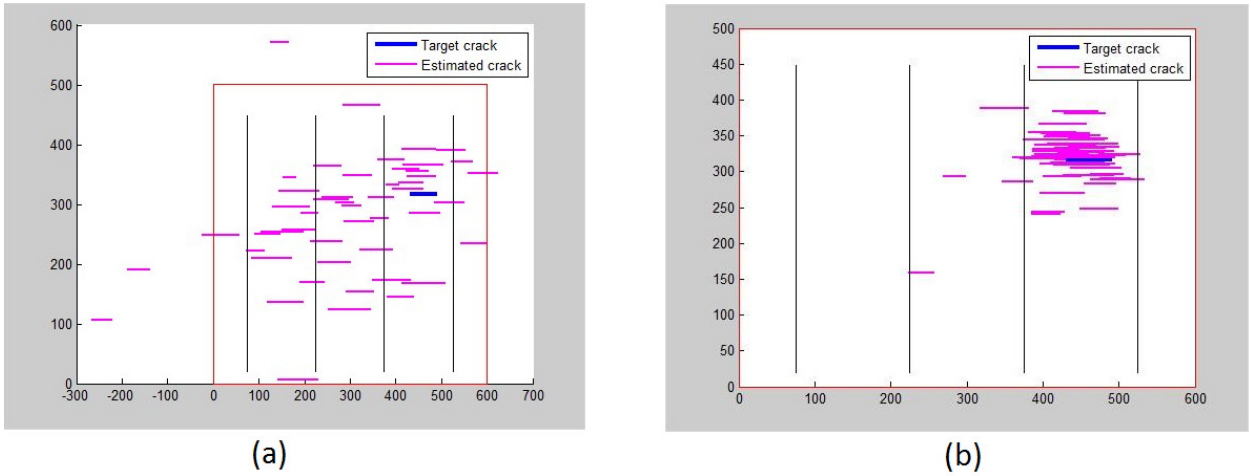


Figure 4: performances of ANN trained to recognize skin crack position and length over the schematized panel (a) with no noise in training (b) with estimated noise added in training phase.

Useful Life (RUL) of the component or sub-region under analysis.

The MATLAB Function Fitting tool is used herein, trained in order to associate a strain pattern given as input (only damaged cases are included at this level) with 3 continuous variables respectively indicating:

- X-Position of crack center
- Y-Position of crack center
- Crack Length

In particular, it is important to add a certain level of noise to the deterministic FE model, in order to increase the performances of the trained ANN. Results for one crack case are reported in Fig. 4, when the noised simulated strain pattern produced by the same target crack is processed with an ANN (a) trained with pure FEM data and (b) trained with noised FEM database. The vertical lines and the rectangular shape respectively represent the stringers and the skin panel.

A *7-sensor-per-stringer* configuration was selected from paragraph 4.1, and tested herein. It has been demonstrated that the ANN is able to estimate the position and extent of the damage, though some advances in post-processing are needed concerning damage quantification. Moreover, stringer failure localization is again by far more efficient than simple skin crack localization, due to the larger effect on strain field modifications.

5 Preliminary test results

Some tests have been recently executed in LTSM-UP (*Laboratory of Technology and Strength of Materials – University of Patras*), monitoring the area of the demonstrator panel through the sensor network designed ad selected in Ch. 4, that is to say the *7-sensor-per-stringer* configuration to obtain a good level of detection for a 60mm skin crack. Fiber Bragg Grating sensor were used in order to minimize the impact on the helicopter structure. The output of the Anomaly Detection algorithm receiving real measurements as input with a growing skin crack is reported in Fig. 5.

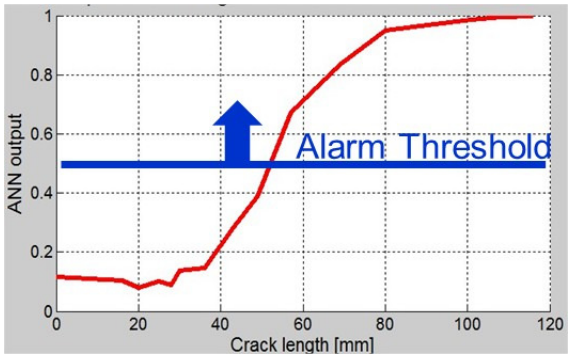


Figure 5: ANN output for a growing skin crack case

As one can notice, the system has been able to detect the damage, generating an alarm, when the crack length was between 50mm and 60mm, in agreement with the design assumptions.

6 Conclusions and forthcoming work

The results obtained up to now encourage continuing in this research. In particular, about 95% of Probability of Detection (PoD) with less than 5% Probability of False Alarm (PFA) seems to be the limit obtained for the selected sensor network (*7-sensors-per-stringer*) in recognizing 60mm long cracks all around the monitored structure. However, performances can be improved increasing the sensor number and, more efficiently, improving the FE model, in particular concerning modeling of strain field near boundaries.

However, deeper analysis are scheduled, both at software and hardware level, connected with further tests on the same structures with different sensor technologies (FBGs, Strain Gauges, Acoustic Wave Piezo-electric network, Crack Gauges, Comparative Vacuum Monitoring). HECTOR is expected to be ended in December 2011, though the ASTYANAX (*Aircraft fuSelage crack moniToring sYstem And progNosis through on-boArd eXpert sensor network*) follow up is going to be proposed to increase the Technology Readiness Level (TRL) of the research, aiming to apply on-board the whole HECTOR SHM methodology to real components and regions for real time aircraft SHM.

Acknowledgements

The authors gratefully acknowledge all the partners contributing to the HECTOR framework for their fruitful collaboration, in particular AGH (M. Sikora, J. Gozdecki, Poland), AgustaWestland (Italy), Consorzio Milano Ricerche (Italy), Laboratory of technology and Strength of Materials (I. Diamantakos, C. Katsikeros, G. Labeas, University of Patras, Greece), Sintef (Norway), University of Zilina (Slovakia) and Vitrociset (Italy).

References

[1] C. Sbarufatti, A. Manes, M. Giglio et al., *Application of Structural Health Monitoring over a Critical Helicopter Component*, Proc. AIRTEC – HELI World (2010).

[2] C. Sbarufatti, A. Manes and M. Giglio, *Probability of detection and false alarms for metallic aerospace panel health monitoring*, Proc. 7th Int. Conf. on CM & MFPT, BINDT, (2010).

[3] C.E. Katsikeros and G.N. Labeas, *Development and validation of a strain-based Structural Health Monitoring system*. Mechanical Systems and Signal Processing, Vol. 23/ 2, (2009).

[4] K. Worden and A.P. Burrows, *Optimal sensor placement for fault detection*, Engineering Structures 23 885–901 (2001).

[5] C. Bishop, *Neural Networks and Pattern Recognition*, Oxford University Press, (1995).

[6] T. Mitchell, *Machine Learning*, McGraw Hill, (1997).

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

System reliability analysis: a Model-Based approach and a case study in the avionics industry

Alfredo Garro and Andrea Tundis
Department of Electronics, Computer and System Sciences (DEIS), University of Calabria, Via P. Bucci 41C, 87036, Rende (CS), Italy
{garro,atundis}@deis.unical.it

Nicola Chirillo
Z Lab Engineering S.r.L., Via Lombardia 1, 37044, Cologna Veneta (VR), Italy
n.chirillo@zeta-lab.it

Keywords: *Reliability Analysis, Model-Based System Engineering, System Modeling and Simulation, SysML, Aircraft Systems*.

Abstract

The increase in both system complexity and accuracy required in the reliability analysis of mission-critical systems requires new, more powerful and flexible analysis tools and techniques. In this context, the paper proposes a model-based approach for system reliability analysis which combines in a unified framework the benefits of popular OMG modeling languages (UML, SysML) with the wide adopted Mathworks simulation and analysis environments (Matlab, Simulink). To show the flexibility and scalability of the proposal as well as its effectiveness in valuating and evaluating through simulation the system reliability performances, a case study, concerning the reliability analysis of a Landing Gear System of a military transonic trainer aircraft, is also presented and discussed.

1 Introduction

Dependability, which represents an important requirement to be satisfied for a wide range of systems, becomes even crucial when a system failure may lead to catastrophic losses in terms of cost, environmental damages, or human lives, as in several industrial and service domains such as avionics, railway, automotive, energy and health care [7].

According to the definition provided by the IEC (International Electrotechnical Commission), Dependability is "the collective term used to describe the availability performance and its influencing factors: reliability performance, maintainability performance and maintenance support performance". Thus, the engineering discipline which aims at providing an integrated and methodological approach to deal with system dependability is commonly indicated by the acronym RAMS (*Reliability*, *Availability*, *Maintainability* and *Safety*). *RAMS analyses*, whose main objective is to identify causes and consequences of system failures, are typically facing using a layered approach and through both *quantitative* and *qualitative* analysis techniques as *series-parallel system reliability analysis*, *Markov Chain models*, *FMECA* (*Failure Modes Effects and Critical Analysis*), and *FTA* (*Fault Tree Analysis*) [3, 12].

The increase in both system complexity and accuracy required in the reliability analysis often makes inadequate the above mentioned techniques which are mainly based on statistical and probabilistic tools and on the hierarchical decomposition of the system in terms of its components. Thus, there is a strongly demand

for new, more powerful and flexible analysis tools and techniques which are also centered on model-based approaches so to benefit from the available modeling practices and which incorporate the use of simulation to flexibly evaluate the system reliability indices and compare different design choices.

Despite a general consensus on the advantages that could derive from the exploitation of model-based approaches for system reliability analysis, the use of these techniques has been traditionally unusual and has not been recommended by international standards until recently [2, 6, 11]. This delay in the adoption is mainly due to the lack of methods able to integrate available modeling languages, tools and techniques in a consistent modeling framework.

In this context, the paper proposes a model-based approach for system reliability analysis which is centered on a popular UML-based language for system modeling (*SysML*) [4, 17] and on a *de facto* standard platform for the simulation of multi-domain dynamic and embedded systems (*Mathworks Simulink*) [8].

In particular, the approach is centered on a classical iterative process [16] which consists of four main phases: *Objectives Definition*, *System Modeling*, *System Simulation*, and *Results Assessment*. In the first phase, the objectives of the reliability analysis are specified; then, in the *System Modeling* phase, the structure and behavior of the System are modeled by using a *SysML* based notation; moreover, specific behaviors, which model the onset, propagation and management of failures, are introduced (a wide set of basic failure behavior patterns have been defined). In the *System Simulation* phase, the previously obtained *Models* of the system are represented in terms of the constructs offered by Mathworks Simulink. Finally, simulation results are analyzed with respect to the objectives of the reliability analysis and, if necessary, new partial or complete process iterations are executed.

The proposed model-based approach has been applied in the reliability analysis of the *Landing Gear System* of a military transonic trainer aircraft produced by a leading European Avionics Industry (for reasons of confidentiality, the names of the Industry and of the aircraft are not reported). The flexibility and scalability of the model-based approach allowed valuating and evaluating several configuration scenarios and settings of system parameters; the analysis of the simulation data provided useful indications to obtain a more descriptive and predictive reliability system model and suggested design choices which could allow improving the system reliability indicators.

The rest of the paper is structured as follow: Section 2 introduces the reliability analysis discipline along with the main reliability functions and indicators; Section 3 presents the proposed model-based approach to system reliability analysis; Section 4 shows the application of the proposal to the reliability analysis of the landing gear system of a military transonic trainer aircraft; finally, conclusion are drawn and future work delineated.

2 System Reliability Analysis

The main objective of the System Reliability Analysis is to identify the causes and consequences of failure of a system and its components [3]. Specifically, *Reliability* represents the ability of a system to perform its required functions under stated conditions, identified during its design, for a specified period of time. This property is strongly related to the following ones: *Availability*, which is the proportion of time a system is in a functioning condition defined at design time; *Maintainability*, which represents the ease with which maintenance of a system can be performed in accordance with prescribed requirements; *Safety*, which takes into account the effects of the system on its surrounding environment to prevent, eliminate and control hazards.

The formal definition of *Reliability* relies on the important concept of *time-to-failure* which is modeled as a random variable T and by the following related functions:

1) Reliability function, which indicates the probability that the system fails after time t:

$$R(t) = Pr(T > t) \quad (1)$$

2) Failure distribution, which is *a cumulative distribution function* and indicates the probability that a system fails before time *t*:

$$F(t) = Pr(T \leq t) = 1 - R(t) \quad (2)$$

3) Failure density function, which indicates how the probability of failure is distributed over the life of the system:

$$f(t) = \frac{d}{dt} F(t) = \dot{F}(t) \quad (3)$$

4) Hazard function (or hazard rate), which indicates the probability of a system failure between *t* and $t+\Delta t$ given that it was operating at time *t*, and becomes the *instantaneous failure rate* as Δt tends to zero:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{R(t) - R(t+\Delta t)}{R(t) \cdot \Delta t} = \frac{-\dot{R}(t)}{R(t)} \quad (4)$$

For a wide range of mechanical systems the function $\lambda(t)$ assumes a typical shape called *Bathtub curve* which shows three different stages of failures during the system life:

– *early failures*, the failure rate is high due to the infant mortality of system components but rapidly decreasing;

– *random failures*, the failure rate is low and constant (useful life of the system);

– *wear-out failures*, the failure rate increases as age and wear take their toll on the system components.

5) Mean Time To Failure (MTTF), which represents the expected value of the *time-to-failure* random variable *T*:

$$MTTF = E(T) = \int_0^{\infty} t f(t)\, dt \quad (5)$$

In Table I the expressions of the above introduced functions in terms of each of them are reported. The valuation and evaluation of these reliability functions represent a crucial task for a quantitative analysis of the reliability properties of a system. Beside quantitative analyses, qualitative analyses, which aim to identify the possible system failures, their rate of occurrence, and (local and global) effects on the system, should be also executed [12] so to complement the information obtained from the quantitative ones.

Table I. – Main Reliability Functions

	f(t) Failure density function	*F(t) Failure distribution*	*R(t) Reliability function*	*λ(t) Hazard function*	*MTTF Mean Time to Failure*
f(t)		$\int_0^t f(\tau)\, d\tau$	$\int_t^{\infty} f(\tau)\, d\tau$	$\frac{f(t)}{1 - \int_0^t f(\tau)\, d\tau}$	$\int_0^{\infty} t\, f(t) dt$
F(t)	$\frac{d\,F(t)}{dt}$		$1 - F(t)$	$\frac{d\,F(t)}{dt} * \frac{1}{1 - F(t)}$	$\int_0^{\infty} (1 - F(t)) dt$
R(t)	$\frac{-d\,R(t)}{dt}$	$1 - R(t)$		$\frac{-d\,R(t)}{dt} * \frac{1}{R(t)}$	$\int_0^{\infty} R(t) dt$
λ(t)	$\lambda(t) e^{-\int_0^t \lambda(\tau) d\tau}$	$1 - e^{-\int_0^t \lambda(\tau) d\tau}$	$e^{-\int_0^t \lambda(\tau) d\,\tau}$		$\int_0^{\infty} e^{-\int_0^t \lambda(\tau) d\,\tau} dt$

3 A Model-Based approach for System Reliability Analysis

The proposed model-based approach, which aims to support both quantitative and qualitative system reliability analyses, relies on an iterative process which consists of the following four main phases: *Objectives Definition*, *System Modeling*, *System Simulation*, and *Results Assessment*.

In the *Objectives Definition* phase, the objectives of the system reliability analysis are specified. The inputs of this phase are the work-products resulting from the System Design and the reliability system requirements. Starting from this documentation, the functions that the system has to perform, the related operative conditions, and the reference time horizons must be clearly individuate along with the main system failures and their local and global effects. Finally, the reliability functions and indicators, to be derived from the analysis of the simulation results, must be identified along with the main analysis techniques to be applied to the data gathered from simulation.

In the *System Modeling* phase the structure and behavior of the System are modeled by using a *SysML* based notation. In this phase the System is decomposed in component entities by applying *in-out zooming mechanisms* [10]. In particular, the following main decomposition levels should be considered: *system*, *subsystems*, *equipment*, and *components*; however, different and deeper hierarchies can be also introduced. Each component entity is modeled as a SysML

Block whose structure is defined by both a *Block Definition Diagram* (BDD) and an *Internal Block Diagram* (IBD) and whose behavior is defined by a *SysML Activity Diagram*. Specifically, the BDD describes the *Block* with its port interfaces, internal attributes, operations, constraints, parts and relationships with other blocks; whereas, the IBD provides a description of the *Block* internal structure, the organization of its component blocks, the type of composition and the topology of internal communication. The behavior of a Block is specified trough a set of *Tasks* whose execution is characterized by pre and post conditions and can be periodically scheduled or triggered by events [14]. Each *Task* is modeled as a *SysML Activity Diagram* and consists in a flow of *actions*. Finally, special T*asks*, which model the onset, propagation and management of Block *failures*, are specified (a wide set of basic failure behavior patterns have been defined). In addition, *SysML Parametric Diagrams* are introduced for supporting specific analysis by defining constraint blocks which express mathematical equations and their parameters that may correspond to block properties.

In the *System Simulation* phase, the previously obtained *Models* of the System are represented in terms of the constructs offered by Mathworks Simulink, the framework which is currently exploited for the simulation. The model transformation is enabled by IBM Rational Rhapsody [5], a model-driven development environment based on UML and SysML which provides an integrated modeling environment for SysML and allows to generate Simulink models starting from SysML ones. This transformation is based on a mapping between the basic SysML and Simulink constructs [15, 18]; in particular: (i) a (simple) SysML Block is transformed into a Simulink Block; (ii) a (composite) SysML Block, consisting of other blocks (its parts), is transformed into a Simulink Subsystem Block; (iii) SysML FlowPorts are transformed into Input and Output Simulink Blocks; (iv) SysML Flow Specifications, used to type FlowPorts, are transformed into Simulink Bus Objects. Moreover, the Mealy Machines which model the behavior of a Simulink Block is obtained by the corresponding SysML Activity Diagrams. The resulting Simulink System is a network of blocks which is executed according to a synchronous reactive model of computation [1]: at each step, Simulink computes, for each block, the set of outputs as a function of the current inputs and the block state, then it updates the block state.

In the *Results Assessment* phase, the data gathered from the simulations are analyzed with reference to the objectives of the reliability analysis identified in the initial phase of the process. Specifically, several analyses can be directly performed in Simulink by using useful add-on like *SIMLOG*, whereas more advanced analysis can be performed by external analysis tools by exporting the obtained results though the Matlab Workspace. As for any iterative process, new (partial or complete) iterations can be executed for achieving new or missed analysis objectives.

4 A case study in the avionics industry

In this Section, the exploitation of the model-based approach presented in Section 3 is exemplified thought a case-study concerning the reliability analysis of a Landing Gear System of a military transonic trainer aircraft [9]. In particular, a synthetic description of the System is given in Section 4.1, whereas the subsequent subsections are related to the different phases of the proposed process.

4.1 System Description

Aircraft landing gears can be of various kinds and sizes [9] depending on the aircraft typologies. The Landing Gear System under consideration has the following main characteristics: (i) *tricycle* as it consists of three independent wheels; (ii) *retractable* as during the flight its wheels are stowed in special accommodation for improving the aerodynamics of the aircraft; (iii) *telescopic* as it comprises telescopic struts which are capable of extension. The System consists of the three following subsystems:

- Left Main Gear;
- Right Main Gear;
- Nose Gear.

The first two subsystems are located within compartments of the aircraft under the fuselage just behind the main air intakes and are composed of a single wheel on which operates a hydraulic brake. The third subsystem is placed under the nose of the aircraft and it is equipped with one wheel steering moved by a hydraulic system. The operations of extension and retraction are controlled by selector valves and hydraulic actuators that are powered by another hydraulic power system, whereas an accumulator provides the hydraulic power for the brakes or for the landing gear in emergency situations.

Each subsystem is composed of various equipment which in turn consist of several components. As an example, the Nose Gear subsystem consists of the following equipment: *Nose Gear Assembly, Extension and Retraction, Emergency Extension, Nose Wheel, Nose Wheel Steering*. In turn, the *Nose Wheel Steering* equipment consists of the following main components: *Hose Flexibles, Pipes, Wiring and Connectors, Attachment Parts, Pedal Sensor, Hydraulic Steering Control Valve, Steering Control Module, Pressure Switch.*

The considered Landing Gear System consists of a total of 128 components thus its complete description is beyond the scope of this paper. However, in the phases of the proposed process, which are described in the following sections, the complete system configuration has been considered for valuating and evaluating its reliability performances.

4.2 Objectives Definition

The reliability analysis of the Landing Gear System is an important task in the aircraft design as this system represents a critical aircraft component. The main objective is to obtain a dependable System able to guarantee the stability of the aircraft during its landing/take-off. In these phases, several critical events may happen, as an example, the following can be identified: (i) collapse of the landing gear due to a structural failure; (ii) failure in extending and locking of the main gears; (iii) inadvertent extension of the landing gear during the flight above 250 kts; (iv) unexpected braking during takeoff after the critical speed; (v) total loss of braking during the landing phase; (vi) steering angle excessive during takeoff/landing phase; (vii) missed warning for failure of the landing gear extension. Therefore, a qualitative analysis of the System reliability requires accurately identifying the weaknesses of the system, its critical components, how they can break down, the way in which faults propagate and the impacts of each failure on the entire system.

Moving from qualitative to quantitative reliability evaluations, the *Reliability Function R(t)* of the System should be evaluated along with main reliability indices as the *Mean Time To Failure* (MTTF) (see Section 2). By combining qualitative and quantitative results a clear picture of the reliability performances of the system can be obtained and alternative design solutions to effectively address the design and maintenance of the system can be compared.

4.3 System Modeling

In this phase the structure and behavior of the System described in Section 4.1 are modeled through SysML diagrams (see Section 3). In particular, the system is decomposed according to the following four-stage hierarchy: *system*, *subsystems*, *equipment*, and *components*.

In Fig. 1, the *Block Definition Diagram* (BDD) of the Landing Gear System is reported. The diagram represents the System parts and their cardinality along with their in/out port interfaces. The internal structure of the System is described by its *Internal Block Diagram* (IBD) which shows the component subsystems, their connections and interaction paths (see Fig. 2).

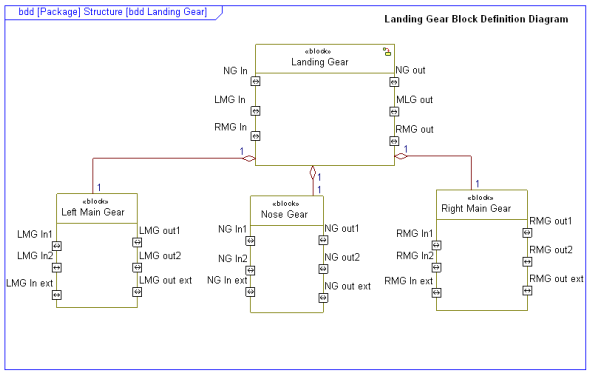


Fig. 1 The BDD of the Landing Gear System.

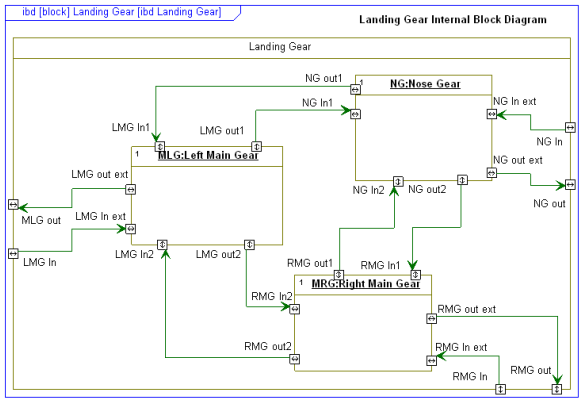


Fig. 2 The IBD of the Landing Gear System.

By applying the well-known *zooming-in* mechanisms, the internal structure of each *subsystem* highlighted in Fig. 2 is represented in details by an IBD which shows its constituting *equipment*; in turn, each equipment is further specified in terms of its *components* by a related IDB diagram. As an example, the IBD of the *Nose Gear* subsystem is shown in Fig. 3; whereas the IBD of the *Nose Wheel Steering* equipment (which belongs to the Nose Gear subsystem) is shown in Fig. 4.

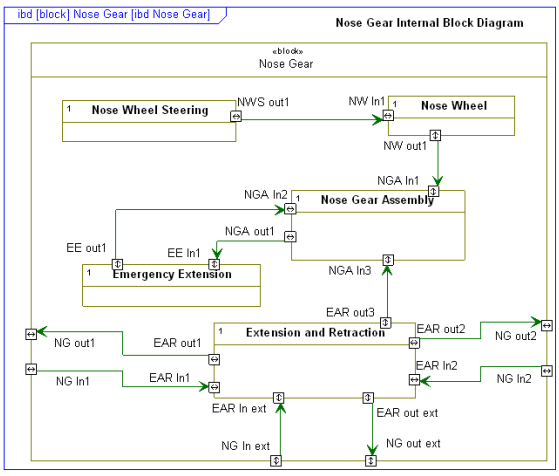


Fig. 3 The IBD of the Nose Gear subsystem.

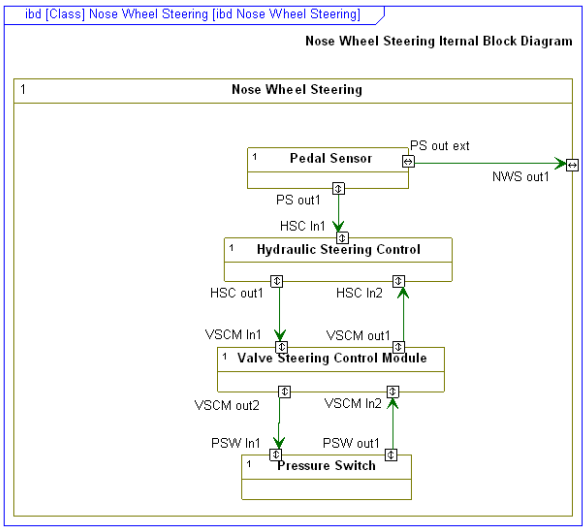


Fig. 4 The IBD of the Nose Wheel Steering equipment.

Beside the System structure, the behavior of the System is also specified following the same layered approach but in a *bottom-up* fashion: from the *component* level to the *equipment*, *subsystem*, and *system* level.

Specifically, the behavior of each *component* is specified as a set of Tasks represented by SysML Activity Diagrams (see Section 3). As an example, the specification of the behavior of the *Pressure Switch*, which is a component of the Nose Wheel Steering equipment (see Fig. 4), is presented.

Table II summarizes the *Tasks* of the *Pressure Switch* component as well as their pre and post-conditions and the execution schedule associated with them according to the UML/SysML specifications (see Section 3).

Table II. – Tasks of the Pressure Switch component.

Task	Pre conditions	Post conditions	Execution Schedule
Measure and Commutation	Pressure value acquired	(Possible) component state transition	Triggered
Failure Generation	Component is working	(Possible) fault generation	Periodical
Failure Evaluation	Fault generated	(Possible) component failure	Triggered
Failure Propagation	Component failure	(Possible) failure propagation	Triggered

As for each component, beside the Tasks which define the specific component behavior (*Measure and Commutation* in this case), a set of other *Tasks* which model the generation, propagation and management of component

failures are introduced (*Failure Generation*, *Failure Evaluation*, *Failure Propagation*). When the model is executed, these *Tasks* are crucial for simulating the onset of component failures, their propagation and their impacts on the entire system.

The SysML Activity Diagrams of two *Tasks* of the Pressure Switch component are reported in Fig. 5 and 6.

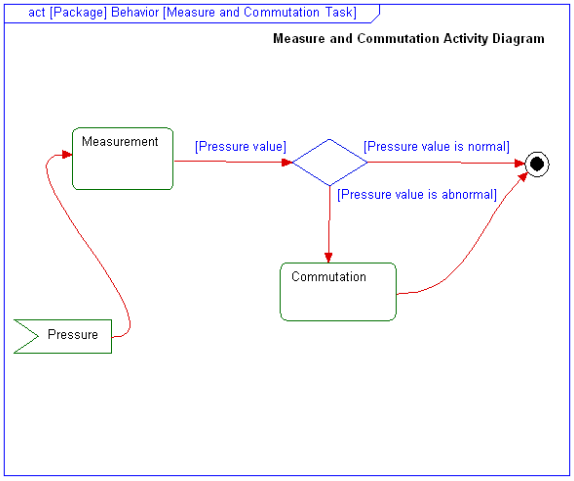


Fig. 5 The Measure and Commutation Task of the Pressure Switch Component.

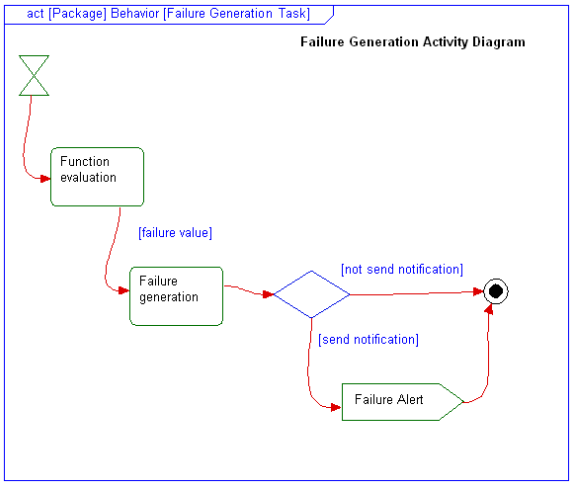


Fig. 6 The Failure Generation Task of the Pressure Switch Component.

The *Measure and Commutation* Task (see Fig. 5) defines the specific behavior of the component. When the component receives the pressure signal evaluates the pressure value and decides to switch its internal state if the measured level of pressure exceeds a fixed threshold.

The *Failure Generation* Task, which models a fault generation process according to a function specifically defined for this component, is reported in Fig. 6. This Task is activated with a given time period; upon activation, the failure function, which depends on the component *Failure Rate*, is evaluated and possibly a component fault generated.

According to the adopted layered approach, the behavior of an *equipment* is given by the composition of the behaviors of its *components*; the behavior of a *subsystem* derives from that of its constituting *equipment* and so on until the *system* level is reached. After the structure and behavior of the System under consideration has been specified through the described layered approach, the Modeling phase can be concluded. It is worth noting that, as the process is iterative, the resulting models can be refined during consecutive process iterations.

4.4 System Simulation

In the *Simulation* phase the SysML based model of the System is simulated in the *Mathworks Simulink* environment.

The *System Simulation* is a semi-automatic process that starts by exporting the SysML diagrams defined by the *IBM Rational Rhapsody* tool for loading them in the *Simulink* environment. As an example, Fig. 7 shows a view of the *Simulink Model* of the *Landing Gear System* as derived from the SysML diagrams reported in Fig. 2 and 3; whereas, Fig. 8 shows a *Simulation Model* for the *Measure and Commutation* Task of the *Pressure Switch* component which has been generated from the Activity Diagram of Fig. 5.

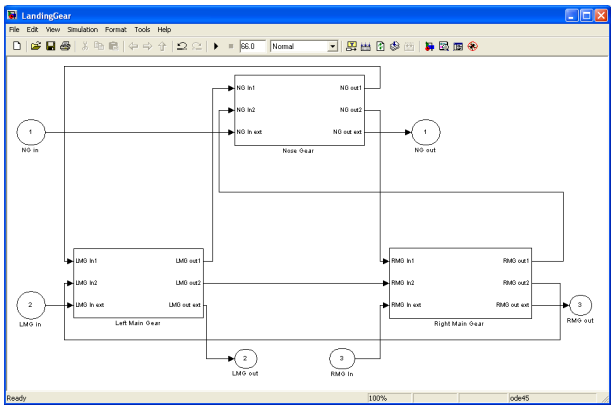


Fig. 7 The Simulation Model of the Landing Gear System.

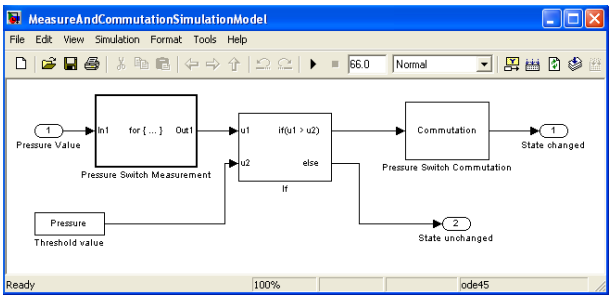


Fig. 8 The Simulation Model of the Measure and Commutation Task of the Pressure Switch component.

After obtaining the *Simulation Model*, the setting of the simulation parameters are performed according to the analysis objectives, then simulation are executed and data gathered for the analysis phase.

4.5 Results Assessment

Several simulations have been executed for analyzing the behavior of the System in different scenarios and evaluating its reliability functions and indicators as R(t) and MTTF (see Section 2). Usually, an analytic definition of these reliability characteristics is very difficult due to the complexity of the System in terms both of its structure and behavior. The scalability and flexibility of the proposed modeling and simulation approach, instead, allowed the evaluation of the main reliability functions and indices in an inductive fashion as well as the observation of macro-level *phenomena* which are hardly captured by classical analytical/deductive models. As an example, in Fig. 9 the *Reliability function* of the System obtained by analyzing the simulation data of two different and extreme operative scenarios are reported. Starting from the calculated *Reliability functions*, two *MTTF* values equals to 20 and 100 hours has been estimated for the solid and dashed R(t) respectively.

Finally, the analysis of the simulation data provided useful indications which allowed obtaining a more descriptive and predictive reliability system model and suggested some design choices which could improve the system reliability indicators.

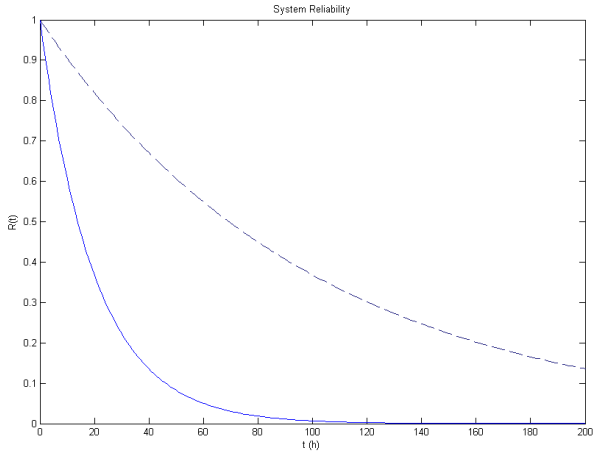


Fig. 9 Reliability Function of the Landing Gear System.

5 Conclusions and future work

The paper has proposed a model-based approach for the Reliability Analysis of systems which aims to combine in a unified framework the benefits of popular OMG modeling languages (UML, SysML) with the wide adopted Mathworks simulation and analysis environment (Matlab, Simulink). This result allows leveraging the strengths both of powerful visual languages, suitable to flexibly model the architectural and behavioral aspects of complex, dynamics, and heterogeneous systems, and of mature and popular tools, suitable for the simulation and analysis of multi-domain systems.

The exploitation of the proposed approach in a case study, concerning the reliability analysis of a Landing Gear System of a military transonic trainer aircraft, has allowed appreciating its flexibility and scalability in complex system modeling as well as its effectiveness in valuating and evaluating through simulation the system reliability performances.

Future research efforts will be geared to the definition of a full-fledged modeling and simulation framework for RAMS analysis and its extensive experimentation in the analysis of mission-critical systems in different application domains. Finally, another research line will concern the definition, starting from the system models which are produced during the proposed process, of a distributed monitoring and control

system able to guarantee the RAMS performances of the system during its whole life-cycle.

References

[1] Benveniste A., Caspi P., Edwards S.A., Halbwachs N., Le Guernic P., and De Simone R., "The synchronous languages 12 years later", *Proceedings of the IEEE*, Vol. 91, No. 1, 2003, pp. 64-83.

[2] D'Ambrogio A., Iazeolla G., Pieroni A. and Gianni D., "A Model Transformation approach for the development of HLA-based distributed simulation systems," *International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, Noordwikerhout, The Netherlands, July 29–31, 2011.

[3] Dodson B., Nolan D., *Practical Reliability Engineering*, John Wiley & Sons Ltd., 2002.

[4] Hause M., "The sysML Modeling Language," *Fifth European Systems Engineering Conference*, Edinburgh, UK, September 18-20, 2006.

[5] IBM Rational Rhapsody Web Site, *http://www-01.ibm.com/software/awdtools/ rhapsody/*

[6] IEC 61508, *Functional safety of electrical/electronic/programmable electronic safety-related systems*, Parts 1-7, 2010.

[7] Laprie J.C., *Dependability: Basic Concepts and Terminology*, Springer-Verlag, 1992.

[8] Mathworks Simulink Web Site: *http://www.mathworks.com/products/simulink/*

[9] Moir I., Seabridge A., *Aircraft Systems: Mechanical, Electrical and Avionics Subsystems Integration*, 3rd ed., Wiley, 2008.

[10] Molesini A., Omicini A., Ricci A. and Denti E., "Zooming multi-agent systems," *6th International Workshop on Agent-Oriented Software Engineering*, Utrecht, The Netherlands, July 25-29, 2005.

[11] Nicolescu, G., Mosterman, P. J., *Model Based Design for Embedded Systems (Computational Analysis, Synthesis and Design of Dynamic Systems)*, CRC Press, 2009.

[12] O'Connor P., *Reliability Engineering Handbook (Quality and Reliability)*, CRC Press, 1999.

[13] Schmidt D.C., "Model Driven Engineering", *IEEE Computer*, Vol. 39, No.2, 2006, pp. 25-31.

[14] Schrott G., "An experimental environment for task-level programming of robots," *Lecture Notes in Control and Information Sciences*, Spinger-Verlag, Vol. 190, 1993, pp. 196-206.

[15] REF8_2, Sindico A., Di Natale M. and Panci G., "Integrating SysML with Simulink using Open-Source model transformations," *International Conference on Simulation and Modeling Methodologies, Technologies and Applications*, Noordwikerhout, The Netherlands, July 29–31, 2011.

[16] Sommerville I., *Software Engineering*, 9th ed., Addison Wesley, 2010.

[17] SysML: Systems Modeling Language *http://www.omgsysml.org/*

[18] Vanderperren Y., Dehaene W., "From UML/SysMLto Matlab/Simulink: Current State and Future Perspective," *Conference on Design, Automation and Test in Europe*, Munich, Germany, March 6-10, 2006.

MCDL - Multi Configurable data link

P. G. Crema
TESEO S.p.A. – Italy
pgcrema@teseo.net

Keywords: *Flight Test, Fiber Optic, Sensors Data Fusion*

Abstract: *This paper presents an innovative tool in the data acquisition and flight test instrumentation. The target is to drastically improve the effectiveness of Flight Test, reducing the set up time, the volume and weight of copper cabling required and the risk of human error in arranging all connections. TESEO high performing and modular fiber-optic based Multi Configurable Data Link (MCDL) is able to acquire a large number of signals from analog and digital sources and transfer them over-fiber across the aircraft. MCDL acts close the signal sources a local hub where the signals to be acquired or monitored are conveyed transferred over a digital optical link to the MCDL concentrator which then reconstruct them faithfully in analog/digital form.*

The fiber-optic based system offers an additional advantage: over-fiber signals are immune to harsh EMI environments and inert in the presence of explosive materials, just one simple for cabling across aircraft and immune to environmental conditions.

MCDL is conceived as a modular system for enhanced flexibility and configurability with unmatched compactness.

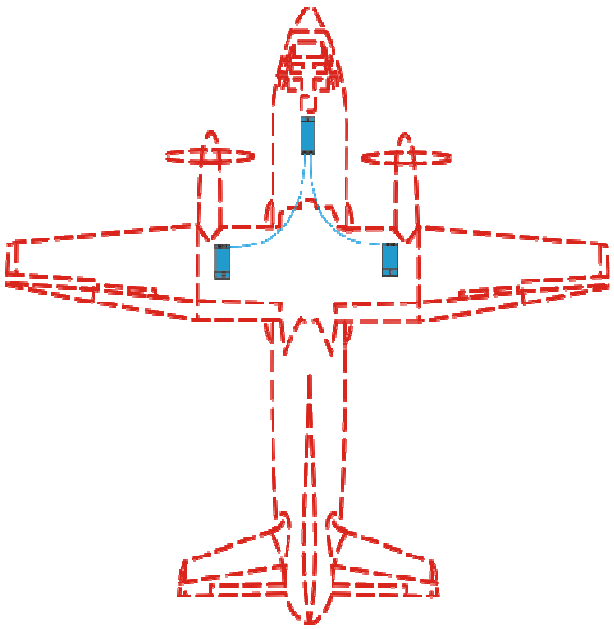


Fig. 1 – Illustrative configuration of MCDL hubs

1 Introduction

Preparing for flight test and executing it could be quite of an exercise: hundreds of sensors have to be positioned, connected and monitored such as pressure, temperature, vibration, noise level, etc. Then several data acquisition systems which are never close enough to the sensors

location – complexity and weight. Or ensuring that instructions are clear for cabling technicians – right first time. All this means costs and time required for each set-up that impact significantly budgets

The TESEO Multi Configurable Data Link (MCDL) is a double point to point (P2P) system: it means that each node (Optical Node, ON) has two optical links in order to make available several possible architectures.

In the following figures we can see four different configurations of a whole system. Figure 2 is a possible configuration in which many analog signals from various sensors are connected to two very far apart ON. A third ON is used to replicate the signals generated in a very distant point, in order to acquire them with an acquisition system .

The portable computer (Host Computer) could work like a monitor for debugging, or it could be used for parameters setting, before the system installation, or for system configuration uploading.

In Figure 3 we see the simplest configuration: a single P2P connection. One of the two ON collects analog signals coming from nearby and it transmits them via optical link to a distant ON in charge of reconstructing the signal to be acquired.

The configuration shown in Figure 4 is a hybrid configuration to transfer bi-directional signals.

Finally, the configuration in Figure 5 relates a deploying example of a single source to two remote acquisition points.

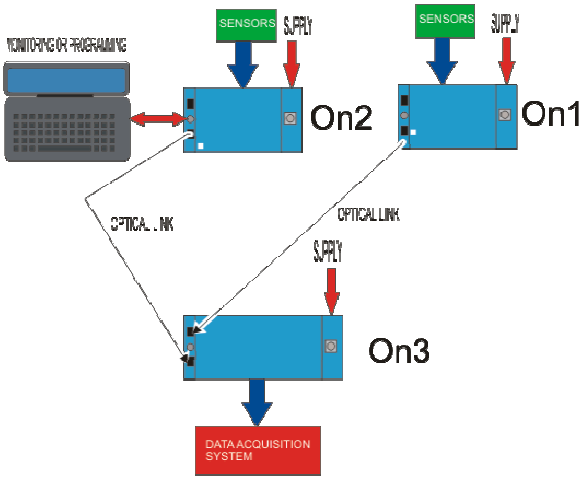


Fig. 2 – Architecture 1

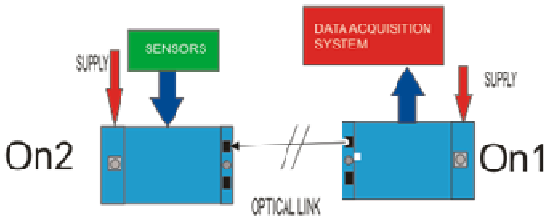


Fig. 3 – Archietcture 2

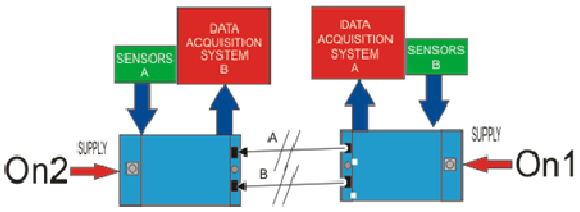


Fig. 4 – Architecture 3

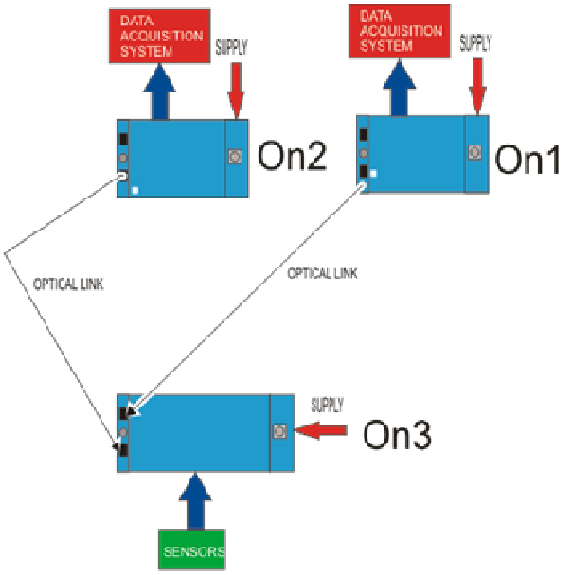


Fig. 5 – Architecture 4

2 The Optical Node

Each ON is a very small rack: 140x88x214 [H x W x L] (or 140x88x295 in the extended version) in which there are five slots (nine in the extended version).

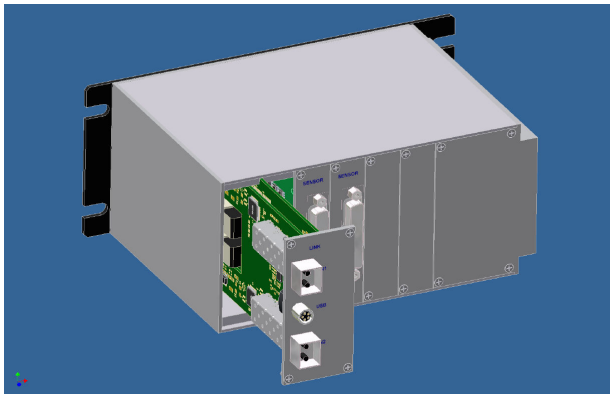


Fig. 6: Optical Node Box

In slot 0 position is inserted a link card, it is the heart of the system. In the other slots it is possible to insert the system specialization cards. In the case described in this paper we used a simple single P2P configuration as described in Figure 3. The ON that collects the signals from the field uses two cards with 16 sensor channels. The ON which reconstructs signals uses two 16 channels DAC cards.

3 The Link Card

The Link Card is the heart of the system. It manages the optical communications, the communication with any host computer through a USB (Universal Serial Bus) port and two back-plane serial buses.

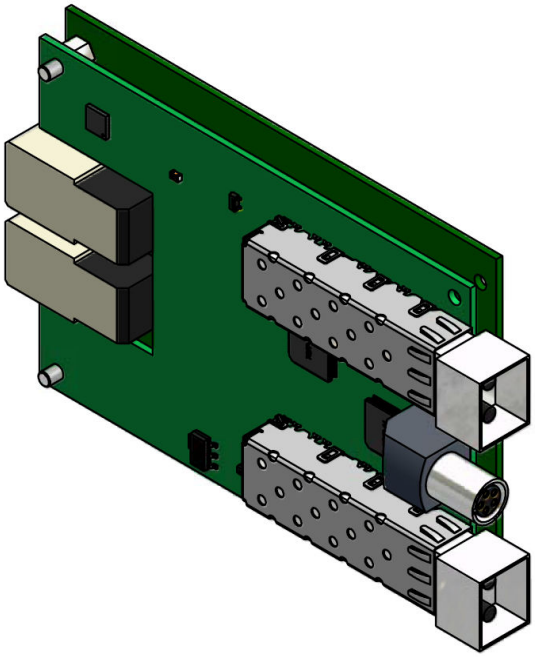


Fig. 7: Link Card

The Link Card is the back-plane bus controller and it synchronizes the whole system and all the devices connected on the back-plane bus. The devices present on the back-plane are called back-plane nodes (BN).
The Fig. 8 shows the Link Card blocks diagram.
Both the ser-deser devices perform the data parallel-to-serial, serial-to-parallel conversion, and clock extraction functions. The serial transceiver interface operates at a maximum speed of 3.125Gbps.
The optical link uses 8-bit/10-bit encoding format. The receiver section performs the serial-to-parallel conversion of the input data and extracts the data reference clock.

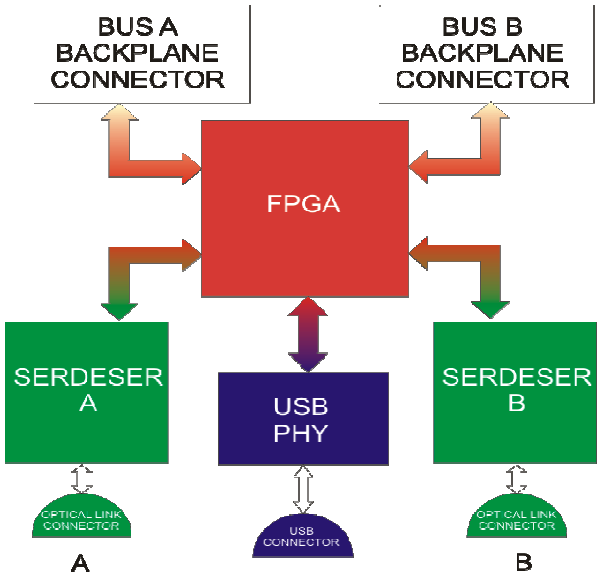


Fig. 8 – Block Diagram

The Link Card, within a MCDL system, manages the generation and the acquisition devices connected to the back-plane (BN) via a TESEO proprietary SPI based bus. It manages too a full duplex optical link in order to transmit, through a point-to-point communication, the acquired data.
The BN management is a Time-Division Multiple Access (TDMA) approach. This means that the Link Card works as a bus master clock. The tick-period, that is the time spent for data exchange between each card slot, could be set by the host computer, which communicates with the Link Card via Universal Serial Bus (USB).
The two ON are synchronized by a tick-frame sent at the beginning of each frame execution.

Each P2P optical link has an ON master and an ON slave. The ON master synchronizes cyclically the slave.
A host computer could be connected directly to the ON, but it is not part of the Time-Triggered system and it cannot manage time critical and/or safety critical functions.
The TESEO SERIAL BUS back-plane uses Serial Peripheral Interface (SPI) for the communication.
The clock (SPI_CLK) is in common to the two shift-register blocks. As we see, the data flow is done simultaneously from the LINK card to the BN and from the BN to the LINK card.

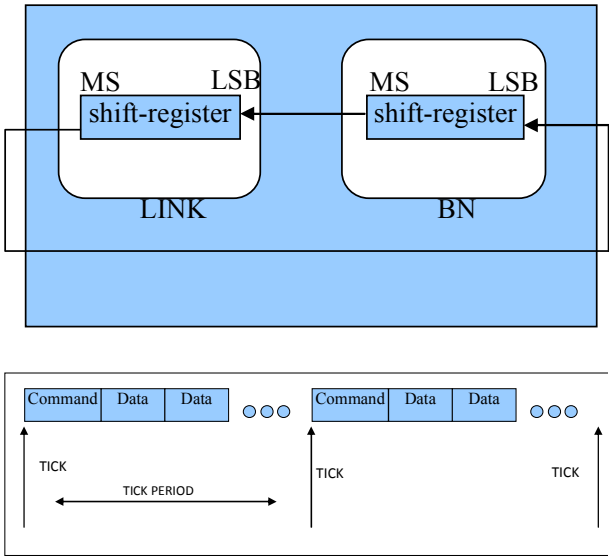


Fig. 9 – Data Flow Diagram

The messages structure is frame-based. At every tick-time a command with data is exchanged between the Link Card and the BN.
A frame is composed by a sequence of several command words followed by data words or NOP (No Operation Performed) to form a continuous data stream cyclically executed. Each command contains BN address information.

4 Time Triggered System Architecture

A time triggered (TT) system manages only one interrupt. This single interrupt is usually linked to a timer "tick" which occurs at a fixed time. In a TT design, all other inputs are polled with a regular cadency.
In a TT architecture, we can determine in advance, before the system begins executing, exactly what it will do at every moment of time in which it is running.
In a TT design, there would be no conflicts over shared resources at any time: this means that all processes can be tested completely in isolation and that changing tasks (or adding additional tasks to a system) does not have unforeseen side effects.
The TT designs tend to have a very simple architecture: this makes them easy to understand, and easy to maintain.
System testing also tends to be more straightforward with TT designs and, because of the simplicity of the architecture, maintenance costs can be substantially reduced too, then a TT architecture improves the system reliability and testability.
The key idea behind TT is to synchronize the execution of tasks on the individual nodes by sharing a single clock source between the various BN boards.

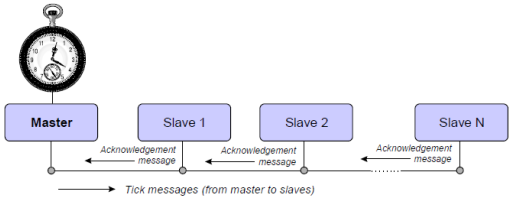


Fig. 10 – Illustrative Time Triggered Architecture

When a timer interrupt occurs, the Master node generates a message that is sent to the Slave nodes. The Slave node, which recognize the ID, responds to the message.
Besides synchronizing the individual nodes on the network, the TT master is responsible for detecting network and node failures. This means that, at worst, the Master node will know the status of all its Slaves after one TDMA round.
With each response message, the Master node identifies which Slave should return message by embedding that particular Slave ID in the data stream. Only the Slave with this particular ID will send an response message back to the Master. The Master will then check the status of this Slave, and send the next message out with a new Slave ID.

5 The Backplane Node (BN)

In this paper we consider two simple types of BN: two identical BN responsible for supplying power to sensors and acquire the analog signal coming from them and two identical BN which reconstructs analog signals after the transmission via fiber optics in the other ON.

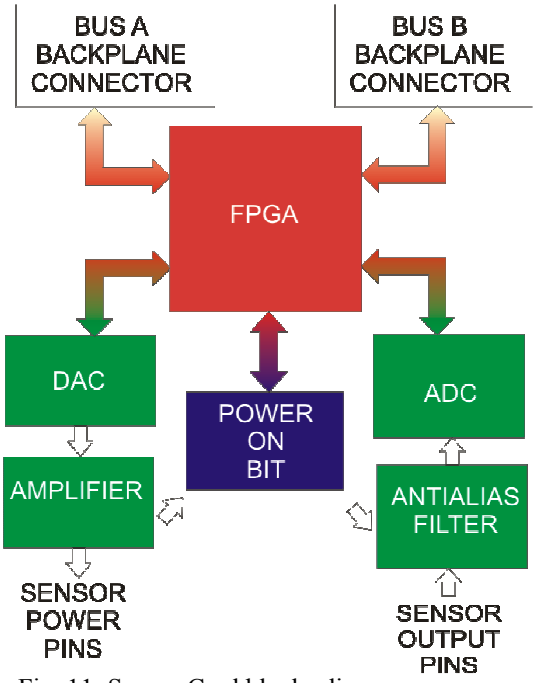


Fig. 11: Sensor Card blocks diagram

A Sensor Card has sixteen analog to digital converters to acquire analog signals from sensors and sixteen digital to analog converters to provide power to such sensors. The main characteristics are:

- Max input frequency: 20 KHz
- Max input voltage: +/- 2.5V
- ADC resolution: 16 bit
- Sensor Supply voltage: programmable 0-10V
- Sensor Supply Current: max 20 mA

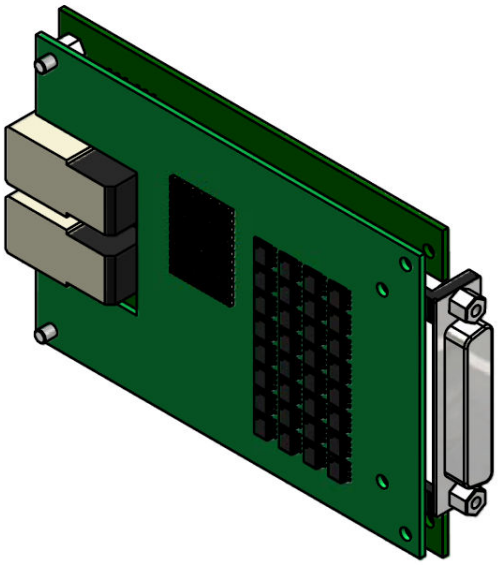


Fig. 12: Sensor Card

The DAC has sixteen digital to analog converters to generate and reconstruct the analog signals coming from the other ON. A sixteen channels ADC converter is also provided in order to make a power-on BIT circuitry.

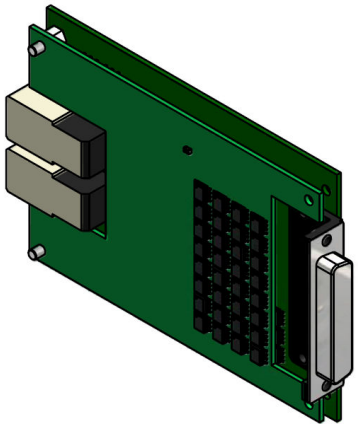


Fig. 11: DAC Card

In order to have 32 analog channels, two sensor cards and two DAC cards are used. The gain of the entire system is one and the delay between the sensor signal and the reconstructed waveform is only due to digitalization, transmission and reconstruction time.

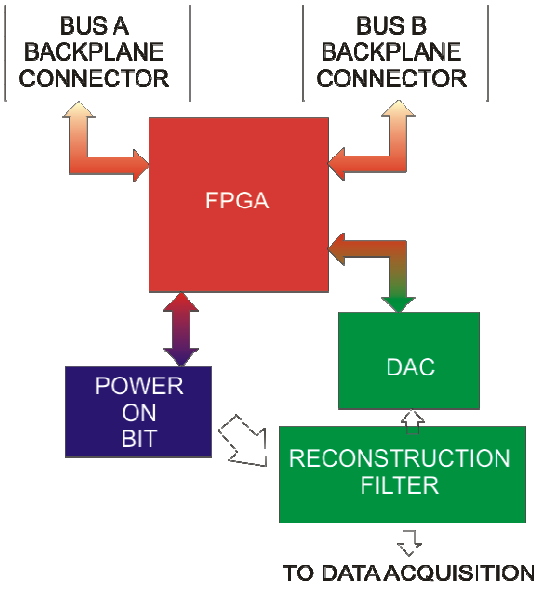


Fig. 12: DAC Card blocks diagram

6 Conclusions

Fiber Optics has been selected for 4 main core values:

- Fiber-Optic is not conductive, electrically passive, immune to EMI even in harsh environments.
- Fiber-Optic is secure even in presence of explosive materials.
- Fiber-Optic simplifies cable management and makes set-up easier.
- Fiber is intrinsically safe and does not have signal losses over distance

The final results achieved in developing MCDL Architecture has been a fiber optic enabled solution to a demand driven request. Flight Test Engineers now can leverage on:

- No Complexity: just 1 fiber crossing the aircraft means On-Board space and short wiring from sensors to MCDL hub.
- Accessibility: Optical Node (ON) always set next to sensors and acquisition side.
- Weight: Fiber is virtually weightless.
- Time to Set-Up: No complexity, reduced time to Set-Up.
- Cost: less cabling, less time, less risk of connection errors.
- Damages: Mitigated risk of connectors damages.
- Risk:, Mitigated risk of connection errors.

List of Symbols

MCDL	TESEO Multi Configurable Data Link.
ON	Optical Node.
BN	Backplane Node.
LK	Link Card.
P2P	Point to Point
USB	Universal Serial Bus
TDMA	Time-Division Multiple Access
TT	Time Triggered
SPI	Serial Peripheral Interface

References

Pont M. J., Ayavoo D., “Rapid Development of Reliable Embedded Systems”, TTE Systems, 2009.



CEAS 2011

GENERAL LECTURE
'Vision for research: Netherlands, Spain and Italy perspectives'

CHANGES IN THE EU AERONAUTICAL MIX - INDUSTRY AND SERVICES - AFTER TEN YEARS INTO THE XXI CENTURY. A REWIEW OF THE MAJOR EVENTS THAT HAVE AFFECTED COLLABORATION. THE STATE OF THE SITUATION - A VIEW FROM SPAIN

Leandro B. Fernandez Sainz, Ph.D.
Professor at CEPADE, Board Member COIAE /AIAE, CEAS Trustee

The present paper intends to review the major events and factors that have affected the Aeronautical Sector in the past ten years, then tries to foresee the future scenario for manpower education and training that would be required throughout the whole Aeronautical Mix, and in particular, in the case of Spain.
More than ten years of the present century have already passed and in that time a major restructuring in the Aeronautical Industry of Europe has taken place: the creation of the pan-European company EADS out of the most important French, German, British and Spanish aeronautical firms at that time.
During this time new aircrafts have been put into the markets that have changed the commercial scenario in the world aeronautical market. Besides two elements of the aircraft manufacturing industry, Engine and Equipment – Electronics on-board and ground equipment - there have also been mergers and restructurings.
Because of the innovative characteristics of those new aircraft, the rest of the related activities and infrastructures that make up the Aeronautical Mode of transport, i.e., the Aeronautical Mix: Operators, Airports and Navigation Authorities, have also been obliged to change quickly.
It has also been necessary to implement new Safety and Security Regulations.
The finances to support the whole Sector are under review and it has been necessary to make a new reconfiguration of the airlines Map.
Furthermore, a new world economic scenario has appeared due to new innovations in technology, and increased manpower and economical growth in the Emerging Countries has caused a shift in the localisation of the most dense operations scenario from the North Atlantic towards the Pacific Rim.
Forecasts of the future number of passengers and recent atmospheric emission data and events have brought about new concerns about the impact of the increase of the Aeronautical Transport Mode. The impact of future atmospheric regulations will eventually bring about new aircraft configurations, leading to the redesign of engines models and the search for new bio-combustibles, and new financial stress will be placed on the other elements of the Aeronautical Mix.
New recent EU Directives- Bologna- on the University Curriculum, together with new Safety and Security Regulations, have also made new changes regarding future employment.

1. Summary

The present paper deals with the Aerospace industry in Spain, which consists of Industrial

activity – Production – and Services. The focus of this paper is that looked for by the young professional – student – when attempting to orient himself on the different options and possibilities available to him to carry out his professional career, rather than repeating the group of well-known statistics which can be found in the references indicated.
Teaching activities, in their varying ranges, are not dealt with here as they are presently under revision due to the setting up of the different Directives affecting them which will bring about structural changes to former and new qualifications. However I would like to point out here that new colleges and departments dedicated to the teaching of basic Aeronautical Engineering are being created and it is foreseeable that the availability of new graduates and administrators will be considerably in excess of real needs.
The *production* activity comprises the Spanish civil aeronautics industry of fixed-wing aircraft directed at the transport of passengers and cargo (EADS); military aeronautics (EADS), helicopters (Eurocopter), engines (ITP), onboard and ground electronic equipment (INDRA), as well as maintenance (IBERIA) etc.
At the present time the *Services* industry is practically all carried out by AENA, which currently provides the services of Airports and Air Navigation with security, quality, efficiency and respect for the environment. AENA is a group of companies dedicated to airport management and the provision of air navigation services.
By way of Aena Aeropuertos S.A. (100% owned by Aena), they manage 47 airports and 2 heliports in Spain and participate directly and indirectly in the management of another 28 airports worldwide. It is the largest airport operator in the world based on number of passengers – around 200 million.
By way of the public entity, Aena offers air navigation services. Aena is the fourth largest provider of air navigation services in Europe and has played an active and important role in all European Union projects related to setting up the *One Sky* since even before it was known by that name.

2. Spanish Institutions

Outlined below, on a national level, is the Spanish institutional environment relating to the functioning and development of air navigation:

- Asociación Española de Compañías Aéreas:

The Spanish Association of Airline Companies (AECA) was founded in 1988 as an association of private airline companies. As well as forming a link between the private airline industry and the public administration, it also defends the common interests of Spanish air transport. AECA also forms a platform for carrying out activities for increasing efficiency, economy and productivity in the management of air transport services in Spain. More information: www.aecaweb.com

- Asociación de Líneas Aéreas (ALA)

The Association of Airline Companies (ALA) was founded in 1952 with the name of FLA (Federación de Líneas Aéreas) and defends the interest of the associated airline companies. The main matters dealt with by this group are economic-financial (taxes and prices), legal-labour as well as handling practices. ALA is made up of 56 passenger transport companies, both schedule and non-schedule, that operate in Spain. More information: http://www.alaspain.com

- Asociación de Compañías Españolas de Transporte Aéreo (ACETA)

The Association of Spanish Air Transport Companies (ACETA) is a private non-profit body whose members include Air Europa, Air Nostrum, Binter Canarias, Gestair, Iberia, Naysa, Orbest and Vueling and covers all sections of the air transport business: scheduled traffic, cargo, executive aviation, regional and charter. It was founded in 2004 and its aims are to defend and strengthen the interests of its associate companies, contribute to the drawing up and development of standards and decisions which affect air transport or promote quality standards which satisfy its clients' expectations. More information: www.aceta.com

- Asociaciones de Operadores de Compañías Locales

The Airline companies that operate in each airport group together in the Association of

Local Operating Companies (AOC). At the same time they also form part of the Spanish national AOC. The aims of this group are the following:

- Improve coordination between the different companies and the airport authorities (AENA).
- Study and jointly resolve problems related to security, air navigation, infrastructures, traffic, etc.

- Comisión Interministerial Defensa-Transportes (CIDETRA)

After the creation in 1977 of the Ministry of Transport, Tourism and Communications, responsibility for civil aeronautics was transferred to it from the Ministry of Defence who had had responsibility up until then. In 1979, the Inter-ministerial Defence-Transport Commission (CIDETRA) was created to permit and ensure coordination between civil and military aeronautics taking into account the competencies of both Ministries. Currently, matters under study are sent to the Commission by the Ministries of Fomento and Defensa.

- Comisión Interministerial de Política Aérea Internacional (CIPAI)

The Inter-ministerial Commission for International Air Policy (CIPAI) was created in 1946 and is made up of representatives of several ministries and presided by a representative from the Ministry of Foreign Affairs. The Commission has competency for everything referring to agreements or treaties related to policy and international air traffic.

- Dirección General de Aviación Civil (DGAC)

The Directorate General for Civil Aviation (DGAC) is the organ by which the Ministry of Fomento exercises the management and planning of civil aeronautical policy and the development of the administrative functions deriving from their being the civil aeronautics authority.

The DGAC offer their services in several fields: general and sports aviation, air agreements and relations with international organisations, user service, research, infrastructures and inspections during aircraft manufacture. More information: www.mfom.es/aviacioncivil

- Subcomité de Circulación Aérea (SCA)

The Committee for Operative Coordination was formed in 1986 in order to establish a suitable forum for communication between the civil aviation administration and users of the air navigation system. At the time of its creation it was divided into five sub-committees, amongst which is the Sub-Committee for Air Circulation. Its mission is to coordinate all the operational requirements identified and justified by its users. AENA, the DGAC, The National Institute for Meteorology and Spanish airline companies participate in this sub-committee.

3. European and International Organisations

The following are European and International organisations related to air navigation:

- Consejo Internacional de Aeropuertos (ACI)

More than 130 countries, through their airport authorities, participate in the Airports Council International (ACI). It covers more than 450 airports. This association was started in 1970 and has its headquarters in Geneva. Its aims are:

- To develop on a worldwide scale, a common airport policy for the member airports and present this before the most relevant world organisations.
- To promote the interchange of information on technical, financial, commercial, marketing and environmental matters. More information: www.airports.org

- Organización de Proveedores de Servicios de Navegación Aérea (CANSO)

Aena is a founding member of the Organisation of Air Navigation Service Providers, to which 16 air navigation service providers belong. This Association has its headquarters in Geneva and was created in 1997. The main aim of CANSO is the defence of aeronautical service supplying entities, based on the principles of safety, efficiency and effectiveness, as well as providing support to the sector's industry. More information: www.canso.org

- Conferencia Europea de Aviación Civil (CEAC)

The European Civil Aviation Conference (ECAC) was founded in 1955 by the European

Commission with the active support of the OAI Secretariat. This inter-governmental organisation, with headquarters in Paris, is made up of 36 European countries. Spain has been a member since its founding. Its objectives are:

- The promotion of orderly and safe development of civil aviation in Europe.
- The harmonization of air transport policies in economic and technical matters.
- The adoption of common measures for controlling airspace.

More information: www.ecac-ceac.org

- Agencia Espacial Europea(ESA)

The main aim of the European Space Agency, created in 1975 and of which Spain is a founding member, is the promotion of collaboration in space for exclusively pacific purposes between its member states. Its activities centre on the development of scientific missions and the practical use of space, from meteorology to telecommunications, as well as in the observation field. More information: http://www.esa.int

- Eurocontrol(Organización Europea para la Seguridad de la Navegación Aérea)

Eurocontrol, created in 1960, is the most important European air navigation organisation. It is made up of 27 member states and has its headquarters in Brussels (Belgium). Spain became a full member in 1997. The aims of Eurocontrol are:

- The coordination and regulation of European airspace.
- The joint exploitation of air traffic control services.
- The detection of rates for aids to air navigation
- Personnel training and the development of programmes, such as EATCHIP (for the integration of air traffic control systems) or CFMU (programme for centralising flow control activities).

More information: http://www.eurocontrol.be

- Asociación Internacional del Transporte Aéreo (IATA)

The International Air Transport Association (IATA) was created in 1985 and is the main international group in defence of the interest of airline companies. It is made up of 189 of these, and its aims are to:

- Promote air transport and commerce.
- Encourage cooperation between companies in the industry and analyse associated problems.
- Collaborate with OACI and other international organisations.

They also elaborate studies and traffic forecasts, as well as the IATA manuals.
More information: www.iata.org

- Autoridades Conjuntas de Aviación

The JAA (Joint Aviation Authority) is the most recently created European body. It is associated to CEAC and its main aims are:

- Establishing technical requirements allowing for common acceptation of aeronautical material by its members.
- Decreeing common air navigation codes and standards relating to the licensing of personnel.

More information: www.jaa.nl

- Organización de Aviación Civil Internacional (OACI)

The International Civil Aviation Organisation is the specialised United Nations Organisation for civil aviation. Created in 1944 with the signing of the Chicago Convention, it has its headquarters in Montreal (Canada). 185 states belong to it. The OACI sets up international principles, standards and agreements for guaranteeing the safety, effectiveness and regularity of air transport. Spain has belonged to this organisation since 1951 and with no interruptions, to its Board, which acts as an executing organ between assemblies. More information: www.icao.int

- European Union (EU)

The EU was founded in 1957 with the Treaty of Rome. It is presently made up of 27 states, with headquarters in Brussels. Spain joined the EU in 1986. The most important EU activities in which AENA participates are the Galileo and EGNOS programmes. Its aims, with relation to air navigation, are:

- Freedom of access to the aeronautical market and tariffs.
- Setting up trans-European transportation networks and their safety policies.

- Promoting research and development. More information: www.ce.europa.eu

4. Community of Madrid

The Community of Madrid, Spain, also has numerous institutions, academic and research centres related to aeronautics and space. The following are some of the most representative:

- Instituto Nacional de Técnica Aeroespacial Esteban Terradas (INTA) (www.inta.es).
- Centro de Astrobiología (www.cab.inta.es).
- Aeropuertos Españoles y Navegación Aérea (AENA) (www.aena.es).
- Universidad Politécnica de Madrid (www.upm.es).
- Escuela Técnica Superior de Ingeniería Aeronáutica y del Espacio, which today jointly substitutes the former Escuela Superior de Ingenieros Aeronáuticos, (ETSIA) (www.aero.upm.es) and the Escuela Universitaria de Ingeniería Técnica Aeronáutica, (www.euita.upm.es/).
- Escuela Técnica Superior de Ingenieros Industriales, (ETSII) (www.etsii.upm.es).
- Escuela Técnica Superior de Ingenieros de Telecomunicación, (ETSIT) (www.etsit.upm.es).
- CSIC, Consejo Superior de Investigaciones Científicas (www.csic.es).
- CIEMAT, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (www.ciemat.es).
- TEDAE, Asociación Española de Tecnologías de Defensa, Aeronáutica y Espacio was formed in Madrid on 26th February 2009 and is made up of Spanish technological industries with a presence in Defence, Aeronautics and Space. This non-profit organisation was created in order to represent and promote its associated members at both national and international level. The ultimate aim of FEDAE is the representation and promotion of its associate's interests within the ambit of the Association. The industries represented by TEDAE employ more than 40,000 people - with a turnover of €8,000 million - of whom 65% are dedicated to exports and 12% to innovative R&D. This Association jointly substitutes in some way ATECMA (Asociación Española de Constructores de Material Aeroespacial) (www.atecma.org), PROESPACIO, (Asociación Española de Empresas del Sector Espacial, (www.proespacio.org), and AFARMADE (Asociación Española de Fabricantes de Material para la Defensa).
- Centro de Astronomía Espacial Europea (ESAC - European Space Astronomy Centre), European Space Agency (www.esa.int/SPECIALS/ESAC/index.html).
- Complejo de Comunicaciones del Espacio Profundo de Madrid (MDSCC -Madrid Deep Space Communication Complex) (www.mdscc.org).
- El Centro de Control de Satélites de la Unión Europea de Torrejón de Ardoz (European Union Satellite Centre - EUSC). (www.eusc.europa.eu/).

4.1 The aerospace industry in Madrid in figures

The aerospace industry in the Community of Madrid represents 3.3% of European aerospace business, 1% on a worldwide level and 60% of the industry's national turnover and employment.

- 54% of the aerospace companies, members of the former ATECMA, are located in the Community of Madrid.
- The aerospace industry in the Community of Madrid had a turnover of €3,209 million in 2006.
- The aerospace industry in the Community of Madrid gives employment to 17,204 workers.
- Madrid employs 72% of the qualified aeronautical engineers in Spain.

5. The Sector Context and background (EU, O.J., 2009/C 175/09)

5.1 Finally, the perspectives for future employment are being pointed out.

In 2007 analysts predicted that air traffic would more than double within the next twenty years at an average growth rate of 6% per annum (5 billion passengers by 2025). In order to meet the expected increase in traffic, forecasts for orders for new aircraft (90-plus seating capacity) over the next twenty years are

optimistic and range between 22,600 (source Airbus) and 23,600 (source Boeing) aircraft. Regional air traffic is rising by 8% per annum. Orders reached a peak in 2007 for regional jets (RJ) and turboprop aircraft (for which orders doubled). Given the background (rising cost of fuel and the financial crisis), the success of turboprops is expected to continue with a probable transfer from the RJ to the turboprop market. Nevertheless, the growth of the RJ market should continue since demand for this category is strong and is expected to encroach on the Boeing and Airbus segment, with new ranges of aircraft, such as the Bombardier C Series and new entrants such as Sukhoi and the Chinese Avic.

With regard to short-haul air traffic, Europe should take timely steps to develop an R&D programme for aircraft of this type. This has now become urgent due to impending fundamental industrial changes concerning single-aisle planes with 100-plus seating capacity.

The growing liberalisation of air transport, the explosion of demand in emerging countries (Asia-Pacific and the Middle East) and the financial recovery of airlines in 2007, should be able to sustain the process.

New entrants, such as the recently merged Avic 1 + 2 (China), Sukhoi (Russia), Bombardier (Canada) and Embraer (Brazil), will probably be operational in 2015-2020. Europe will not necessarily win the price war in this sector but, on the other hand, it can hold its own by maintaining its technological edge through innovation.

Let us point out now that the aeronautics industry is of fundamental importance to European industry as a whole, due to its weight in terms of production, exports, employment and investment in R&D. It is also the driving force for a number of industries (subcontractors and downstream sectors such as aircraft maintenance) and for entire regions. No less important, it is a standard bearer for European added value and proves that joint efforts enable Europe to compete against global competitors, such as the United States and new emergent countries.

On the other hand the impact of the current international crisis cannot yet be predicted accurately. At least in the short term, the drop in oil prices may be beneficial for airlines. However, irrespective of possible developments in air transport and despite current growth, the economic and social impact of ongoing and future reorganisation in this sector in Europe is real and there is a serious risk of the further decline of the European air transport sector. This process could entail major risks such as the loss of key skills and Europe's world leadership due to its inability to make the necessary investment in developing new key employment and technologies, as well as the disappearance of a substantial number of European subcontractors from the supply chain and massive job losses.

6. Main goals and challenges for the EU aeronautics industry

The key issues for the sector are maintaining competitiveness and employment for keeping its world position and improving its public image. It is difficult for new operators to enter the aeronautics sector, and becoming a leading player has become impossible. There are only two manufacturers of 100-seater plus aircraft left on the world market for the time being: Airbus and Boeing. The technology, skills and infrastructure that are being permanently eroded or lost are extremely difficult to recreate, except in the case of transference.

Indeed, the next ten years should see the end of the Boeing-Airbus duopoly in this strategic sector, which is expected to account for approximately 65% of the 29,400 new aircraft to be built by 2027 (about 19,160 aircraft - Source: Boeing Forecast 2008-2027), but only 40% of value, a sign of growing competition and strong downward pressure on the price of this type of aircraft.

New entrants, such as the recently merged Avic 1 + 2 (China), Sukhoi (Russia), Bombardier (Canada) and Embraer (Brazil), will probably be operational in 2015-2020. Europe will not necessarily win the price war in this sector but, on the other hand, it can hold its own by

maintaining its technological edge through innovation.
Europe must therefore ensure that countries with a recognised aeronautics tradition are able to:

- maintain and build on their skills, especially by focusing on high technologies, and participate in the construction of a European network of tiers-1 and subcontractors with the capacity to give effective support to principals such as Airbus-CASA, Saab, ATR etc.;
- play a greater role in developing partnerships with the world of research (universities, other higher education institutions, government laboratories, etc.) in the field of fundamental research.

Europe cannot ignore the strong links between military and civilian research and it is necessary to increase coordination between the European Commission and the European Defence Agency (EDA) in order to promote the development of new dual-use technologies to be implemented in both the military and the civilian segments of the aeronautics industry.
The clusters in the Communities of Madrid, Andalusia and Basque country are interesting cases. Since 2000, when the A380 was launched, those regions have been implementing their own plans (Action Plan for the Development of Regional Enterprises) to support SME-SMI subcontracting aeronautics companies. This initiative produced interesting results and is to be continued in order to help SME-SMI subcontractors adapt to the Airbus POWER8 plan.
The new measure will provide case-by-case support to help businesses form groupings, increase their technological capacity, enter new markets, pool resources for purchasing raw materials, etc.
The globalisation of aeronautics has a strong impact on salaries and jobs. To counter this, research and training measures must be strengthened and supported in order to tap job creation potential. One option could be human resource planning.
Human resource planning should serve to anticipate future change. It should allow workers to cope better with future uncertainty, to develop long-term projects, to give meaning to their work, and to develop their career and investment in their company, while taking account of their needs and aspirations. It should enable businesses to adapt to change and competition.
Human resource planning must be part of a proper career and skills framework with very long-term objectives (30 years). Its goal should be to define training and skills development needs that it would be desirable to implement in the medium term in the supply of initial and lifelong training from the perspective of the industry as well as that of trainers and workers' representatives. It could be part of a debate on the employment pool.

Survival of the Fittest: 90 Years of Dutch Aeronautics RDT&E in a Globalising Environment

Fred Abbink
Senior Advisor to the General Director
The Netherlands National Aerospace Laboratory – NLR
Amsterdam, The Netherlands

Abstract

In 1919 Royal Dutch Airlines – KLM, Fokker Aircraft Industry and The National Aerospace Laboratory – NLR, were founded. In the following 90 years, in three main time periods, many political and technological developments shaped KLM, Fokker and NLR into today's form and structure. These three time periods are 1919 to 1945, the Cold War period to 1989, and the post Cold War period to the present day.

The main developments in these time periods were:

- *the consolidation of the multiple national aircraft industries, through National Champions and Joint Ventures between National Champions into European aerospace industries*
- *the consolidation of the multiple national airlines, through national flag carriers, airline alliances and mergers into global airlines*
- *the consolidation of Europe and the development of European organizations and agencies*
- *the development of European and international cooperation in aeronautical research*

Against these global and European developments, the lecture will address the development and implementation of new technology to support the Netherlands Aeronautical Industry, and civil and military operators in the changing environments of the periods between the First and Second World Wars, during the Cold War and during the post Cold War.

It will address the NLR RDT&E support of:

- *the Dutch Military Forces in the field of threat analysis, self protection, training and maintenance, as well as helicopter-ship qualification*
- *Fokker in the field of aerodynamic design, structural testing and flight testing of the Fokker F.27, F.28, Fokker 50 and Fokker 100, and the support in composites design and qualification in the period after the Fokker bankruptcy in 1996*
- *Schiphol Airport and ATC The Netherlands in order to increase the airport*

capacity, within the (legal) boundaries of safety, noise and emissions.

Furthermore it will address how NLR increases its RDT&E effectivity and efficiency by increased participation in the EU Framework Programmes, AGARD/RTO, GARTEUR, ACARE and the Association of European Research Establishments in Aeronautics, and in the field of wind tunnels, by jointly building, operating and exploiting large wind tunnels like the Large Low-speed German-Netherlands Windtunnel (DNW-LLF) and the European Transonic Windtunnel (ETW).

CEAS 2011

ATM Automation

3rd CEAS Air&Space Conference
21st AIDAA Congress

UAS Air Traffic Insertion Starts Now (Real-time simulation of UAS in ATC)

Henk. H. Hesselink
AT-One, The Netherlands
henk.hesselink@at-one.aero

Dirk-Roger Schmitt
AT-One, Germany
dirk.roger.schmitt@at-one.aero

Keywords: *UAS, UAV, Air Traffic Insertion, Unmanned Aircraft System*

Abstract

Civil Unmanned Aircraft Systems (UAS) are presently deployed in segregated airspace; passage though controlled airspace is taking place only through segregated corridors. With the increased use and the growing size of Unmanned Aircraft (UA), the need for insertion in non-segregated airspace increases, with first steps being taken in environments with air traffic control services in normal density traffic situations (en-route and not too busy TMA's). Already, civil UAS are flying in segregated airspace to carry out maritime surveillance missions and their insertion in ATC can be expected to be requested soon.

The European Commission (EC), European Space Agency (ESA), and the European Defence Agency (EDA) have established a European Framework Cooperation (EFC) in which UAS air traffic insertion is identified as major topic to be addressed.

We will describe the results of two experiments with real-time, man-in-the-loop air traffic control simulations to support UAS air traffic insertion.

1 Introduction

Several studies investigate UAS air traffic insertion, mostly addressing Detect And Avoid (DAA), safety of the operations (related to the aircraft, other airspace users, and population on the ground), architectures for data link communication, and definition of standards for certification. However, little effort is currently dedicated to performing actual flight trials and preparatory simulations for actually achieving the ultimate goal: air traffic insertion.

The major work so far in air traffic insertion is the European Civil UAV Roadmap defined by the European Commission funded UAVNet consortium [1] . A further roadmap was defined by a consortium called Air4All [2] The document defines six consecutive steps until full integration is achieved in step 6, where civil type certified UAS fly Instrument Flight Rules (IFR) and Visual Flight Rules (VFR) across national borders routinely in controlled and uncontrolled airspace (airspace classes A, B, C, D, E, F, and G).

In this paper, we present the set up and results of simulations for UAS in order to prepare for full flights in the near future. Actual flights cannot be carried out before the full set

of routine and emergency procedures has been evaluated in a simulated environment.

In order to evaluate and validate routes, procedures, and emergency situations, we have set up simulations in a real-time man-in-the-loop Air Traffic Control (ATC) environment, including UAS, which were piloted from a realistic Remote Pilot Station (RPS). In two projects, USICO (Unmanned Aerial Vehicle Safety Issues for Civil Operations, in 2007) and SINUE (Satellites enabling the integration in Non-segregated airspace of UAS in Europe, simulations have been organized, where real air traffic controllers participated to experiments for the introduction of UAS in non-segregated airspace.

For the USICO, a dense traffic sample of the Frankfurt Flight Information Region (FIR) was chosen. For SINUE, we have chosen to set up a radar simulation facility which has been configured for running the scenarios around the Canary Islands in Spain. For this, the Spanish airspace was set up and a representative traffic sample with flights from and to Gran Canarias was implemented.

2 Background – earlier work

2.1 Air4All

In June 2008, the Air4All [2] consortium, comprising of European aviation defense companies and major UAS industry partners, presented a roadmap and implementation plan defining the way to the routine use of UAS within European airspace.

On the understanding that all challenges will not practically be completely solved in one step, the route to insertion of UAS was divided into a number of increasingly challenging steps based around the different classes of airspace and the relative difficulty of operating in them.

Step 1: Fly experimental UAS within national borders in segregated airspace (regular, at short timescale) – Unpopulated range

Step 1a: Fly experimental UAS within national borders in segregated airspace (regular, at short timescale) - overflown sparse population

Step 2: Fly an experimental UAS as IFR traffic within national borders in controlled, non segregated airspace (airspace classes A, B and C)

Step 3: Fly a national type certified state UAS as IFR traffic within national borders, routinely in controlled airspace (airspace classes A, B and C)

Step 4: Fly a civil type certified UAS as IFR traffic within national borders, routinely in controlled airspace (airspace classes A, B and C)

Step 5: Fly a civil or state UAS as IFR traffic across national borders, routinely in controlled airspace (airspace classes A, B and C)

Step 5a: Fly a civil or state UAS as IFR traffic across national borders; routinely in controlled airspace (airspace classes A, B, C, D and E)

Step 6: Fly a state UAS as IFR and VFR traffic across national borders, routinely in controlled and uncontrolled airspace (airspace classes A, B, C, D, E, F and G)

Step 6a: Fly a civil type certified UAS as IFR and VFR traffic across national borders, routinely in controlled and uncontrolled airspace (airspace classes A, B, C, D, E, F and G). Summarized Air4All identifies steps from

- experimental UA to type certified aircraft
- state aircraft vs. civil use of UAS
- national airspace use vs. cross border operations

as depicted in figure 1.

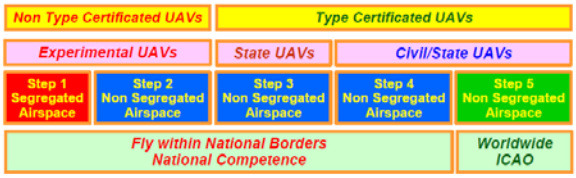


Figure 1 Air4All steps

The Air4All consortium notices that step 1 has already partly been achieved and in progress and the focus for research must therefore be laid at step 2. According to Air4All, step 2 has a number of immature challenges but is considered achievable in a reasonably short timeframe. Step 2 consists of a number of issues, linked to technical challenges and is

deemed feasible to be implemented for demonstration by the end of 2011.

2.2 E4U

A representative group of European research establishments has set up a prioritization of actions to identify the major challenges that need to be considered for air traffic insertion of UAS.

The E4U study describes the topics:

- UAS Air Traffic Insertion
- Single European Sky
- UAS Missions
- Platform and payload
- Radio Bandwidth Allocation
- Ground Control Station
- Propulsion systems

The study is ongoing until the end of 2011; early results show that, as expected, DAA and security are important topics to solve in the near future. Furthermore, LOS/BLOS infrastructures and safe recovery systems are identified as important elements for research. In BLOS operations, automated take-off and landing are required.

2.3 Working groups

Standardisation committees EUROCAE WG-73 and RTCA SC-203 are identifying necessary elements of the architecture of communication systems that will support the operation of UAS in non-segregated airspace. Their role is not to endorse or promote a particular architecture, and consequently there is no consensus on what the architecture should look like.

The European Commission (EC), European Space Agency (ESA), and the European Defence Agency (EDA) have established a European Framework Cooperation (EFC) in which UAS air traffic insertion is identified as major topic to be addressed.

3 Introducing UAS flight in non-segregated airspace

Both Air4All and E4U studies identify the flight of an experimental UAS within national borders in controlled, non-segregated airspace as first steps in UAS air traffic insertion. We assume that the current state-of-the-art in UAS insertion is that the aircraft already fly in segregated airspace without major difficulties. Although not all issues are solved completely, we will consider in this paper the next step: fly an experimental UAS in IFR traffic. Topics that need to be addressed for this are.

Separation. The UA will now fly in airspace together with other aircraft. It is expected that increased separation criteria and dedicated ATC is necessary to separate the UA from other aircraft flying in the same sector.

Collision Avoidance. To avoid other traffic, the simplest step is to fly IFR under control of ATC, where ATC will provide separation between aircraft. Airspace considered will be low density class C en-route airspace and quiet Terminal Maneuvering Areas (TMA). The UAS flight will need to be monitored more closely by ATC, but can also be followed e.g. with a chase aircraft or a ground monitoring system. Collision avoidance also concerns obstacle avoidance and the avoidance of controlled flight into terrain.

Secure and sustainable communications for command and control. Security and integrity of the datalink is assumed to be available. The UA is expected to react ‘immediately’ to the instructions of ATC. The experimental flight must be performed at a geographical location where satellite coverage can be ensured.

Radio Bandwidth allocation. This challenge is only relevant to the real experimental flight and not for initial simulations. It must ensure that national authorities provides allowance; the World Radio Conference 2012 (WRC 2012) will decide on allocation of frequencies for UAS.

ATC Interface. The challenge concerning the ATC interface covers the following items:

- Ensure suitable data exchange with ATC
- Ensure continuous ATC interface with sufficient integrity

- Ensure ATC conformal operator interface
- ATC interface for pre-flight information needs definition

A dedicated architecture needs to be set up for communication between ATC and the UA in those situations where a satellite system is used as main communication means.

Dependable emergency recovery. A procedure must be defined and implemented in the UA to ensure safe recovery in emergency cases. An emergency route and procedures for flying towards and at the route need to be set up, just as well as a "home"-area, where the aircraft will eventually fly to in case of emergencies.

Emergency recovery is a major issue for ensuring the safety of flight.

Health monitoring / Fault detection. This challenge concerns the ability of the UA to detect faults which affect the ability to continue safe flight and to avoid collisions. This feature is scripted in the simulation, so that we assure that possible faults are evaluated in our experiments.

UAS pilot / commander training. The UAS pilot must be able to control the vehicle in normal and in degraded operating states. When the UA or communication with the UA is in a degraded state, the pilot must be able to asses the ability to safely fly and land the aircraft.

Furthermore, the UAS pilot must be able to use standard radio communication procedures in communicating with ATC and must have the capability to react 'immediately' to the instructions of ATC.

Demonstration preparation. All studies identify the need for early demonstration and a thorough preparation of demonstration flights. We propose to carry out simulations in a realistic ATC environment in order to explain and train air traffic controllers in handing the aircraft.

4 Research question

To bring air traffic insertion of UAS further, we start with simulating the environments in which the aircraft will fly. It will be possible to set up the necessary architecture in a network of simulators, for the air traffic control station, the RPS, and the satellite communication link. In a real-time simulation environment, air traffic controllers will be able to experience without risk, the aircraft in operation in their sector and experience themselves the aircraft's characteristics, the use of emergency routes and procedures, communication with a remote pilot, and the interaction with other traffic. Our research question was to

1) Identify a suitable architecture for BLOS operations with UAS.

2) Examine the effects on ATC of UAS in their airspace.

5 The architecture set up

In most scenarios, the aircraft will fly en-route their missions in a remote area and will therefore be flying Beyond Line of Sight (BLOS). Communication between the pilot and the aircraft will have to take place through satellite communication. Just as well, communication between the ATC centre and the RPS will be relayed over satellite. Therefore, the architecture proposed must at least enable the Command and Control (C2) link between the UAS pilot and the UA and the ATC link between the UAS pilot and the ATC center, see the functional decomposition in figure 2 (from [2]) below.

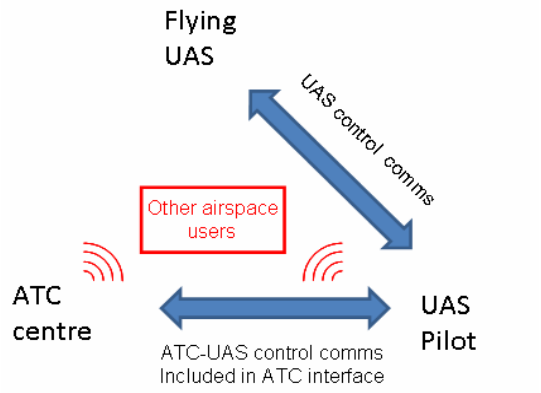


Figure 2 Functional set up

In the simulation experiment, the architecture as depicted in figure 3 was chosen.

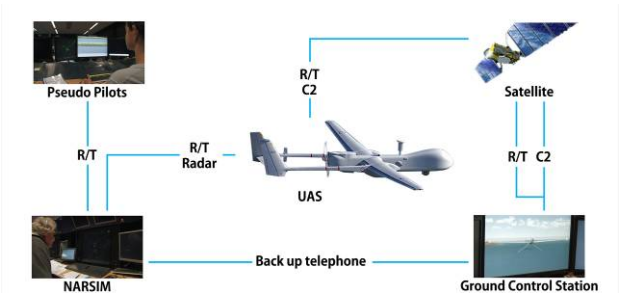


Figure 3 Functional set up

In the center of figure 3, the UA is flying a mission in controlled airspace. The UAS pilot has no line of sight with the aircraft, so command and control will be relayed over satellite. This already is a standard operating procedure for UAS that fly BLOS operations.

Specific attention has been paid to VHF R/T communication between ATC and the UAS pilot. In our set up, the aircraft will receive all R/T on the frequency and relays this signal on a dedicated channel to the satellite. This set up requires significant bandwidth hence operating costs, but our calculations do show that bandwidth does not form a limitation here. A back up for R/T communication is available through a standard telephone line.

A pseudo pilot had to control the other traffic in the simulation. In a typical simulation, depending on the intensity of the required actions, pseudo pilots are capable of dealing with 10 to 20 aircraft at the same time.

6 Emergency Situations/Procedures

The mission of the European Aviation Safety Agency (EASA) is to promote and maintain the highest common standards of safety and environmental protection for civil aviation in Europe and worldwide. EASA performed a study for an impact analysis on safety of communication for unmanned aircraft systems [5].

From this study, through a functional hazard analysis, the following list of relevant emergency situations needs to be covered during experimental simulations and flights:

- Loss of voice communications between UAV/S pilot and ATC
- Interruptions to voice communications between UAV pilot and ATC
- Intelligibility and latency of voice communications between UAV pilot and ATC
- Loss of command and control link between UAV and GCS
- Interruption of command and control link between UAV and ATC (due to system reliability or satcom coverage)
- Loss of surveillance information feed to ATC
- Interruption of surveillance information feed to ATC (due to system reliability or radar coverage)
- Loss of surveillance information to other airspace users
- Interruption of surveillance information to other airspace users (due to system reliability or coverage)

With the exception of the “loss of surveillance information”, all events were considered in the experiments to cover all emergency situations emerging from the use of UAS. As the C2 and ATC signals are relayed through different channels on board the aircraft, the “loss off”- emergency situations can occur for either one of them or for the both simultaneously.

For the design of emergency procedures, three aspects need to be considered: a home area, an emergency route, cleared from the other airspace routes, and a procedure to fly from the current location (where the emergency occurs) to the emergency route.

6.1 Home area

The home area is a base, where the UA will fly to when an emergency occurs. The aircraft will land there or perform a maneuver which will destroy the aircraft without risk of casualties. For each flight and for each experiment, the home area needs to be defined, depending on the local situation. For the two experiments mentioned in this paper, USICO [4]

and SINUE [3], two distinct procedures were defined.

In the SINUE set up, a home area above sea was defined, where the aircraft would fly a circular pattern and be climbing in order to try to re-establish communication with a land based station that would be within line of sight. In USICO, an emergency airport was identified.

6.2 Emergency route

An emergency route must be designed that is fully separated from other air traffic routes, so that the UA can follow a path separated from all other traffic. For every flight with a UAS, the route must be carefully evaluated in order to check whether it is easily and safely reachable from the mission area.

Figure 4 shows the route used in the Spanish experiments, where one route was sufficient for all experiments performed. This route was designed in cooperation with air traffic controllers and was designed such that several entry points were defined towards which the aircraft would fly in case of an emergency. The points were chosen so that the aircraft would never fly over inhabited areas in case of an emergency and was vertically separated from other crossing air routes. The figure below shows the emergency route in red. The home-area is located at the bottom in the middle of the figure.

One special situation is when the aircraft is on final approach. In this case, the UA would fly the standard missed approach procedure to avoid it flying through other aircraft on approach, see the “hook”–pattern on the bottom of the picture.

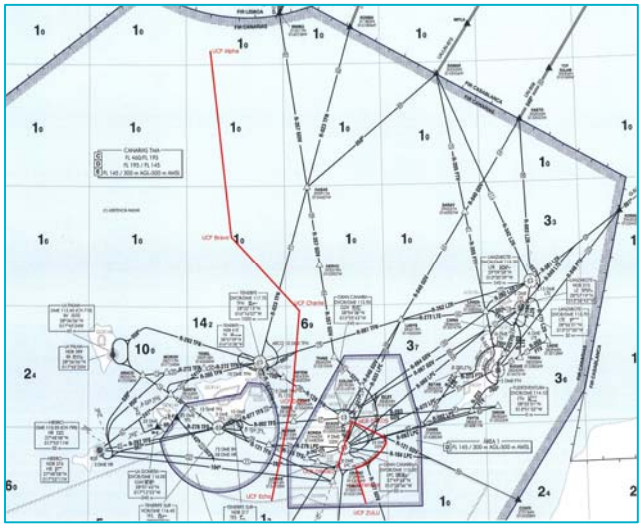


Figure 4 Design of emergency routes

For USICO, the simulated airspace is the TMA Frankfurt controlled by Frankfurt Arrival and the western sector controlled by Langen Radar. Controller working positions of the ATC center (Frankfurt Arrival and Langen Radar) are provided by the employed ATMOS. The simulated traffic in these two sectors is piloted by the pseudo pilots. The traffic in the northern and southern sector is navigating fully autonomously, i.e. it is so called dummy traffic. See figure below. For the emergency, the airport of Hahn was planned as alternate.

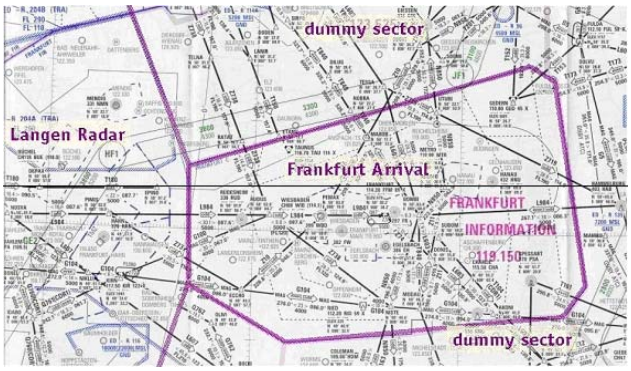


Figure 5 USICO TMA Frankfurt

6.3 Towards the emergency route

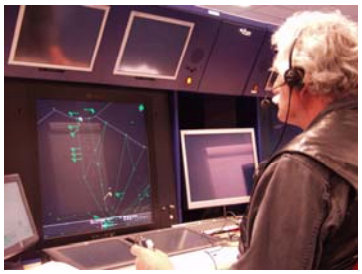
To reach the emergency route, the UAS follows a standard procedure which is known to the controller and the remote pilot. The procedure chosen in the SINUE study follows the procedure that has to be followed by other aircraft as well: the UA will abort its flight path

by turning towards the closest way point on the emergency route and maintain its current altitude for two minutes. After the two minutes the aircraft would climb or descent towards the altitude of the closest way point on the emergency route.

7 Experiments

In the simulations we carried out experiments in air traffic control simulation facilities, where real air traffic controllers participated to carry out the simulation and to evaluate the proposed concept and procedures.

We used an air traffic control simulator, which resembles the airspace where the aircraft is flying as much as possible.



Experienced air traffic controller and pseudo pilots were running the experiments, who were briefed at the beginning of the day and were given the possibility for training.

Through questionnaires directly after each run and at the end of a simulation day a directed list of questions was handled. Just as well, at the end of the day, a discussion session was held with all participants of the simulations.

8 Results

The goal of the two studies was to examine the effects on ATC of UAS in their airspace. From the simulations, questionnaires, and debriefings, we obtain results for the sessions that were held.

Separation and collision avoidance. In the simulations, we did not use different separation criteria than those currently in use. The aim was to see if current separation can be maintained, even though there is no pilot on board the aircraft. We instructed controllers to use current separation criteria for the UAS, which they were able to maintain. In initial practical trials with real aircraft, for safety, the separation between a UA and a piloted aircraft will be increased in actual air traffic insertion experiments. The exact separation will need to be decided by the regulatory authority.

Communication. Communication delay because of the satellite connection will not be an issue when high quality bands are used. For SINUE we have chosen to perform the mission around the Canary Islands, where coverage of the Hispasat satellite system can be ensured. The satellite gives a delay in voice communication of around two seconds. In the scenario (no dense traffic), this was rated acceptable by the air traffic controllers.

ATC interface. We investigated the interface with ATC with respect to these aspects. New special squawk codes are proposed:

7600: comm loss
7660: datalink loss, proceed as planned
7661: datalink loss, return home
7662: datalink loss, fly to emergency field
7700: emergency

Although controllers have indicated that they do not particularly require specific terminology or symbols for UAS guidance. Either they do not feel comfortable with more information, or they expect that more information will not help them in solving the issues at hand.

Dependable emergency recovery

In the simulation, we have defined a “home”-zone, to which the aircraft will fly following a standard route which is separated from other airspace routes. The procedure for flying towards the standard route follows common practice.

Controllers in all cases indicated they felt comfortable with the procedures defined, even where the emergency situation would occur at the “most inconvenient moment”. In our case an UA was flying without control trough an arrival stream and in another situation straight towards two low flying IFR aircraft. As long as emergency situations are defined similar to those of manned aircraft, controllers are well trained for emergency situations.

Situational awareness. It is important for all parties to have a good overview of the traffic situation and to have the same mental picture of a traffic situation. This implies that there is a

need for a good situational awareness for air traffic controllers, UA pilots, and pilots of other traffic.

The air traffic controller will need to know that he is dealing with an unmanned airccraft in an instance. Already at any existing ATCo's (Air Traffic Controller) display, the aircraft type is indicated in the aircraft label. The aircraft types need to be known to controllers. Other options to give more recognition to the UAS can be:

- A special convention for the use of callsigns can be arranged for UAS.
- A dedicated UAS symbol can be used to depict the aircraft.
- The UA label at the ATCo's display can be made more explicit, e.g. by use of a special colour.

The unmanned aircraft must be easily visible by eye for controllers in the control tower, which implies that the colour coding of the aircraft bodies and liveries must be carefully designed.

During the introduction phase of UAS into air traffic control, other pilots will need to be aware of the flying objects around them. Air traffic control must play a role in this, through informing pilots that an unmanned aircraft is flying ahead of them. This can be quite easily accommodated through informing other traffic over the R/T that a special aircraft is flying in their vicinity. This is already common practice, e.g. with hot air balloons and glider traffic.

Just like for air traffic controllers in the tower, the aircraft must be easily visually recognizable for other pilots.

Emergency procedures. In the experiments, all emergency situations as identified by EASA were tested.

We observed that controllers were not always fully aware of the aircraft's behavior at the moment that it was flying towards the emergency route. This was partly due to the fact that they used it for the first time.

From the discussions with controllers, it is suggested to define and discuss the emergency routes in advance of any simulation or real flight trial, based on the planned flight of the UA. The altitude of the points on the route must be defined such that the aircraft will make as little as possible a vertical movement on its way towards the route. Just as well, the route must also be defined as high as possible, to increase the possibility for re-establishing communication, either through satellite or via direct line-of-sight.

The emergency route can be displayed at the controller's display, either at all times or only at request of the controller.

Back up phone. One specific back up element was introduced in the SINUE experiment. The air traffic controller was able to contact the UAS pilot directly by phone. This possibility is especially interesting in case of R/T failure between ATCo and UAS.

The procedure to initiate a phone call was implemented as follows. In case of a communication failure involving the UA, the UAS pilot initiates a phone call between him and the ATCo. The controller has to "answer the phone". The connection was a fixed phone connection that is open as long as the connection was active. Voice on the telephone was relayed over the headset of the ATCo. The ATCo used his microphone to talk and did not have the use the Press To Talk (PTT) button. This mean the UAS pilot was able hear all communications from the ATCo to the other pilots and that all other traffic was able to hear the instructions from the ATCo to the UAS pilot. The UAS pilot and the pilots of the other traffic were not able to hear each other.

In any future concept, it needs to be identified in what particular situations the telephone connection should be activated. Possibly, all communication between the pilot and the ATCo can be performed over telephone, removing the necessity for R/T installations on board the UA.

Workload. One important issue is workload. We measured workload with some of the USICO experiments After adequate training of controllers during the warm-up runs the controllers felt that the simulation set-up represented quite well a real working environment of ATC controllers. Hence there is evidence that the obtained results are representative for a real ATC/ATM environment.

The trials started with the baseline of the normal traffic. Later on, 2 UAVs have been added into the airspace. It could be shown, that the workload increased, which was due to the "new" behavior of the type of aircraft. Figure 6 shows an example of the results of NASA TLX workload determination of the controllers for those different trials

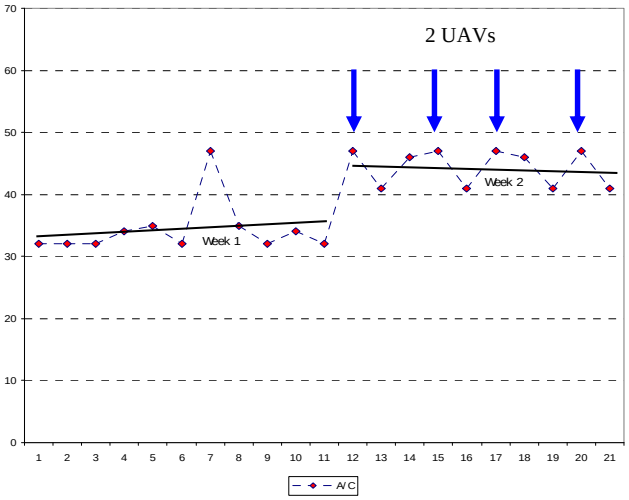


Figure 6 NASA TLX workload vs. trials without (left) and with UAVs (rigth) included in airspace [6]

Recommendations

We asked the controllers whether they would assign a different priority to unmanned aircraft over manned traffic. Interestingly, they considered UA traffic lower priority than commercial traffic and would treat it as small VFR traffic. This needs to be taken into consideration with assigning routes and sequences to the UA.

From the results of the simulations, the following recommendations are given:

- When operating over satellite, keep the RPS on the party line. The UA pilot and the pilots of other traffic must be able to hear each other and hear the instructions given to each of them.
- Dedicated R/T must be developed or existing R/T must be adapted to inform other pilots of the UA in their vicinity.
- The UA does not require new symbology on the ATCo's display, but the ATCo must be able to see in a glance that the aircraft on his display is unmanned. A simple indication by using a dedicated type of call sign will do.
- The ATCo does not require more information on emergency transponder codes.
- ATCo's need good training of emergency situations.
- Benefit can be taken from the fact that communication with the UA pilot can be established over a high quality land line.

9 Conclusions

We have set up a real-time man-in-the-loop simulation environment, where real air traffic controllers, a pilot at a UAS RPS, and a real pilot (for other traffic) participated to evaluate the concept. Several representative scenarios were evaluated, including emergency situations.

We have shown that integration of UAS in controlled airspace is a feasible concept; air traffic controllers indicated that control of the UAS did not differ significantly from control of other, manned, aircraft. We have also demonstrated that UAS emergency procedures can be designed equivalent to those of manned aircraft, such that the air traffic controller will understand them and is able to predict the behaviour of the UAS in several loss-of-satellite-communication situations.

References

[1] *European Civil Unmanned Air Vehicle Roadmap, 25 Nations for an Aerospace Breakthrough*, Volume 1 – 3, UAVNet, editor, 2005

[2] Air4All, *UAS Insertion into General Air Traffic, Final Study Report,* Document No. BAES-AS&FC-BDEV-GEN-RP-005564 Issue 1, June 2008

[3] SINUE, *Satellites for the Integration in Non-segregated airspace of UAS in Europe, Final Report,* Document No. SINUE_EID_00035_FINAL REPORT, January 2011

ann E, *Evaluation of ATC/ATM integra-tion concepts for civil UAV operations. in: Unmanned Aerial Vehicle Safety Issues for Civial Op-erations (USICO)*, Report D4.1, European Project funded within the GROWTH Aeronautical Part of the 5th Framework Programme of the European Com-mission, Contract G4RD-CT-2002-00635, 2004

[5] EASA, *Preliminary Impact Assessment on the Safety of Communications for Unmanned Aircraft Systems, Final Report, volume 1*, European Aviation Safety Agency, Research Project EASA.2008/8

[6] Courtesy of Reimund Küke, Rheinmetall Defence Electronics GmbH, Bremen, Coordinator USICO Project, 2004.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Modeling and validation of aircraft mathematical models for the development of an innovative Flight Management System

M. Battipede, M. Cassaro, and P.Gili
Dipartimento di Ingegneria Aeronautica e Spaziale, Politecnico di Torino, Italy

P. Gunetti
Department of Automatic Control and Systems Engineering, University of Sheffield, UK

Keywords: *aircraft simulation, clean sky, modeling, validation,*

Abstract

This paper reports the experience of the MAS_Lab (Multipurpose Aircraft Simulation Laboratory) project, developed at the Politecnico di Torino within the Clean Sky ITD Systems for Green Operations. The complete project environment is briefly described and the MAS_Lab tasks are discussed. Modeling design choices are investigated, highlighting those concerning simulator structure and automatic flight control laws; the validation process is described, focusing on the techniques used to assess model accuracy. Examples of results obtained during the validation campaign are presented, with the intention of showing how model validation and autopilot performance are considered satisfactory.

List of acronyms:

FMS flight management system; AP autopilot; FD flight director; ALT altitude hold mode; LOC localizer mode; APP approach mode; ATHR autothrottle; FPA flight path angle; HDG heading; V/S vertical speed; s_p stabilizer angle; θ_B body attitude angle; α angle of attack; F_N engine net thrust; H.F. horizontal flight.

1 General Introduction

The Management of Trajectory and Mission (MTM) aims at developing technologies to reduce emissions (CO2, NOx) and noise through the 4D trajectory optimization. Since the systems developed for trajectory and mission optimization need to be inserted into the overall economic model of the operators (the airlines), some knowledge needs to be gained on how those economical models will influence operators attitude when it comes to choosing between environmental impact and economic performance. This will enable to associate a "cost" to each generated trajectory. In this respect, the following items need to be developed, as summarized in figure (1):

- an environmental impact model to compute the noise related to the aircraft operations as well as the amount of carbon dioxide, nitrogen oxides and water vapor released in the high atmosphere;
- a set of aircraft models, sufficiently detailed for the purpose [1] [2];
- cost model to get the required economical data.

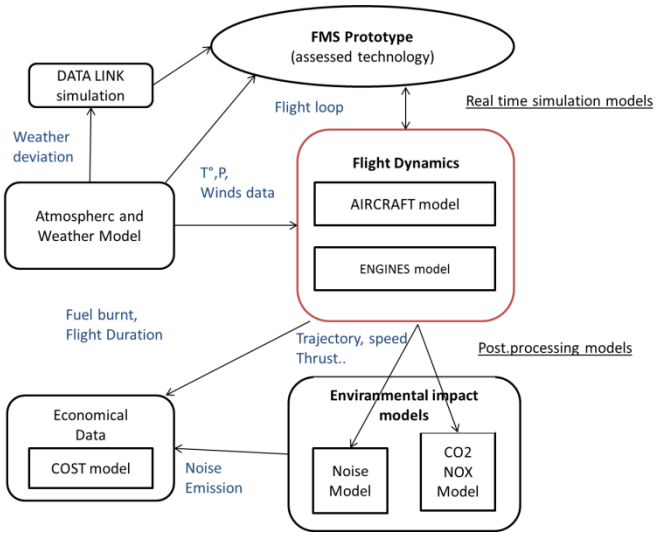


Fig. 1 - Project Environment

The MAS_Lab project is focused on the aircraft models development (red box in Fig. 1). In particular the tasks outlined in the project can be listed as follows:

- Define the simulator specifications in terms of major features, such as hardware/software, development environment, but also in terms of what the simulator must be able to model and to which extent.
- Verify and validate the model according to the definitions given by the VV&A Recommended Practice Guide of the DoD Model & Simulation Coordination Office [6].
- Implement the simulator, including the interfaces.
- Maintain and improve the simulator throughout the whole Clean Sky duration.
- Introduce the knowledge of state-of-the-art simulation technology within the consortium and promote the dissemination of the technical and scientific results achieved during the project, among the partners and the international community also in non- aeronautical contexts.

The MAS_Lab project consists, more in detail, in the design of a so called *Master Model* in which four aircraft, concerning different categories, ranging from wide-body to regional jet, are modeled. Three of the four models are referred to existing aircrafts: B747-100; ATR 42-500 and CESSNA Citation CJ3; while the fourth is a non-existent model, which can be defined by the user through a limited set of parameters.

2 *MAS_Lab Master Model* structure

The MAS_Lab *Master Model* is implemented using the Matlab\Simulink® environment. To achieve the objectives listed in the previous section a modular configuration of the simulator was used. This choice was due to the simplicity of finding and correcting errors with this method, and because it is the most suitable configuration, considering the parallel development of several aircraft models in the same simulator.

Figure 2 shows the main external loop of the simulator composed by the three blocks, as follows: Pilot; AFCS and Aircraft dynamics. The Pilot block concerns the logic interface built to allow the user, or the FMS, to act on the simulated aircraft by means of a cockpit, or direct connection. The AFCS block contains all the autopilot's control laws that can be used to control the airplane or to give visual cues to the pilot when acting as flight director. The Aircraft dynamics block contains the four different aircraft models and the blocks solving the state equations, so it is the core of the simulator. Using various enable-switches, it is possible to change between highly different aircraft without changing the model structure.

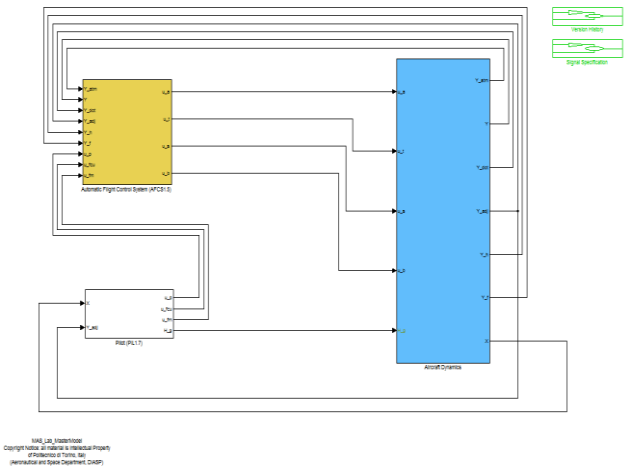


Fig. 2 - *Master Model* main external loop

2.1 Pilot

This block consists of several components in which all the external command are transformed in signals compatible with the rest of the model. More in detail it contains the following blocks:

- *Joystick*

This block enables the user to act on the aircraft control thus commanding the longitudinal, lateral, directional and thrust control inputs. Controls for flaps, stabilizer, speedbrakes, steering wheel and brakes are also present. The vector *u_p* collects all data from this block

- *Flight Control Unit*

The purpose of this device is to enable the user (pilot) to manage the autopilot manually during flight. The FCU implemented is composed by three different types of buttons: switch push buttons, on/off push buttons and triple action buttons.

Switch push buttons are used to select the reference parameter for the AP. For example, they are used to switch from Mach to Speed value and from FPA to VS.

On/off push buttons are used to activate/deactivate the autopilot and their specific functioning modes. For example the AP and ATHR buttons engages the AP and the ATHR while the ALT, LOC and APPR buttons activate the corresponding guidance modes.

Triple action buttons are used to manually control the autopilot target or to enable the FMS control of the same target. For example, by pulling and rotating the knob, the user can control speed, lateral (heading) and vertical (altitude or V/S) targets. Whereas by pushing the buttons the FMS control is enable.

2.2 Automatic Flight Control System

After a deep study on the autopilots and their mode selection logic, a unique AFCS configuration was chosen for all the simulated aircraft. This means that all the aircraft have the same autopilot, divided into longitudinal, lateral-direction autopilot and auto throttle, but obviously with different control laws. The structure of the AFCS and its validation process will be detailed in Section 4.

2.3 Aircraft dynamics

The A/C dynamics block, Fig.3, contains the four aircraft non-linear models and is thus the core of the simulator. It is able to evaluate the response of the aircraft as a function of the different command and atmospheric inputs and return the step-by-step time evolution of the system (a discrete time integrator was used to solve the equations of motion).

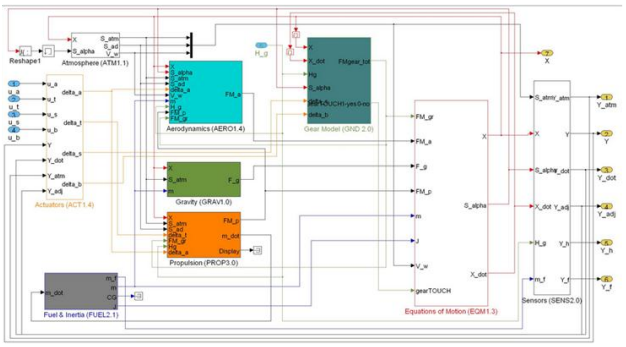


Fig. 3 - Aircraft dynamic model's structure

The block is divided into:

- *Actuators block*

The aircraft actuators (elevators, ailerons, rudders, flaps and stabilizer) are accurately modeled through the use of look-up tables that reproduce the dynamic response, the mechanical inter-connections and the blow-out limitations of each control surface.

- *Aerodynamic block*

On the basis of the control surface deflection signals, coming from the actuators block, this block is used to compute the aerodynamic forces and moments in body axes reference frame: several look-up tables are used in order to take account of various aerodynamic effects such as dynamic coupling and ground effect.

- *Gravity block*

Evaluates the gravity forces in body axes.

- *Propulsion block*

Contains the models of the aircraft engines, computing propulsive forces and moments in body axes in every flight condition.

- *Fuel & inertia block*

It simulates the mass properties (c.g. position and inertial tensor) of the aircraft as a function of the initial gross weight and the fuel consumption.

- *Gear model*

It simulates the gear reactions when in contact with the ground, and expresses the resultant forces and moments in body axes reference frame.

- *Equations of motion block*

This is the real core of the simulator. By taking into account all the external forces and moments acting on the aircraft and their mass and inertial properties, in this block the six DoF non-linear equation of motion are implemented and solved with the numerical Runge-Kutta method. The state vector x of the aircraft, composed by the linear and angular position, velocities and accelerations, are obtained. The body axes reference frame is used for this purpose.

- *Atmosphere block*

It calculates the physical properties of the atmosphere. It implements the ISA model, but can be connected to an external model if a more realistic one is necessary.

- *Sensors block*

It models the sensors of the aircraft, including radio navigational aids which are used to control the ILS landing.

3 *MAS_Lab* Verification and Validation

In order to assess the accuracy, coherence and consistency of the simulator, a significant effort was spent in performing the verification and validation (V&V) process. The V&V Plan is based on documentation published by the Defense Modeling and Simulation Office (DMSO) of the American Department of Defense (DoD). These Recommended Practice Guides (RPG) on Verification, Validation and Accreditation (VV&A) allow to establish guidelines and methodologies that can be expected to result in meaningful V&V campaigns [6]. It is to be noted that the V&V process can be completed only if the aircraft model is completed itself. As a consequence only the B747-100 model was validated at the state of the art of the simulator, as hereafter reported. Following the reference guidelines, table with 21 different tests was created and accomplished for the B747-100 validation. As shown in Fig.4, V&V tests are performed primarily (Phase 1) at a system level (e.g. A1.T3 correspond to engine static thrust test or A1.T6\7 to longitudinal and lateral control system test) and then (Phase 2 and 3) are extended to the whole model on three different levels:

- statics
- dynamics
- performance (statics and dynamics)

For example tests A1.T9 to A1.T11 correspond to trim condition test, whereas tests A1.T12 to A1.T15 correspond to the dynamic test performed on the linearized aircraft model and tests A1.T16 to A1.T21 are the performance test carried out on the complete non-linear aircraft model.

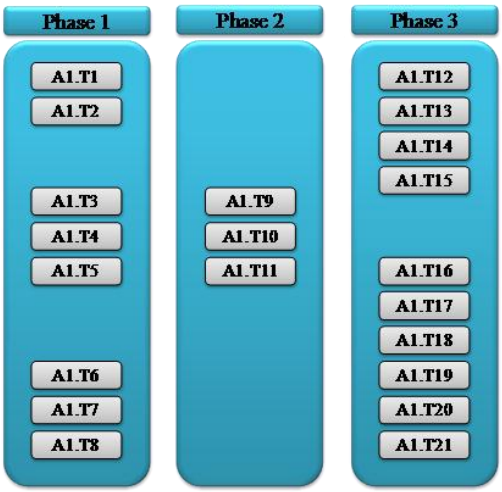


Fig. 4 - B747-100 V&V test scheduling

The level of accuracy of the results obtained is guaranteed by the high fidelity of the sources used as reference for comparison purpose (e.g. NASA certified simulator data and BOEING experimental data [7]).

3.1 Validation results

Because of the project purpose, the propulsive system is the most important component of the simulator. Since showing the results of the entire validation campaign is beyond the scope of this paper, the propulsion system will be used as an example to show the validation process.

In the first phase, the test purpose is to verify the accuracy of the system models, as aerodynamic, propulsive and control system.

First of all, each block is statically tested. Correspondences of input\output combinations with reference data are checked. For example, two different test types are performed on the propulsion system. In the first case the maximum excursions of the Thrust Lever Angle are checked, for the forward and reverse positions, for both cases maximum and idle static thrust is checked by comparing the engine block output data. In the second test, the same comparison is done for the static thrust obtained for different Mach and EPR values.

In the second phase, the whole aircraft model is tested by analyzing the equilibrium points of the model. The simulated airplane is statically trimmed at several variations of weight, c.g. positions, altitude, speed and flap positions. Computed value of s_p, θ_B, α and F_N (engine thrust) are compared to literature data. The effects of configuration changes, landing gear and speed brake positions are evaluated.

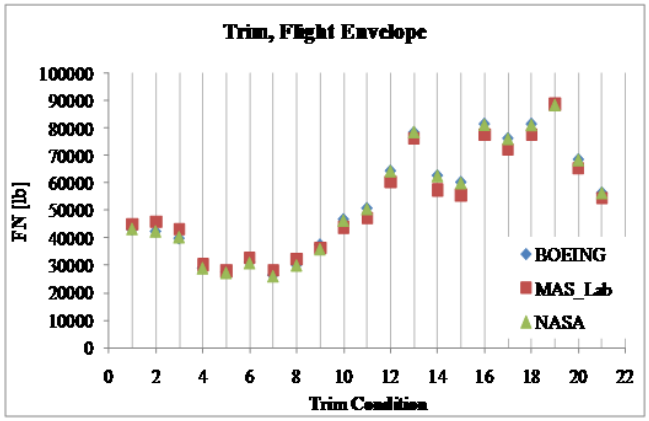


Fig. 5 - B747model's thrust validation (H.F.)

Figure 5 is an example of the test carried out in this phase and of the most relevant results. The test sequence does not follow well defined criteria, for this reason it is impossible to highlight a specific trend in the parameter's variation. However it is clear that the level of accuracy of the thrust's results follows the stringent requirements established for the simulator purpose, as reported in Table 1.

	maximum abs.error %	minimum abs. error %	average on 22 conditions
F_N	7.31	0.45	4.62%

Table. 1 - B747 model thrust percentage error (H.F.)

These results, matched with the fuel model, bring to an average error of fuel consumption of 5% as expected by the simulator requirements.

In the third phase, the model dynamic and performance are tested by means of several computational tools (e.g. trim and linearization algorithms) and non-linear simulation. In particular two different tests concerning the dynamic analysis of the aircraft model are performed.

In the first test, the aircraft modes (short period, phugoid, dutch roll and roll mode) dynamic features are calculated in terms of damping ζ and natural frequency ω_n: the aircraft is trimmed in a specific flight conditions and the equations of motion are linearized. In practice, every flight condition generates a new linear model expressed in terms of transfer functions.

In the second test, the aircraft modes are excited by an elevator input (longitudinal dynamics) or a $\beta \neq 0$ initial condition (lateral-directional dynamics). Results are obtained running a simulation of the non-linear aircraft model, recording the time history and comparing them with the literature data.

The aircraft performance tests are performed by manual or automatic piloting of the simulated aircraft. Different mission profiles are reproduced by the non-linear simulation and then recorded data compared with literature reference data. Three types of maneuvers are accomplished: take-off, climb and acceleration/deceleration.

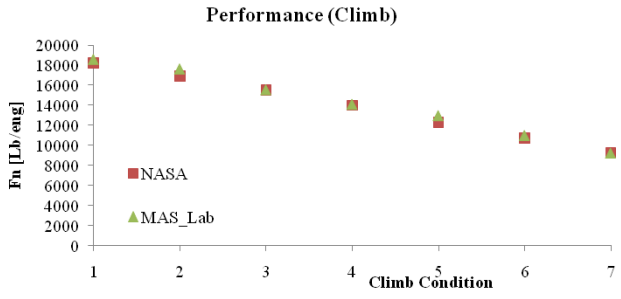


Fig. 6 - B747 model thrust validation (Climb)

As shown in Figure 6, six different climb conditions are analyzed, covering the entire flight envelope in terms of altitude (from sea level to maximum altitude) and at maximum thrust. The test sequence shows the high level of consistency of the MAS_Lab model response,

compared with literature data. For example, for the parameters reported in Table 2, it is possible to notice that not only the reference trend is followed, but also that the absolute errors are very small. Because the project goals require excellent modeling of fuel consumption and trajectory management, the accuracy and consistency of climb maneuvers are critical in assessing MAS_Lab, and the obtained results are satisfactory.

	maximum abs.error %	minimum abs. error %	average on 6 conditions
F_N	5.29	0.21	2.26%

Table. 2 - B747 model thrust percentage error(Climb)

4 Automatic Flight Control System: modeling and validation

Modern Flight Management Systems (FMS) are capable not only of automatically generating a detailed flight plan, but also of executing most of it. To do this, they are appropriately interfaced with the Automatic Flight Control System (AFCS) on-board the aircraft. As the MAS_Lab is primarily intended as a tool to test and validate an innovative FMS, the aircraft AFCS should be accurately modeled and an interface with the FMS ensured.

Because autopilots are a commercial product, it is very difficult to retrieve detailed information regarding their performance and the implemented control laws. The NASA document used throughout MAS_Lab for the Boeing 747-100 model [8] provides a rather detailed description of the autopilot system; however, this description focuses on the electromechanical aspects of the system, and its implementation in Simulink would require modeling all the electrical connections and the related mechanical actuator dynamics. Furthermore, autopilots of the 747-100 period (1960-70) provided specific control modes which are not used anymore, while missing control modes that are now commonly implemented; also, these autopilots were never meant to be controlled by a FMS.

For these reasons, it was decided that the MAS_Lab AFCS was to be based on the Boeing 747-400 AFCS rather than the Boeing 747-100; all relevant control modes were to be implemented from scratch, so as to ensure the possibility to use the AFCS both manually (pilot controlling the AFCS) and automatically (FMS controlling the AFCS). Seamless transition between the two possibilities was to be ensured, as in fact is normal on real modern aircraft, where the FMS controls most of the flight with the pilot intervening at times to change parts of the flight (for example as a response to an ATC directive).

In the following subsections, the 747 AFCS as modeled in MAS_Lab will be described, focusing first on the available control modes, then on the actual control laws, and finally on the corresponding validation process. It is to be noted that the 747 AFCS model will also be used for the other aircraft, though of course the control law gains will have to be recalculated and some control modes might not be available on a specific aircraft.

4.1 AFCS modes

The AFCS functionality can easily be decomposed into three separate parts: pitch control, roll control and auto-throttle. Several control modes are available for each, which can mostly be activated independently (although some interdependency is present).

The following pitch control modes are modeled in the MAS_Lab AFCS:

- *Manual or MAN (1)*: the autopilot keeps a fixed pitch angle provided by the pilot;
- *Altitude Hold or ALT (2):* the autopilot captures and keeps the target altitude;
- *Vertical Speed Hold or VS (3):* the autopilot keeps a pilot-defined vertical speed;
- *IAS Hold or IAS (4):* the autopilot maneuvers the aircraft on the pitch axis in order to maintain a desired airspeed (note that the throttle must be set to a fixed value);
- *Glide-Slope or GS (5):* the autopilot maneuvers the aircraft to follow the Glide-Slope signal of an ILS, executing flare and rollout once landing is obtained;

• *Take-Off/Go-Around or TOGA (6):* the autopilot keeps a pre-defined constant pitch attitude that ensures a positive rate of climb, to be used during take-off and go-around maneuvers;
• *VNAV Path (7):* this is equivalent to the Manual mode, but the commanded pitch angle is provided by the Flight Management System;
• *VNAV SPD (8):* this is equivalent to the IAS Hold mode, but the commanded airspeed is provided by the Flight Management System;
• *VNAV ALT (9):* this is equivalent to the Altitude Hold mode, but the commanded altitude is provided by the Flight Management System;

The following roll control modes are modeled in the MAS_Lab AFCS:

• *Manual or MAN (1):* the autopilot keeps a fixed roll angle provided by the pilot;
• *Heading Hold or HDG (2):* the autopilot captures and keeps the target heading;
• *Track Hold or TRK (3):* a variant of HDG, the autopilot keeps a precise constant-heading route defined by a starting position;
• *Localizer or LOC (4):* the autopilot maneuvers the aircraft to follow the Localizer signal of an ILS;
• *Take-Off/Go-Around or TOGA (5):* the autopilot keeps wings leveled during take-off and go-around maneuvers;
• *LNAV (6):* this is equivalent to the Manual mode, but the commanded roll angle is provided by the Flight Management System;

The following auto-throttle modes are modeled in the MAS_Lab AFCS:

• *Thrust Reference or THR REF (1):* in this mode, one of several pre-defined fixed throttle values is kept (e.g. IDLE, TAKE-OFF, etc.);
• *Speed Hold or SPD (2):* the auto-throttle controls thrust to maintain a desired airspeed;
• *Mach Hold or Mach (3):* the auto-throttle controls thrust to maintain a desired Mach;
• *Managed THR REF(4):* a direct throttle command is provided by the FMS;
• *Managed SPD(5):* this is equivalent to the SPD mode, but the commanded airspeed is provided by the Flight Management System;
• *Managed Mach (6):* this is equivalent to the Mach mode, but the commanded Mach is provided by the Flight Management System;

The mode selection logic has been implemented using Stateflow, in order to represent the many safeguards that are applied on the actual AFCS; for example, activation of the *IAS* pitch mode inhibits activation of the *SPD* auto-throttle mode (as the two are not compatible). Also, some modes can be armed and then activated when appropriate conditions are met; for example, the *S* and *LOC* modes are armed and then activated when the relative signal deviation is below a specified threshold.

4.2 Control laws

The main pitch control loop is shown in Figure 7; it is a traditional autopilot design which uses pitch angle and pitch rate signals to calculate the elevator displacement. The signals are appropriately filtered and gain scheduling is employed to ensure a good response throughout the flight envelope.

The nine pitch control modes are applied externally to the main loop, providing a desired pitch angle value. For each mode, appropriate sensor signals are used to compute the pitch angle required to achieve the condition. As an example, the altitude and altitude rate signals are processed to compute the pitch angle required to reach and maintain the desired altitude in the ALT mode. A roll compensation signal is added to avoid altitude loss during lateral maneuvers. Figure 8 shows how the model switches between the various control laws.

A similar strategy is used for the roll control part: the main control loop uses roll angle and roll rate signals to calculate the aileron displacement, whereas the desired roll angle value is provided by the various roll control modes. The roll control modes are usually simpler than pitch control modes; due to the slow latero-directional response of the aircraft, in fact, a proportional signal is sufficient, without the need of derivative signals.

The auto-throttle is conceived differently, as there is no inner loop and each mode directly controls the throttle needed to achieve the desired reference value. As an example, Figure 9 shows the Speed Hold control law. It uses filtered airspeed proportional and derivative

signals to calculate the throttle value; gain scheduling is applied to ensure smooth response throughout the flight envelope and specific rules are used for speed control during landing (e.g. throttle retard function).

4.3 AFCS model validation

As previously stated, autopilots are commercial products and it is difficult and costly to obtain actual performance data. For this reason, thorough validation of the dynamic response of the MAS_Lab AFCS was not feasible. Instead, the validation process was focused on checking static behavior and on a qualitative analysis of dynamic response.

A total of nine test types have been prepared in order to validate AFCS functionality. To ensure sufficient coverage of the flight envelope, each test is repeated starting with three different initial trimmed configurations:

• *condition 1*: 30000 ft, 0.72 Mach, 400000 lb gross weight, c.g. position at 32% M.A.C.

• *condition 2*: 5000 ft, 0.41 Mach, 564000 lb gross weight, c.g. position at 25% M.A.C.

• *condition 3*: 15000 ft, 0.55 Mach, 680000 lb gross weight, c.g. position at 14% M.A.C.

The validation tests for the MAS_Lab AFCS are as follows:

• *test 1*, ALT mode test: starting from the trim condition, a limited altitude change (±500 ft and ±1000 ft) is commanded, with the auto-throttle in SPD mode; ALT mode and main pitch loop dynamic response are evaluated;

• *test 2*, VS mode test: starting from the trim condition, an altitude change (±3000 ft and ±5000 ft) is commanded, with the auto-throttle in SPD mode and a required VS (±1000 ft/min and ±1500 ft/min); VS mode and main pitch loop dynamic response are evaluated;

• *test 3*, IAS mode test: starting from the trim condition, an altitude change (±3000 ft and ±5000 ft) is commanded, with the auto-throttle in THR REF mode and a required IAS (current speed ±20 kt); IAS mode and main pitch loop dynamic response are evaluated;

• *test 4*, SPD mode test: starting from the trim condition, a speed variation (±20 kt, ±40 kt and ±80 kt) is commanded, with ALT mode keeping constant altitude; SPD mode dynamic response is evaluated;

• *test 5*, Mach mode test: starting from the trim condition, a Mach variation (±0.05 and ±0.1) is commanded, with ALT mode keeping constant altitude; Mach mode dynamic response is evaluated;

• *test 6*, HDG mode test: starting from the trim condition, a heading angle change (±30 deg, ±60 deg and ±135 deg) is commanded; HDG mode and main roll loop dynamic response are evaluated;

• *test 7*, auto-transition tests: various tests are run to verify the functionality of automatic mode transitions (such as VS/IAS to ALT when target altitude is reached, SPD to THR REF when IAS is engaged);

• *test 8*, auto take-off test: the automatic take-off function is tested using different configurations (weight and c.g. position);

• *test 9*, ILS test: starting from various low-altitude trim conditions, ILS capture and track are tested; focus is placed on activation of modes and verification of approach profile (vertical – GS and lateral - LOC);

Figures 10 and 11 show a test example; these are the plots of altitude and pitch angle obtained for test 2, executed from condition 3 with an altitude change of 5000 ft and a VS of 1000 ft/min. In this case, the AFCS was judged satisfactory.

It is to be noted that AFCS dynamic response cannot be guaranteed to exactly reproduce the actual behavior of the aircraft, unless the data describing the performance of the actual AFCS is available.

5 Conclusion

Modeling and Validation process was completed with satisfactory results for the B747-100 model. Accuracy, consistency and coherence are the main features of the simulator. It is already able to fly with a Microsoft® Flight Simulator graphical interface. It is also ready to be interfaced with the innovative Flight Management System provided by a partner of the Clean Sky project. For the next future it is planned the completion of the other three

aircraft models. However their modeling seems to be faster starting from the knowledge and experience acquired in this first phase of the MAS_Lab project.

References

[1] Stevens B. and Lewis F., *Aircraft Control and Simulation*, J. Wiley & Sons, 2004.

[2] Etkin B. and Reid L., *Dynamics of flight: Stability and Control*, J. Wiley & Sons, 1996.

[3] Rauw M.O., "*Using SIMULINK for Nonlinear Analysis of Aircraft Control Systems*", Interim report, Delft University of Technology (Aerospace Engineering), Delft, 1992.

[4] Rauw, M.O., "*A SIMULINK Toolbox for Flight Dynamics and Control Analysis*", FDC 1.4, Draft version 7, May 25, 2005.

[5] Rauw, M.O., "*A SIMULINK environment for Flight Dynamics and Control analysis-application to the DHC-2 'Beaver'"*, Delft University of Technology (Aerospace Engineering), Delft, 1993.

[6] U.S. Department of Defense. "*DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)*". DoD Instruction 5000.61. May 13, 2003.

[7] Boeing Document D6-13703, "*Boeing 747 Flight Manual*", December 30, 1969.

[8] Boeing Document D6-13645, "*747 IEFCS System Description*", April 03, 1968.

[9] Boeing Document D6-301-51-425, "*747 Operations Manual"*, May 01, 1992.

[10] "*Flight control systems: practical issues in design and implementation*", Edited by Roger Pratt,Institution of Electrical Engineers, IEE Control Engineering Series, 2000.

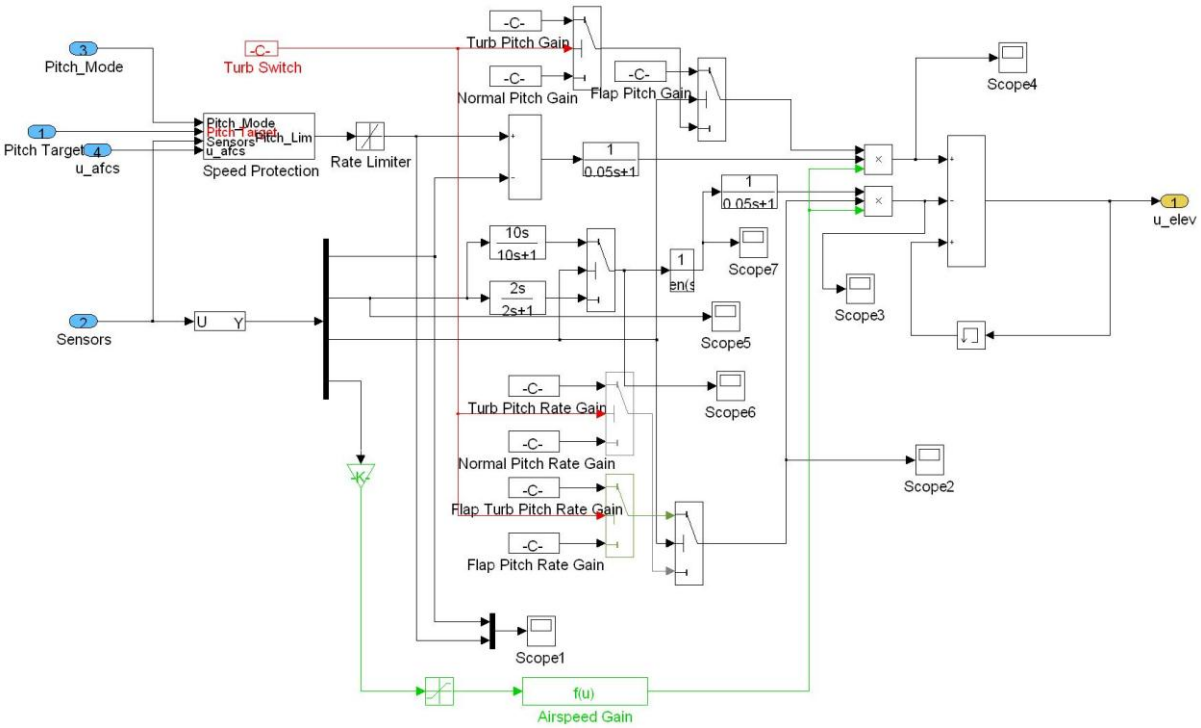


Fig. 7 - Main Pitch control loop

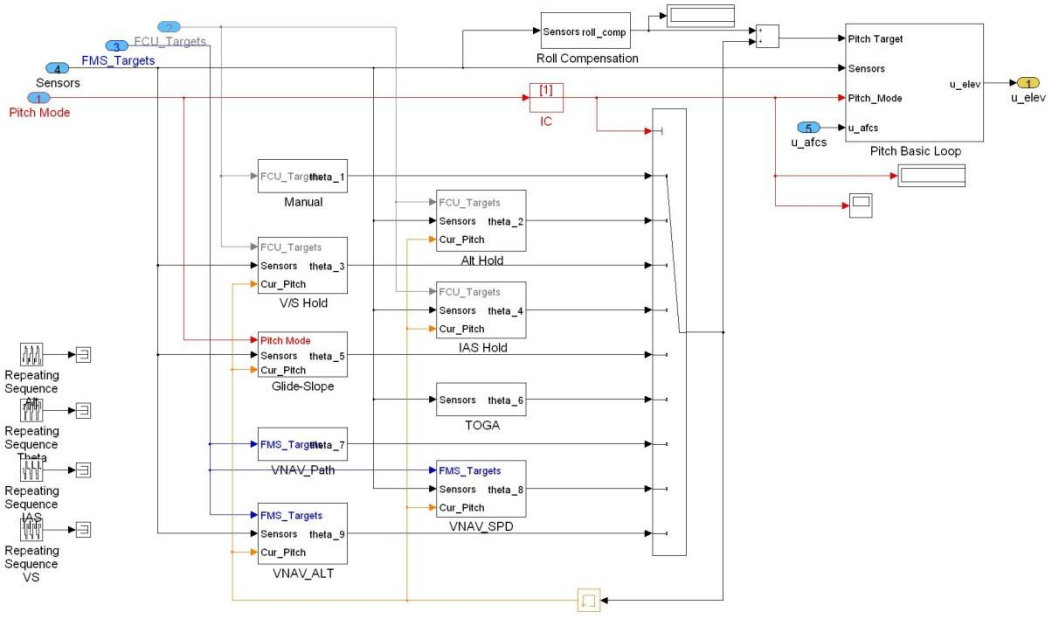


Fig. 8 - Pitch control modes

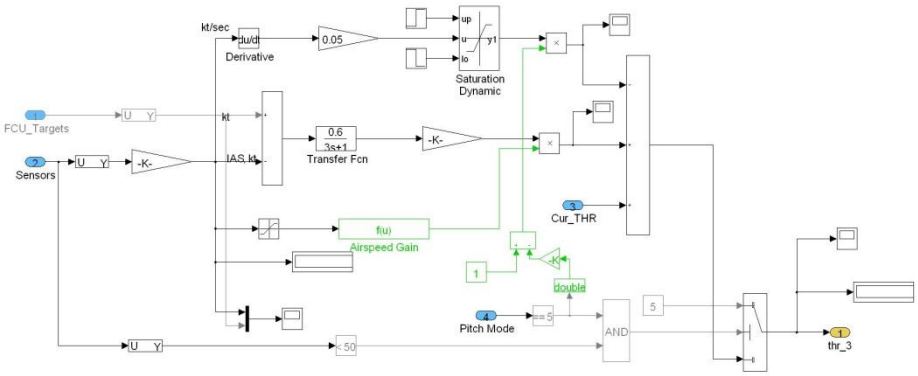


Fig. 9 - Speed hold control law

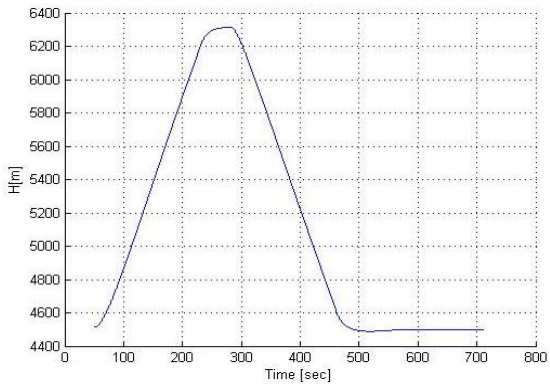


Fig. 10 - Altitude profile, test 2

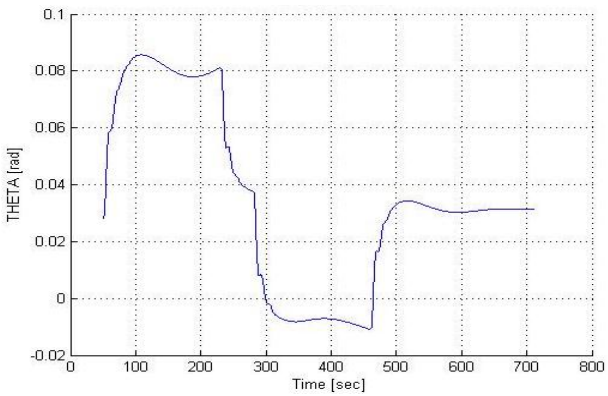


Fig. 11 - Pitch profile, test 2

3rd CEAS Air&Space Conference
21st AIDAA Congress

ADS/IRS/GPS Tightly Integrated Navigation Technology Based on the Fuzzy Adaptive Federated Kalman Filters

Lai Jizhou, Wu Xiuzhi, Wan Hui, and Liu Jianye
Navigation Research Center, Nanjing University of Aeronautics and Astronautics, China

Keywords: *Integrated Navigation System; Tight Integrated; Fuzzy Reasoning; Adaptive Federated Filter*

Abstract

In order to ensure the accuracy and reliability of integrated navigation system, scheme of multisource fusion navigation is usually used. Scheme of Inertial Reference System (IRS)/ Air Data System (ADS)/ Global Positioning System (GPS) integrated navigation is designed. The federal filter structure based on adaptive information distribution is designed for information fusion of multisource system. Fuzzy reasoning technology is introduced to ADS/IRS Kalman filter, and tight integrated mode is used in IRS/GPS filter. The optimal fusion and processing of multisource navigation information can be done through proposed algorithm. The fuzzy adaptive Kalman filtering model of ADS/IRS and tight integrated filtering model of INRS/GPS are built in this paper, and federal filtering algorithm based on adaptive information distribution is designed. Simulation is carried out based on above algorithms, and the results verify the effectiveness of proposed tightly-coupled ADS/ IRS/GPS integrated navigation algorithm based on adaptive fuzzy federated filter.

1 General Introduction

With the development of the sensor technology, the accuracy and reliability of the navigation system are more and more high, but higher request is put forward by the large passenger aircraft. Integrated navigation system based on INS is the main navigation system for most modern large aircrafts. Since INS is a diverging system, its error spreads over time. So INS is generally used combined with another navigation system whose error is relatively constant over time, and Global Position System (GPS) is a good choice. Extra navigation facilities are not need to be introduced to aircraft for ADS. Aircraft airspeed tube and altimeter of ADS belong to the autonomous navigation method like INS. At present, air data inertial reference units (ADIRU) are the main navigation system of the large passenger aircraft [1].

For different types of navigation systems, Kalman filtering technology is often used to fuse information from multiple sources, which occurs a problem that statistical properties of measurement noise is unpredictable. Fuzzy reasoning technology is introduced to Kalman filter technology, and measurement noise is adaptively adjusted through the properties of filter residual, so the system state can be estimated accurately [2].

At present the INS/GPS loose integrated mode is widely used which is easy to be realized, but the problem is obvious: firstly, at least 4 satellites are needed which can't be guaranteed in the actual situation, and the accuracy will be decreased seriously if the receiver doesn't work; secondly, since measurement noise and measurement state are time correlation as position, velocity of receiver are time correlated with its errors, the output of GPS affects navigation precision and error estimation

precision. While GPS doesn't need navigation solution in tight integrated mode, and its output can revise INS through Kalman filter, causing better navigation performance [3-6].

ADS/IRS/GPS federal filtering structure is designed in this paper. ADS/IRS Kalman filter based on fuzzy reasoning technology and IRS/GPS Kalman filter of tight integrated mode are used, and adaptive information distribution algorithm is designed in federal filtering to improve the accuracy and reliability of the integrated navigation system. At last fusion model of ADS/IRS/GPS integrated navigation system is built and simulation is done, the results verify the effectiveness of proposed algorithm.

2 IRS/GPS tight integrated navigation technology

2.1 Theory of IRS/GPS tight integrated navigation

The theory of IRS/GPS tight integrated navigation is shown in Fig. 1: baseline information are pseudorange and pseudorange rate which are receiver's output, and measurement information are pseudorange and pseudorange rate which are between body and satellites. The difference of baseline information and measurement information are observation information. The error of INS and clock error of receiver are estimated through filtering, then open loop correction or close-loop feedback correction can be done [7].

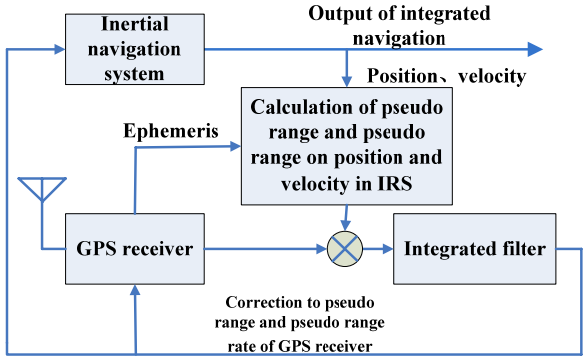


Fig.1 Diagram of pseudorange and pseudorange rate integrated navigation system

2.2 Model of IRS/GPS tight integrated navigation

The state of filter is $\boldsymbol{X}_{IRS}=[\boldsymbol{X}_{IRS}^T,\boldsymbol{X}_{GPS}^T]^T$. And $\boldsymbol{X}_{IRS}=[\varphi_E,\varphi_N,\varphi_U,\delta_{vE},\delta_{vN},\delta_{vU},\delta_L,\delta_\lambda,\delta_h,\varepsilon_{bx},\varepsilon_{by},\varepsilon_{bz},\varepsilon_{rx},\varepsilon_{ry},\varepsilon_{rz},\nabla_x,\nabla_y,\nabla_z]^T$, $\boldsymbol{X}_{GPS}=[\delta\boldsymbol{t}_u,\delta\boldsymbol{t}_{ru}]^T$. The noise of state equation is $\boldsymbol{W}=[w_{gx},w_{gy},w_{gz},w_{rx},w_{ry},w_{rz},w_{ax},w_{ay},w_{az},w_{tu},w_{tru}]$.

The definition of state and noise in INS are the same with ADS/IRS integrated navigation system. Error states of GPS receiver are the distance $\delta\boldsymbol{t}_u$ equivalent with clock error and the distance rate $\delta\boldsymbol{t}_{ru}$ equivalent with clock frequency error. So the state equation of the pseudorange and pseudorange rate integrated navigation system is:

$$\begin{bmatrix}\dot{\boldsymbol{X}}_{INS}(t)\\ \dot{\boldsymbol{X}}_{GPS}(t)\end{bmatrix}=\begin{bmatrix}F_{INS}(t) & 0\\ 0 & F_{ADS}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{X}_{INS}(t)\\ \boldsymbol{X}_{GPS}(t)\end{bmatrix}+\begin{bmatrix}G_{INS}(t) & 0\\ 0 & G_{ADS}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{W}_{INS}(t)\\ \boldsymbol{W}_{GPS}(t)\end{bmatrix} \quad (1)$$

Where $F_{GPS}=\begin{bmatrix}0 & 1\\ 0 & -\beta_{tru}\end{bmatrix}, G=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. The measurements of GPS/IRS integrated navigation are pseudorange difference and pseudorange rate difference. $Z(t)=H(t)X(t)+V(t)$, where $Z(t)=[\rho_{I1}-\rho_{G1},\cdots,\rho_{In}-\rho_{Gn},\dot{\rho}_{I1}-\dot{\rho}_{G1},\cdots,\dot{\rho}_{In}-\dot{\rho}_{Gn}]$,

and $H=\begin{bmatrix}0&0&0&0&0&0&a_{11}&a_{12}&a_{13}&\cdots&1&0\\ \vdots&&\ddots&&\vdots&\vdots&\ddots&\vdots&&\vdots&&\\ 0&0&0&0&0&0&a_{n1}&a_{n2}&a_{n3}&\cdots&1&0\\ 0&0&0&b_{11}&b_{12}&b_{13}&0&0&0&\cdots&0&1\\ \vdots&\ddots&\vdots&\vdots&\vdots&\vdots&&\vdots&\ddots&\vdots&\ddots&\\ 0&0&0&b_{n1}&b_{n2}&b_{n3}&0&0&0&\cdots&0&1\end{bmatrix}_{20\times20}$.

The coefficients are calculated as follows:

$$\begin{cases}a_{n1}=(R_N+h)[-e_{n1}\sin L\cos\lambda-\sin L\sin\lambda]\\ \quad+[R_N(1-e^2)+h]e_{n3}\cos L\\ a_{n2}=(R_N+h)[e_{n2}\cos L\cos\lambda-\cos L\sin\lambda]\\ a_{n3}=e_{n1}\cos L\cos\lambda+e_{n2}\cos L\sin\lambda\end{cases}.$$

$$\begin{cases}b_{n1}=-e_{n1}\sin\lambda-e_{n2}\cos\lambda\\ b_{n2}=e_{n1}\sin L\cos\lambda-e_{n2}\sin L\sin\lambda+e_{n3}\cos L\\ b_{n3}=e_{n1}\cos L\cos\lambda+e_{n2}\cos L\sin\lambda+e_{n3}\sin L\end{cases}$$

λ, L, H are longitude, latitude, and height separately, e_{n1}, e_{n2}, e_{n3} are orientation cosines between the nth satellite and the receiver.

3 ADS/IRS integrated navigation technology based on fuzzy adaptive filtering

3.1 Design of ADS/IRS fuzzy adaptive filtering

System noise $\boldsymbol{W}_k$ and measurement noise $\boldsymbol{V}_k$ are zero-mean Gauss whit noise in traditional Kalman filter, and covariance matrixes are known as Q and R. Statistical properties of measurement noise are unknown for INS/ADS integrated navigation system. System state can be estimated well by adjusting the measurement noise statistical properties through measurement information [8-9].

If the mathematical model is accuracy enough, the residual sequence should be zero-mean white noise and residual variance should be the theoretical residual variance got by Kalman filter. If the deviation is large, it means that measurement noise level has changed and measurement noise covariance matrix R need to be adjusted. According to the above idea, Kalman filter is designed as follows:

$$\begin{cases} \hat{\boldsymbol{X}}_{k/k-1} = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{k-1} \\ \hat{\boldsymbol{X}}_k = \hat{\boldsymbol{X}}_{k/k-1} + \boldsymbol{K}_k[\boldsymbol{Z}_k - \boldsymbol{H}_k\hat{\boldsymbol{X}}_{k/k-1}] \\ \boldsymbol{K}_k = \boldsymbol{P}_{k/k-1}\boldsymbol{H}_k^T[\boldsymbol{H}_k\boldsymbol{P}_{k/k-1}\boldsymbol{H}_k^T + \boldsymbol{R}_k]^{-1} \\ \boldsymbol{P}_{k/k-1} = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}_{k,k-1}^T + \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{\Gamma}_{k-1}^T \\ \boldsymbol{P}_k = [\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k]\boldsymbol{P}_{k/k-1}[\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k]^T + \boldsymbol{K}_k\boldsymbol{R}_k\boldsymbol{K}_k^T \\ \boldsymbol{R}_k = \mu_k\boldsymbol{R}_{k-1} \end{cases} \quad (2)$$

Where $\boldsymbol{R}_k$ is the estimated value of the k step measurement noise matrix, μ_k is the adjustment coefficient of $\boldsymbol{R}_k$.

The ratio of innovations nominal variance matrix's trace and actual variance matrix's trace is set as input of the fuzzy reasoning system in reference [10], and the output is used to adjust the measurement noise, so the differences of measurement noise are omitted during the adjustment. The measurement noise of each state is adjusted separately and μ_k=diag[$\mu_{k1}, \mu_{k2}, \mu_{k3}, \mu_{k4}$], where $\mu_{k1}, \mu_{k2}, \mu_{k3}, \mu_{k4}$ are the adjustment coefficients of east velocity, north velocity, up velocity and height measurement. The difference function of residual variance value are defined as follows:

$$ROR_{kj} = \frac{P_r(j,j)}{C_r(j,j)} \quad (3)$$

where $\boldsymbol{P}_r$, C_r are nominal value, measurement value of residual variance, $C_r = \frac{1}{M}\sum_{i=i_0}^{k} r_i r_i^T$, $\boldsymbol{P}_r = \boldsymbol{H}_k\left(\boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}_{k,k-1}^T + \boldsymbol{\Gamma}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{\Gamma}_{k-1}^T\right)\boldsymbol{H}_k^T + \boldsymbol{R}_{k-1}$. $i_0 = k - M + 1$, M is set through experience.

The rules of fuzzy reasoning system which adjusts $\boldsymbol{R}_k(j,j)$ are set as follows: if the model is accuracy, the value of ROR_{kj} should be 1; if ROR_{kj} is 1, it means that the model accords with the real one. If $ROR_{kj} > 1$, it means that measurement noise increases, so measurement noise variance matrix $\boldsymbol{R}_k(j,j)$ should increase too; if $ROR_{kj} < 1$, it means that $\boldsymbol{R}_k(j,j)$ should decrease.

"Mamdani" fuzzy system is used in this paper, and center of gravity defuzzification is used. Figure 2 is the diagram of fuzzy Kalman filter.

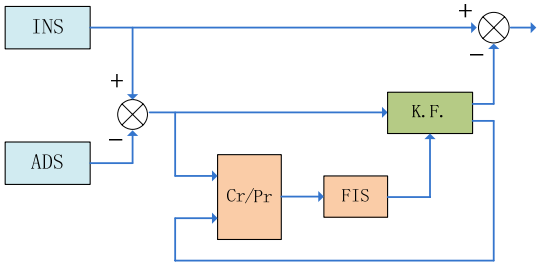


Fig. 2 The principle diagram of fuzzy adaptive Kalman filter

Since statistical properties of measurement noise of each observation are different, so the membership functions of four fuzzy reasoning systems should differ from each other. Take the example of the adjustment of up velocity measurement noise, the input membership functions ROR_{k3} and output membership

functions μ_{k3} of fuzzy reasoning system are set as Fig. 3 and Fig. 4.

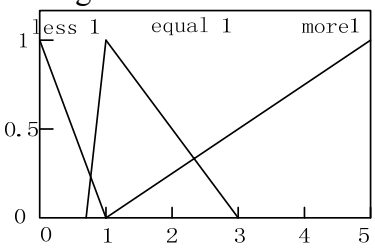


Fig. 3 Membership functions of ROR_{k3}

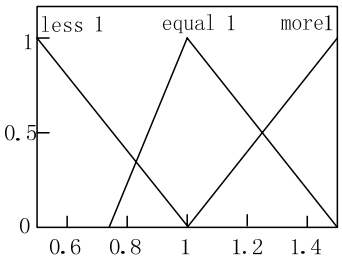


Fig. 4 Membership functions of μ_{k3}

3.2 Filtering model of ADS/INS integrated navigation system

The state of filter is $\boldsymbol{X}_{INS}=[\boldsymbol{X}_{INS}^T,\boldsymbol{X}_{ADS}^T]^T$. And $\boldsymbol{X}_{INS}=[\varphi_E,\varphi_N,\varphi_U,\delta_{vE},\delta_{vN},\delta_{vU},\delta_L,\delta_\lambda,\delta_h,\varepsilon_{bx},\varepsilon_{by},\varepsilon_{bz},\varepsilon_{rx},\varepsilon_{ry},\varepsilon_{rz},\nabla_x,\nabla_y,\nabla_z]^T$, $\boldsymbol{X}_{ADS}=[\boldsymbol{v}_{wbx},\boldsymbol{v}_{wby},\boldsymbol{v}_{wbz},\boldsymbol{v}_{wrx},\boldsymbol{v}_{wry},\boldsymbol{v}_{wrz},\delta\boldsymbol{K}_{ADS},\delta\boldsymbol{h}_b]^T$. The error equation of state equation can be built according to the sensor error model and INS error model. The state variables of INS are defined as reference 3. The noise of state equation is $\boldsymbol{W}=[w_{gx},w_{gy},w_{gz},w_{rx},w_{ry},w_{rz},w_{ax},w_{ay},w_{az},w_b,w_d,w_{vx},w_{vy},w_{vz},w_h]$. The state equation is $\dot{\boldsymbol{X}}(t)=\boldsymbol{F}(t)\boldsymbol{X}(t)+\boldsymbol{G}(t)\boldsymbol{W}$, and all parameters of the equation can be set according to the above differential equation, where

$$F=\begin{bmatrix}F_{INS}&0\\0&F_{ADS}\end{bmatrix},G=\begin{bmatrix}G_{INS}&0\\0&G_{ADS}\end{bmatrix}.$$

The measurement of INS/ADS integrated navigation is set as follows:

$$Z_{ADS}=\left[(v_{EINS}-v_{EADS}),(v_{NINS}-v_{NADS}),(v_{UINS}-v_{UADS}),(h_{INS}-h_{ADS})\right]^T \quad (4)$$

The output of ADS in navigation frame is:

$$\begin{aligned}v_{ADS}^n&=C_b^n\left(1+\delta K_{ADS}\right)\left(v^n+v_w^n\right)\\&=C_b^n\left(1+\delta K_{ADS}\right)\begin{bmatrix}v_E+v_{wbE}+v_{wrE}\\v_N+v_{wbN}+v_{wrN}\\v_U+v_{wbU}+v_{wrU}\end{bmatrix}\end{aligned} \quad (5)$$

The measurement equation is:

$$Z_{ADS}=\begin{bmatrix}H_1&H_{ADS}\end{bmatrix}\begin{bmatrix}X_{INS}\\X_{ADS}\end{bmatrix}+V_{ADS} \quad (6)$$

Where

$$H_1=\left[\begin{array}{ccc:c:c:c}0&v_U&-v_N&&&\\-v_U&0&v_E&I_{3\times3}&0_{3\times3}&0_{3\times6}\\v_N&v_E&0&&&\\\hline&&&0_{1\times6}&0\ 0\ 1&0_{1\times6}\end{array}\right],$$

$$H_{ADS}=\left[\begin{array}{ccc:c}&&&-v_E\ 0\\-I_{3\times3}&-I_{3\times3}&&-v_N\ 0\\&&&-v_U\ 0\\\hline&0_{1\times6}&&0\ -1\end{array}\right].$$

4 Filter configuration of ADS/IRS/GPS integrated navigation system

In order to get the optimal fusion of ADS, IRS and GPS information, the information fusion configuration is designed as Fig. 5. Sub-filters of the federal filter are designed to optimize filter's performance and navigation precision: fuzzy adaptive Kalman filtering technology is used in ADS/IRS sub-filter; tight integrated mode is used in IRS/GPS sub-filter; adaptive information distribution technology is used in level two fusions [11-12].

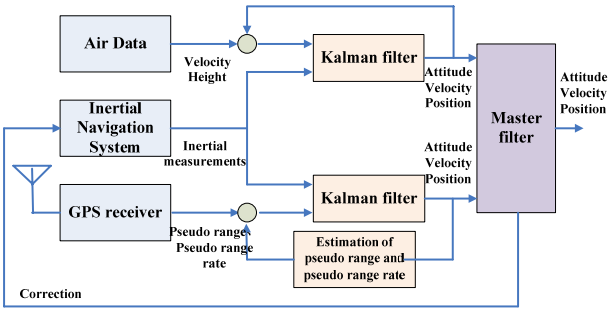


Fig.5 Federal filter structure of ADS/IRS/GPS integrated navigation system

4.1 ADS/IRS/GPS federal filter algorithm

The federal filter is a two-level data fusion structure. Sub-filter and main filter work together. The outputs of all sub-filters are fused through the main filter, and then the global optimal estimation $\hat{\boldsymbol{X}}_{g,k}$ and variance $\boldsymbol{P}_{g,k}$ are got. Then the sub-filters and main filter are reset according to the information distribution principle. The state equation and the measurement equation of the ith sub-filter are set as follows:

$$\boldsymbol{X}_k = \boldsymbol{\Phi}_{k,k\text{-}1}\boldsymbol{X}_{k-1} + \boldsymbol{\Gamma}_{k-1}\boldsymbol{W}_{k-1} \qquad (7)$$

$$\boldsymbol{Z}_{i,k} = \boldsymbol{H}_{i,k}\boldsymbol{X}_{i,k} + \boldsymbol{V}_{i,k} \qquad (8)$$

The time updating, optimal fusion of main filter, information distribution and reset processes can be described as follows:

$$\begin{cases} \boldsymbol{P}_{i,k/k-1} = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{i,k-1}\boldsymbol{\Phi}_{k,k-1} + \boldsymbol{\Gamma}_{k,k-1}\boldsymbol{Q}_{i,k-1}\boldsymbol{\Gamma}^T_{k,k-1} \\ \hat{\boldsymbol{X}}_{i,k/k-1} = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{X}}_{i,k-1} \end{cases} \qquad (9)$$

$$\boldsymbol{P}^{-1}_{g,k/k} = \sum_{i=1}^{n,m}\boldsymbol{P}^{-1}_{i,k},\ \boldsymbol{P}^{-1}_{g,k}\hat{\boldsymbol{X}}_{g,k} = \sum_{i=1}^{n,m}\boldsymbol{P}^{-1}_{i,k}\hat{\boldsymbol{X}}_{i,k} \qquad (10)$$

$$\boldsymbol{P}^{-1}_{i,k} = \beta_i\boldsymbol{P}^{-1}_{g,k},\ \boldsymbol{Q}^{-1}_{i,k} = \beta_i\boldsymbol{Q}^{-1}_{g,k},\ \sum_{i=1}^{n,m}\beta_i = 1\ (i = 1,\cdots,n,m) \qquad (11)$$

4.2 Adaptive information distribution technology used in ADS/IRS/GPS federal filter

The fault of sub-filter can affect the state valuations since sub-filter's contribution to the filter. So distance function is built and distribution coefficients of sub-filters are self-adaptive adjusted, so as to improve the filtering effect. For the public state variables of two sub-filter, adaptive factor adjusting principles are designed as follows:

$$\beta_i(k) = \frac{1/\lambda_i(k)}{1/\lambda_1(k) + 1/\lambda_2(k) + \cdots + 1/\lambda_n(k)} \qquad (12)$$

Wher β_i is weighted factor, $\lambda_i(k)$ are functions of constructed distance, and $\lambda_i(k) = \boldsymbol{r}_{i,r}(k)\boldsymbol{P}^{-1}_{i,r}(k)\boldsymbol{r}^T_{i,r}(k)$. $\boldsymbol{r}_{i,r}(k)$ are prediction errors of sensors based on the global optimal filtering valuation. $\boldsymbol{r}_{i,r}(k) = \boldsymbol{Z}_i(k) - \hat{\boldsymbol{Z}}_i(k)$, $\boldsymbol{P}_{i,r}(k) = E[\boldsymbol{r}_{i,r}(k)\boldsymbol{r}^T_{i,r}(k)]$.

It can be seen from Eq. 13, if accuracy of the ith sensor decreases, the coefficient of the ith sub-filter will adaptively decrease too. The filter error of each system changes all the time in dynamic system, $\lambda(k)$ changes too, so β_i is real-time, and the method can be used to adjust the weighting factor real-time.

5 Simulation and Results

Simulation is carried out based on the fusion model of the integrated navigation system, the output frequency of INS is set as 50Hz. The bias of gyros of two sets of INS is set as 0.03 (°)/h, and bias of accelerometers is set as $10^{-4}g$. The bias of gyros of another set of INS is set as 0.1 (°)/h, and bias of accelerometers is set as $5*10^{-4}g$. The output frequency of ADS is 1Hz, the errors of ease, north, and up wind constant velocity are set as 5m/s,3m/s and 1m/s separately, and the errors of ease, north, and up wind correlation velocity are set as 5m/s,3m/s and 1m/s separately, correlation time is 30 seconds. The error of altimeter is set as 50m, correlation time is 1000 seconds. Scale factor error is 0.003. A typical flight path are taken which includes level flight, take off and landing, and slalom-like flight. The initial position is $[29.4^\circ, 106.6^\circ, 100m]$, and the initial heading is 90°.

IRS/GPS simulation results can be seen in Fig. 6 and Fig. 7, the dotted line represents loose integrated mode and the solid line represents tight integrated mode. It can be obviously seen that tight integrated mode has a higher accuracy than loose integrated mode. Both latitude error and longitude error are less than 5m, and height error is less than 10m.

The residual value function of Kalman filter is shown in Fig. 8. It can be concluded from Fig. 8

that the residual value function defined in Kalman filter tends to 1 after using fuzzy reasoning system to adaptive control, it means that residual variance is almost same with the theoretical value.

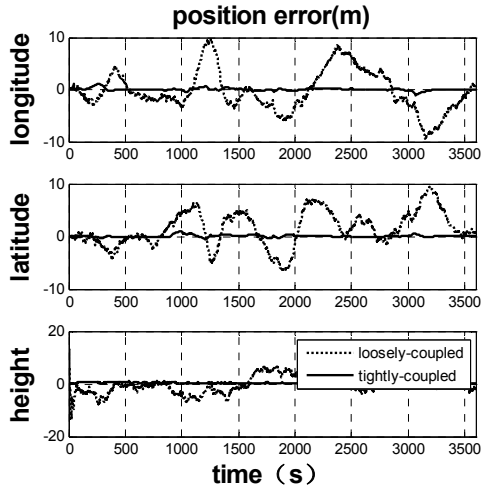


Fig. 6 Position error of IRS/GPS integrated and tight integrated navigation system

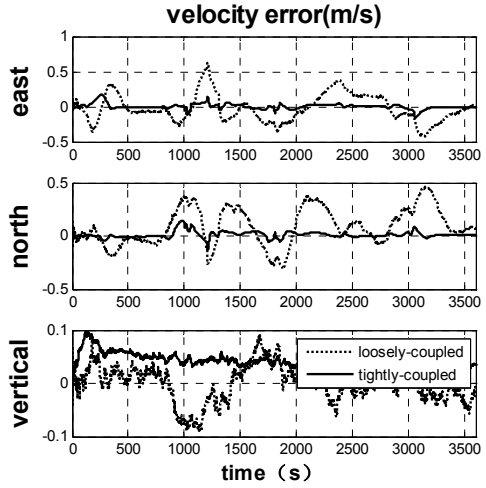


Fig. 7 Velocity error of IRS/GPS integrated and tight integrated navigation system

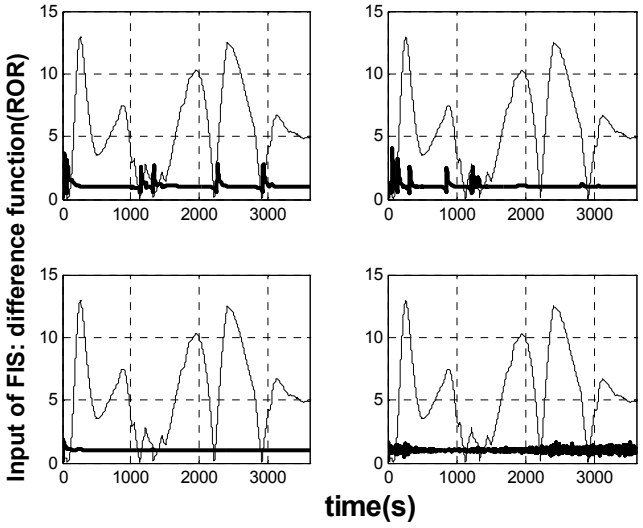


Fig. 8 Residual of Kalman filter covariance function of IRS/ADS, before and after using fuzzy reasoning system

The velocity error of INS/ADS integrated navigation system is shown in Fig. 8: thick solid line represents integrated navigation results; two thin lines represent two same sets of INS/ADS integrated navigation results; the dashed represents a set of INS/ADS integrated navigation results which has different accuracy. In order to analysis the performance of the integrated navigation system, the velocity errors of integrated navigation and one set of INS/ADS integrated navigation are compared in Table 1.

It can be seen from Fig. 8 that the error of INS/ADS integrated navigation system with low accuracy INS is larger than the other two sets of INS/ADS integrated navigation system with higher accuracy INS, but the integrated navigation results between two conditions are almost same. It can be concluded that when several sets of INS are working together, adaptive weighted algorithm can make integrated navigation as the same accuracy with the good subsystem. From Table 1 it can be seen that the velocity error variance of integrated navigation is low and the navigation performance is good. So during the navigation processing, if fault occurs in one set of INS, its contribution to the total integrated navigation will be decreases, and integrated navigation system will remain high accuracy.

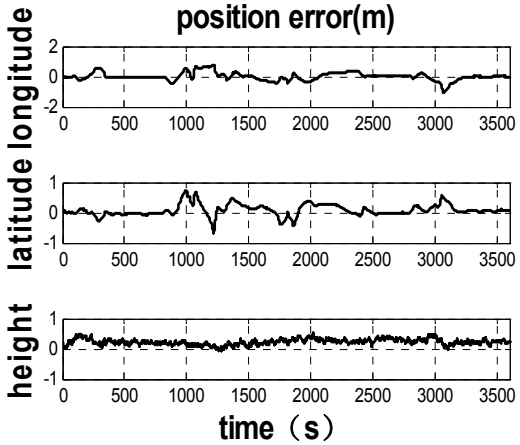


Fig. 9 Position error of IRS/GPS/ADS integrated navigation system

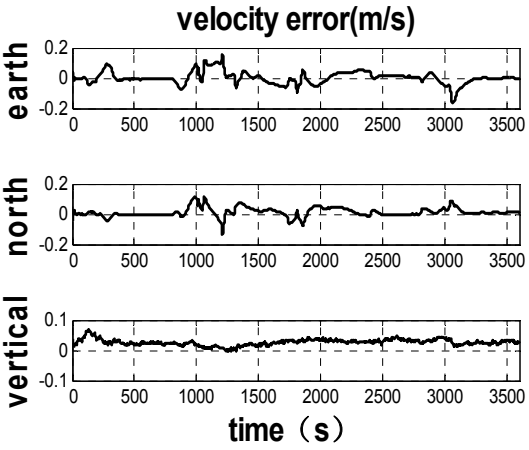


Fig. 10 Velocity error of IRS/GPS/ADS integrated navigation system

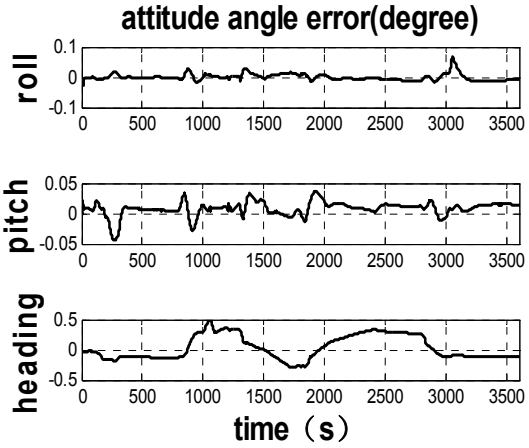


Fig. 11 Attitude error of IRS/GPS/ADS integrated navigation system

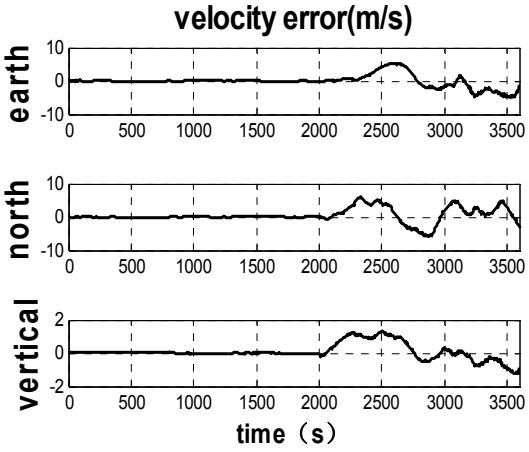


Fig. 12 Velocity error of GPS/ADIRS integrated navigation system, GPS is failure after 2000 seconds

The position, velocity and attitude error of ADS/IRS/GPS can be seen in Fig. 9, Fig. 10 and Fig. 11 separately. A good performance can be got through the integrated algorithm designed in this paper. The velocity of ADS/IRS/GPS integrated navigation system is shown in Fig. 12, and GPS is failure after 2000 seconds. It can be seen from Fig. 12 ADS/IRS integrated navigation system works well after GPS's outage, and the error of integrated velocity doesn't diverge since velocity/height integrated mode is adopted.

6 Conclusions

Multi-source information fusion algorithm is designed to the ADS/IRS/GPS integrated navigation system which is widely used in civil aviation passenger plane. Fuzzy inference technology is used to improve the performance of Kalman filter, and IRS/GPS tight integrated mode is adopted which is better than loose integrated mode. All information of sub-filters can be effectively used by using adaptive information distribution technology, positioning accuracy can be guaranteed when GPS is failure. The scheme supplies a method for information fusion of ADS/IRS/GPS multi-source navigation system.

References

[1] Eugenio Denti, Roberto Galatolo, Francesco Schettini, "Aircraft State Estimation:Inertial&Air Data Systems", *2nd European Conference for Aero-Space Sciences,* .Brussels, Belgium, 2007, pp. 1-6.

[2] Fu J, Zhang X F, Bian H W, "Adaptive Kalman filter based on fuzzy logic for marine integrated navigation system", Journal of Chinese Inertial Technology, Vol. 15, No. 4, 2007, pp. 412~417.

[3] Soloviev A, "Tight coupling of GPS, laser scanner, and inertial measurements for navigation in urban environments", Position, Location and Navigation Symposium, 2008, pp. 511~525.

[4] Miller I, Schimpf B, Campbell M, "Tightly-coupled GPS/INS system design for autonomous urban navigation". Position, Location and Navigation Symposium, 2008, pp. 1297~1310.

[5] Mark Manfred, Thomas Ryno, "Tightly Coupled GPS/IRS Navigation for ADS-B", Integrated Communications Navigation and Surveillance Conference, 2008.

[6] Weidong Zhou, Xiang-wei Qiao, Fanbin Meng and Hebing Zhang, "Study on SINS/GPS Tightly Integrated Navigation Based on Adaptive Extended Kalman Filter", Proceedings of International Conference on Information and Automation, 2010, pp. 2344~2347.

[7] Liu Jianye, Zeng Qinghua, Zhao Wei, "Navigation system theory and application",

Nanjing university of aeronautics press, 2007.
[8] Yang Yuanxi, Gao Weiguang, "Comparison of Adaptive Factors in Kalm an Filters on Navigation Results", The Journal of Navigation, Vol. 58, No. 3, 2005, pp. 471~478.
[9] Xu Tianlai, You Wenhu, Cui Pingyuan, "Research on GPS/INS Integrated navigation System Based on Fuzzy Adaptive Kalman Filtering", Journal of Astronautics, No.5, 2005, pp. 571-575.
[10] Qin Yongyuan, Zhang Hongyue, "Kalman filter and integrated navigation principle", Northwest industry university press, 1998.
[11] Call Curt, Mike Ibis, Jim McDonald, "Kevin Vanderwerf. Performance of Honeywell's Inertial/GPS Hybrid (HIGH) for RNP Operations", Proceedings of Position Location and Navigation Symposium, 2006.
[12] McDonald, Jim, Joshua Kendrick, "Benefits of Tightly Coupled GPS/IRS for RNP Operations in Terrain Challenged Airport", Proceedings of Position Location and Navigation Symposium, 2008.

CEAS 2011

ATM Airport Operations 2

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Description of Complex Runway Usage with Queuing Theory

R. Mori
Electronic Navigation Research Institute, Japan

Keywords: *Queuing theory, Runway allocation, Priority queue, Airport capacity*

Abstract

The runway operation at Tokyo International Airport became complicated with the opening of the fourth runway in October 2010. However, the runway allocation is basically fixed by destination/departure airports, which is not optimized in terms of fuel consumption and waiting time. In this paper, the runway operation is described by queuing theory, and the characteristics of airport operations are revealed. First, actual runway operation data is obtained and analyzed. Next, assuming certain simplifications, the runway operation is described by queuing theory. Finally, considering several possible traffic scenarios, the optimal runway allocation is examined, and the impact of factors such as take-off separation and position of the runway is revealed.

Nomenclature

$E(W)$: average waiting time of all aircraft.

$E(W_i)$: average waiting time on runway *i*.

$E(B_i)$: average processing time (take-off or landing separation time) on runway *i*.

$E(R_i)\left(=\frac{E(B_i^2)}{2E(B_i)}\right)$: average residual processing time on runway *i*.

$E(L_i^q)\left(=\lambda_i E(W_i)\right)$: average number of aircraft waiting on runway *i*.

λ_i : traffic amount on runway *i*.

$\mu_i\left(=\frac{1}{E(B_i)}\right)$: average service rate on runway *i*.

$\rho_i\left(=\frac{\lambda_i}{\mu_i}<1\right)$: probability of runway occupation on runway *i*.

1 Introduction

The airport congestion is recently a critical issue in many airports in the world. The bottleneck of the traffic flow is usually found on runways because the number of the take-offs and landings are limited due to the aircraft separation. As a result, long waiting queues are found on the ground taxiway and in the air. The decrease of departure and arrival queues will lead to the decrease in the fuel consumption.

A 4D trajectory operation can decrease queues even at the same traffic volume. Under a 4D trajectory operation, the flight plans of all aircraft are decided in advance, i.e., static queuing, and all aircraft are to be controlled under strict time restrictions. In such a case, in theory, departure and arrival queues will disappear if the flight plans are developed well. However, under this concept, once an aircraft cannot satisfy the restriction, many following aircraft become affected. Therefore, dynamic queuing is also required for the future system, and the goal of this research is to develop a real-time dynamic runway allocation system.

Several approaches for dynamic runway allocation system have been proposed. Issacson et al. proposed a knowledge-based runway

assignment algorithm[1], while Fritzsche et al. tried to calculate the potential improvement of the dynamic runway allocations[2]. Although both approaches are promising, the target airport runway operation is relatively simple. In addition, the degree to which the results are optimized remains unclear. Janic proposed a heuristic runway allocation algorithm[3] where a complicated airport is considered. However, only a rough estimation of the runway usage is considered, and the waiting time of the individual aircraft is not taken into account.

The target airport of this research is Tokyo International Airport where the fourth runway was opened in October 2010. As the runway operation becomes more complicated, a sophisticated runway allocation algorithm is required. In order to achieve a dynamic allocation of such a complicated airport, the understanding of airport operation characteristics is necessary. Therefore, in this paper, as a first step, the runway operation is simplified, and it is described by queuing theory[4][5]. This is an analytical method, so it is very fast to calculate and easy to understand. Then, the characteristics of the runway operation are examined, and the relationship between the parameters (e.g., the amount of traffic, the trend of the waiting time) is revealed.

This paper is organized as follows. Section 2 starts with the details of the runway operation in Tokyo International Airport, and the actual take-off and landing data is obtained. In Section 3, a simple queuing theory is introduced, and the runway operation is described. In Section 4, based on a simple model, the characteristics of the runway operation are revealed. Finally, the paper is summarized in Section 5.

2 Runway Operations at Tokyo International Airport

2.1 Tokyo International Airport and its runway operation

Tokyo International Airport is the biggest airport in Japan, and is mostly used for domestic flights. In 2010, the fourth new runway was opened and more spots for international flights became available. It is expected that the annual traffic volume will increase to 447,000 flights in the future, although currently it stands at 303,000 flights[6]. Although there are four runways, the runways cross each other so a mutual interaction of the runways exists. Under this condition, the optimized runway allocation is not obvious.

The maximum capacity of take-off and landing is expected to be 40 aircraft each per hour. Figure 1 shows the airport map and a typical operation under south wind at maximum capacity. As mentioned earlier, there are four runways in the airport. A and C runways are 3,000 meters long used for take-off aircraft, while B and D runways are 2,500 meters long for landing aircraft. However, there are some restrictions of the runway usage. First, when aircraft lands on D runway, aircraft cannot take off from A and C runways, or they would collide. Second, A and B runway intersect at the end. These runways can be used independently as long as the aircraft accept the short take-off. However, heavy aircraft sometimes need more length of the runway for take-off, and in such a case, the two runways simultaneously cannot be used. In addition, D runway is located further from the terminal buildings, so it takes more time for aircraft to taxi.

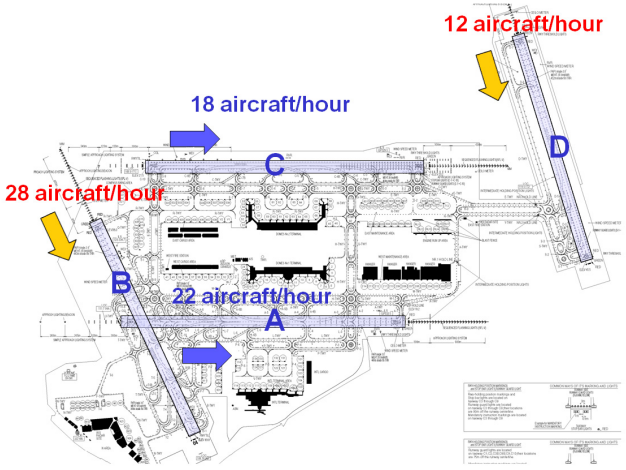


Fig. 1 The airport map and typical operation under south wind.

Currently, the runway usage depends entirely on the direction of departure/arrival airports. A runway is used for westbound departure aircraft with at most 22 aircraft per hour, whereas B

runway serves southbound arrival aircraft with a rate of 28 aircraft per hour, C runway is for eastbound departure aircraft with 18 aircraft per hour, and D runway is for northbound arrival aircraft with 12 aircraft per hour.

This static runway allocation system favors reduced ATC workload, because ATC does not have to occupy themselves with the runway allocation for each aircraft. However, it is also obvious that this allocation is not optimized in terms of fuel consumption and waiting time of aircraft. As mentioned before, a dynamic allocation would be the best solution, but this paper aims at revealing the characteristics of the airport operation first. Therefore, this paper considers the static runway allocation, and answers the following questions based on queuing theory.

- What is the optimized static runway allocation?
- How is the waiting time influenced by traffic volume and airport configuration?

2.2 Take-off and landing aircraft data

Before applying queuing theory, several data are required, i.e. the average take-off and landing time separation, the ratio of aircraft using long runway for take-off, and additional taxiing time to D runway. The data is obtained for one week in December 2010, even though the runway was not operated under the considered south-wind conditions at all times.

First of all, the take-off and landing time separations are considered. These time separations are not constant, as they depend on the combination of preceding and following aircraft, as well as include some randomness. However, this time in order to apply queuing theory, it is assumed that these time separations are always constant.

In extracting take-off separation, the minimum time separation is available only when aircraft take off continuously. Therefore, the take-off time separation data is extracted based on the following conditions: there is an aircraft waiting for take-off just before the runway when the preceding aircraft takes off. As for landing aircraft, it is difficult to judge whether each landing separation can be reduced or not. Therefore, the mode value of the landing separation for all landing data is used as reference landing separation. The landing separation is discretized by 5 seconds. D runway is used less often, so only B runway data is used. Although the landing separation is estimated less accurately, it is sufficient for the purpose of this research. Table 1 shows the take-off and landing separations on each day.

The number of take-off aircraft from C runway is smaller than that from A runway, so the difference of take-off separation on C runway between days is relatively big. However, there is no clear difference of take-off separation between runways, so the take-off separation is assumed to be 104 s. Similarly, the landing separation is assumed to be 126 s.

In terms of long take-off on Runway A, the number of aircraft applying long take-off is obtained. However, only 4 aircraft used long

Table 1 Take-off and landing separation

	Average take-off separation (number of data)		Mode of landing separation (number of data)
	A Runway	C Runway	B Runway
12/20	103.8s (12)	110.9s (16)	115s (56)
12/21	(0)	(0)	(0)
12/22	104.4s (25)	98.2s (14)	115s (64)
12/23	(0)	(0)	(0)
12/24	103.6s (43)	107.6s (35)	125s (151)
12/25	105.1s (94)	106.9s (51)	130s (373)
12/26	104.5s (84)	99.5s (32)	125s (354)
Total	104.5s (258)	104.4s (132)	125.7s (998)

take-off out of 702 aircraft during this week, so the effect of the long take-off aircraft is neglected this time. However, in the future, the number of international flights is expected to increase, which may increase the number of aircraft applying long take-off runway.

Finally, the effect of taxiing time is considered. D runway is located further from the terminal buildings, which requires additional taxiing time compared to other runways. Considering the waiting time of departure aircraft, the total taxiing time is important. Therefore, if an aircraft uses D runway, the additional taxiing time should be added even if it executes smooth taxiing. Consequently, the average taxiing time of landing aircraft is obtained. In general, the landing aircraft go taxiing smoothly, so their taxiing time can be assumed to be the same as the taxiing time without congestion. Table 2 shows the average taxiing time of landing aircraft on each runway. Clearly, the taxiing time from D runway is the longest with more than 1 minute longer than that from B runway. This additional taxiing time will be considered in the simulation.

Table 2 Average taxiing time of landing aircraft on each runway.

	Average taxiing time of landing aircraft	Number of samples
A Runway	216.64 s	1601
B Runway	343.60 s	998
C Runway	215.12 s	523
D Runway	416.38 s	325

3 Description of Runway Usage with Queuing Theory

According to the conditions written in Sec. 2.2, the average waiting time on each runway is calculated. This time, queuing theory is used to describe the runway operation. To simplify the problem, the following assumptions are introduced. All queues are assumed to be M/D/1.

1) Arrival aircraft are always prior to departure aircraft, and all queues are non-preemptive.

2) A and C runways have independent queues.

3) The departure and arrival separations are always the same (deterministic), respectively, i.e.

$$\mu_A = \mu_C = \mu_{dep},\ \mu_B = \mu_D = \mu_{arr},$$

$$E(R_A) = E(R_C) = \frac{1}{2\mu_{dep}},\ E(R_B) = E(R_D) = \frac{1}{2\mu_{arr}}$$

4) The intersection of B runway (arrival) and A runway (departure) is not considered.

3.1 B runway

B runway is totally independent of the other traffic. Only a single flow of arrival aircraft should be considered. Therefore, basic queuing theory is applied to calculate the waiting time.

$$E(W_B) = E(L_B^q)E(B_B) + \rho_B E(R_B) \qquad (1)$$

Finally, the waiting time is obtained as the following expression.

$$E(W_B) = \frac{\rho_B}{2\mu_{arr}(1-\rho_B)} \qquad (2)$$

3.2 D Runway

D runway is not independent of other flows, and is affected by departure aircraft from A and C runways. At D runway queue, priority is given to departure aircraft, so priority queuing theory is applied. However, priority queuing theory is based on a single priority queue and a single low-priority queue, although there are two low-priority queues this time. Therefore, the combined low-priority queue is assumed, and the waiting time on D runway is calculated based on priority queuing theory.

$$E(W_D) = E(L_D^q)E(B_D) + \rho_{AC}E(R_{AC}) + \rho_D E(R_D) \qquad (3)$$

where ρ_{AC} indicates the probability of runway occupation on runway A or C. The flow of A and C runways is randomly distributed between $1-\rho_D$, so the approximate probability of

runway occupation on runway A and C (defined as $\rho_{A\&C}$) is calculated.

$$\rho_{A\&C} = (1-\rho_D)\frac{\rho_A}{1-\rho_D}\frac{\rho_C}{1-\rho_D} = \frac{\rho_A\rho_C}{1-\rho_D} \quad (4)$$

Next, ρ_{AC} is calculated.

$$\rho_{AC} = \rho_A + \rho_C - \rho_{A\&C} \quad (5)$$
$$= \rho_A + \rho_C - \frac{\rho_A\rho_C}{1-\rho_D}$$

Finally, the waiting time on runway D is calculated as follows.

$$\frac{\left(\rho_A + \rho_C - \frac{\rho_A\rho_C}{1-\rho_D}\right)\frac{1}{2\mu_{dep}} + \frac{\rho_D}{2\mu_{arr}}}{1-\rho_D} \quad (6)$$

The D runway is further from the terminal building, which require longer taxiing time, as shown in the last section. Therefore, this factor is considered as additional waiting time (Δw_{D0}).

$$E(W_D) \quad (7)$$
$$= \frac{\left(\rho_A + \rho_C - \frac{\rho_A\rho_C}{1-\rho_D}\right)\frac{1}{2\mu_{dep}} + \frac{\rho_D}{2\mu_{arr}}}{1-\rho_D}$$
$$+\Delta w_{D0}$$

3.3 A and C runways

A and C runways have less priority to arrival aircraft than D runway. Since A and C runways are equally treated, only A runway is considered here. The average waiting time on runway A is formulated based on priority queuing theory.

$$E(W_A) = E(L_A^q)E(B_A) + E(L_D^q)E(B_D) \quad (8)$$
$$+\rho_A E(R_A) + \rho_D E(R_D) + \rho_D E(W_D)$$
$$+w_C$$

This runway has less priority to D runway, so the waiting time on D runway is added as the second, fourth and fifth terms. The last term shows the correction terms to express two low-priority queues. The difference between a single low-priority queue and two low-priority queues are found in the additional processing time on C runway when D runway is queued. The extended waiting time is defined as w_C. The details are shown in Fig. 2. The difference between a single low-priority queue and two low-priority queues becomes obvious in the following scenario. Suppose an aircraft is waiting for take-off on A runway. Before the aircraft takes off, an aircraft is queued on D runway followed by an aircraft on C runway. The additional waiting time Δt_{AC} is equal to t_C. Therefore, w_C can be calculated based on the following formula.

$$w_C = \int_0^{\frac{1}{\mu_{dep}}} d\tau \rho_D \mu_{arr} \tau \int_\tau^{\frac{1}{\mu_{dep}}} dt \rho_C \mu_{dep} \quad (9)$$
$$= \frac{\rho_C\rho_D\mu_{arr}}{6\mu_{dep}^2}$$

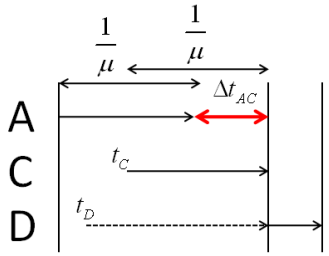


Fig. 2 The difference in the additional waiting time.

Finally, the average waiting time on A runway is calculated.

$$E(W_A) = \left\{\frac{\rho_A}{2\mu_{dep}} + \frac{\rho_D}{2\mu_{arr}} + \rho_D E(W_D) \right. \quad (10)$$
$$\left. + \frac{\rho_C\rho_D\mu_{arr}}{6\mu_{dep}^2}\right\} \Big/ \left(1-\rho_A-\rho_D\right)$$

The average waiting time on C runway is also calculated in the same way.

$$E(W_C) = \left\{\frac{\rho_C}{2\mu_{dep}} + \frac{\rho_D}{2\mu_{arr}} + \rho_D E(W_D) \right. \quad (11)$$
$$\left. + \frac{\rho_A\rho_D\mu_{arr}}{6\mu_{dep}^2}\right\} \Big/ \left(1-\rho_C-\rho_D\right)$$

The average waiting time of all aircraft can be found by the following expression.

$$E(W) = \{\lambda_A E(W_A) + \lambda_B E(W_B) + \lambda_C E(W_C) + \lambda_D E(W_D)\} / (\lambda_A + \lambda_B + \lambda_C + \lambda_D) \quad (12)$$

The obtained formulas are evaluated by a numerical simulation.

3.4 Parameter values

In order to calculate the waiting time, several parameters are required. Actually, μ_{dep}, μ_{arr}, λ_A, λ_B, λ_C, λ_D, w_{D0} are needed. λ_A, λ_B, λ_C and λ_D are the traffic demand, which are the conditions of the simulation. μ_{dep}, μ_{arr}, and w_{D0} are obtained from the actual data as shown in Sec. II, 3600/104, 3600/126, 60/3600, respectively. The unit is hour.

4 Simulation Results

4.1 Calculation assumptions

The waiting time depends on the departure and arrival traffic. To simplify the problem, the departure traffic is evenly distributed between A and C runways. ($\lambda_A = \lambda_C = \frac{\lambda_{dep}}{2}$) The arrival traffic is now distributed with regard to the direction of the departure airport, and the ratio of aircraft using D runway (defined as γ) is currently set to 0.3. Using γ, the arrival traffic demand can be written as follows: $\lambda_B = (1-\gamma)\lambda_{arr}$, $\lambda_D = \gamma\lambda_{arr}$. When the departure and arrival traffic are fixed, the waiting time is calculated in terms of γ.

4.2 Calculation results under several scenarios

4.2.1 Calculation result for normal traffic density

First of all, the characteristic of the waiting time is revealed. The departure and arrival traffic are assumed to be 30 aircraft/hour, which is the current traffic at most times. Figure 3 shows the calculation result. $E(W_A)$ and $E(W_C)$ have the same curve type, because it is assumed that the departure aircraft are distributed evenly to A and C runways. As γ increases, the waiting time of A and C runways increases because the departure is blocked by arrival aircraft on D runway. On the other hand, as γ increases, the waiting time of B runway decreases, because the traffic to B runway decreases with higher γ. At D runway, a priority is always given to departure aircraft, so the waiting time is relatively smaller than that of other runways. The optimal waiting time is 3.31 minutes achieved when γ is 0.246. Even if γ is 0.3, the waiting time is 3.63 minutes, which is just 0.32 minutes higher than the optimal one. Therefore, the current distribution of arrival aircraft is very acceptable when normal traffic density is assumed.

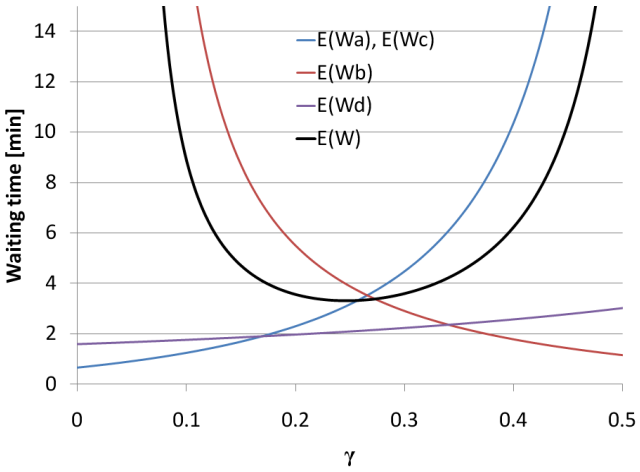


Fig. 3 Average waiting time on each runway. (departure: 30/hour, arrival: 30/hour)

4.2.2 Impact of traffic density of waiting time

However, the traffic volume is not constant throughout the day. In the morning, the number of departure aircraft is much higher than that of arrival aircraft, and vice versa at night. This time, the following two cases are considered based on actual traffic data.

a) Departure traffic: 40 aircraft/hour, arrival traffic: 26 aircraft/hour

b) Departure traffic: 5 aircraft/hour, arrival traffic: 36 aircraft/ hour

The result is shown in Fig. 4. Here, only average waiting time in each scenario is shown. The result clearly shows that the optimal distribution is completely changed depending on the scenarios. When the departure traffic is relatively high, the optimal γ is shifted to the low end so that the arrival aircraft to D runway do not interfere with departure aircraft. In the same way, γ is relatively high when arrival traffic is big. The difference of the average waiting time between $\gamma = 0.3$ and the optimal one is more than 2 minutes, which infers that much improvement can be made.

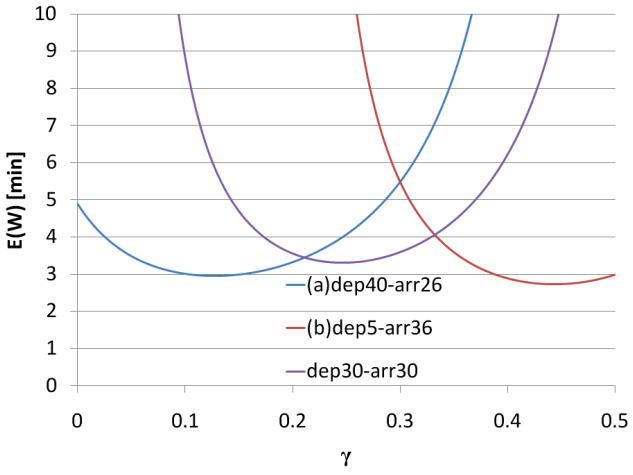


Fig. 4 Average waiting time under various traffic.

4.2.3 Effects of taxiing duration

As mentioned earlier, D runway is located far from the terminal building, so additional taxiing time (Δw_{D0}) is required when D runway is used. According to the obtained data, Δw_{D0} is about 1 minute, but the effect of Δw_{D0} is considered here. Therefore, the relationship between γ and average taxiing time is examined when Δw_{D0} is 0, 1, and 2 minutes. Figure 5 shows the calculation result. The dotted line shows the result when Δw_{D0} is 0, the solid line: 1 minute, and the dashed line: 2 minutes. Naturally, the difference of average taxiing time is bigger when the number of arrival aircraft is bigger. However, we should focus on the optimal γ in each Δw_{D0}. In all scenarios, the optimal γ is almost constant even when Δw_{D0} is changed. Table 3 shows the detailed relationship between Δw_{D0} and γ. According to the result, the optimal γ is changed within 0.011, corresponding to 0.39 aircraft/hour. This result infers that a few minutes additional taxiing time to D runway does not affect the optimal traffic flow. In other words, the additional taxiing time will not be a trigger to change the runway usage.

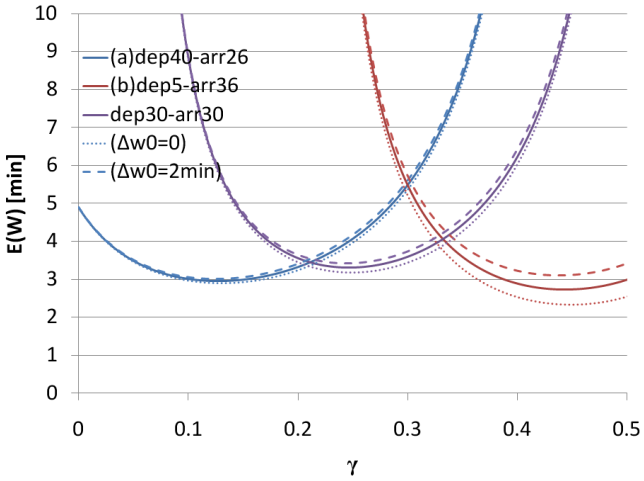


Fig. 5 Influence of taxiing time on average waiting time.

4.2.4 Capacity and take-off/landing separation.

Judging from the obtained traffic data, the airport cannot accommodate 40 departure and 40 arrival traffic because take-off and landing separation is too long. It does not mean that the airport cannot manage the objective capacity technically, because the take-off and landing separation can be reduced. The question is that

Table 3 The effect of Δw_{D0} to optimal γ.

	departure: 30, arrival: 30		departure: 5, arrival: 36		departure: 40, arrival: 26	
Δw_{D0}	Optimal γ	aircraft/hour	Optimal γ	aircraft/hour	Optimal γ	aircraft/hour
0 min	0.2487	7.46	0.4489	16.16	0.1302	3.39
2 min	0.2438	7.32	0.4379	15.77	0.1248	3.24

how much time should be reduced. The possible reduction times of take-off and landing separation are defined as Δt_{dep} and Δt_{arr}, respectively. Therefore, μ_{dep} and μ_{arr} are calculated as follows; $\mu_{dep} = \frac{3600}{104 - \Delta t_{dep}}$, $\mu_{arr} = \frac{3600}{126 - \Delta t_{arr}}$. The average waiting time with various Δt_{dep} and Δt_{arr} is calculated and shown in Fig. 6. To achieve less than 10 minutes average waiting time, at least 15 s reduction of landing separation is required. According to the figure, Δt_{arr} gives about two times greater impact than Δt_{dep}. For example, almost the same average waiting time is observed when $(\Delta t_{dep}, \Delta t_{arr}) = (0, 20), (10, 15)$ or $(\Delta t_{dep}, \Delta t_{arr}) = (0, 25), (10, 20)$. In addition, not surprisingly, the optimal γ is shifted to positive when Δt_{dep} increases, and to negative when Δt_{arr} increases. Unlike in the case of additional taxiing time to D runway, the difference of the optimal γ between the various sets of Δt_{dep} and Δt_{arr} is not negligible. The bigger Δt_{dep} and Δt_{arr}, the smaller the average waiting time. However, the optimal γ is not the same even if the optimal average waiting time is the same. To minimize the average waiting time, the solution should be changed depending on the situation.

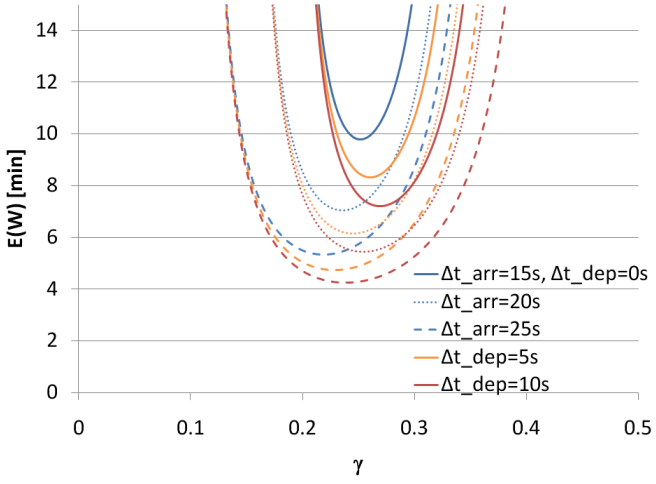


Fig. 6 Impact of the reduction of the take-off and landing separation. (departure: 40/hour, arrival: 40/hour)

5 Conclusions

The congestion of airport runways is one of the biggest issues for aviation. Optimal runway allocation is a promising answer considering the limited airport capacity. This time, the runway operation at Tokyo International Airport was simplified and modeled by queuing theory. Several parameters were obtained with the actual runway operation data. According to the calculation results, the optimal distribution of the runway usage was changed depending on the traffic and the take-off/landing separation. However, the location of the runway would not be a big issue in terms of the distribution of the runway usage. In the future, it should be verified how much the simplified model agrees with the actual runway congestion.

References

[1] Isaacson, D. R., Davis, T. J., Robinson III, J. E., "Knowledge-based Runway Assignment for Arrival Aircraft in the Terminal Area, *AIAA Guidance, Navigation, and Control Conference*, New Orleans, August 11-13, 1997.

[2] Fritzsche, M., Gunther, T., and Fricke, H., "Potential of Dynamic Aircraft to Runway Allocation for Parallel Runways," *4th International Conference on Research in Air Transportation*, Budapest, June 1-4, 2010.

[3] Janic, M., "A Heuristic Algorithm for the Allocation of Airport Runway System Capacity," *Transportation Planning and Technology*, Vol. 30, No. 5, pp. 501-520, 2007.

[4] Bhat, U. N., *An Introduction to Queuing Theory*, Birkhauser, 2008.

[5] Buzacott, J. A., and Shanthikumar, J. G., *Stochastic Models of Manufacturing Systems*, Prentice Hall, 1993

[6] Japan Civil Aviation Bureau, http://www.mlit.go.jp/kisha/kisha04/12/120226/03.pdf (in Japanese)

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Proof-of-Concept of an Algorithm for Aircraft Departure Sequencing

Urbano Tancredi
University of Naples Parthenope, Napoli, Italy

Keywords: *Air Traffic Control, Departure Sequencing, Combinatorial Optimization*

Abstract

An algorithm is presented for establishing an optimal sequence of aircraft departing from an airport. Under a specific choice of the requirements and objectives, the problem is formulated as a combinatorial optimization task. The algorithm blends exhaustive evaluation of the possible sequences with a sliding sub-horizon technique. The validity of the algorithm's concept is evaluated by application on a realistic traffic data sample for Milan Malpensa airport, concerning an entire day period in which a significant traffic load was recorded. Results suggest that the algorithm is capable of delivering the benefits for which it has been designed. Indeed, the algorithm exhibits variations w.r.t. the real traffic sample in the defined performance metrics that are in complete accordance with the expected effects. For instance, the algorithm allows saving around 15 % of the overall time the aircraft spend with active engines, with related cost, pollution and noise benefits.

1 Introduction

In today procedures, the control tower handles departing traffic at an airport on a "first come – first served" principle, which is based on the actual time or order that aircraft request pushback and taxi clearance. Sequencing of departures is done manually when practicable by use of "best practices". The objective of a runway departure sequencing function is instead to establish an optimal sequence in which aircraft can depart from the available runways at an airport and start their initial climb phase. This is done minimizing the waiting times and maximizing the efficiency of the operations, under compliance to applicable technical and operational rules that restrict the usage of runways, such as separation criteria for aircraft, timeslots in which aircraft should depart, and aircraft performance limits. A Departure Manager (DMAN) is a decision support tool for replacing the current manual procedures with a set of ATC tools that accommodate the information exchange and provides an optimized flow of outbound traffic [1]. The DMAN mainly provides departure schedules to achieve an optimal use of the runways and to improve the organization of the outgoing traffic. One of the core elements of a DMAN is the Departure Sequencing (DS) algorithm, which is in charge of determining the optimal sequence of departures, as well as the correct timing of prominent events, such as take-off and off-block times.

The present paper focuses on developing an algorithm for sequencing aircraft departures at

an airport whose complexity is representative of a typical Italian hub and to prove the validity of the developed concept by evaluating the algorithm performances on a realistic traffic sample. The algorithm has been designed for handling DS problems that are consistent with the DMAN operational concept proposed by EUROCONTROL [1 – 4], which reflects the main features of relevant research in progress at European airports [5 – 7].

The DS problem is typically treated as an optimization problem with a discrete component (sequencing) and a continuous one (timing of prominent events). This typology of optimization problems is very complex to tackle, and is usually approached introducing some simplifying assumptions that transform the original intractable problem into a simpler one whose solution can be obtained by applying some well-established methods (e.g. [8]). The approach followed in this paper is instead to avoid optimizing the continuous part of the problem, and to solve the DS problem with combinatorial optimization algorithms. This problem relaxation is possible since most objective functions in DMAN applications are monotone w.r.t. the timing, i.e. earliest is always better [9, 10]. Thus, by opportunely designing the objective function, one can explicitly enforce optimality of the continuous part of the problem, and thus approach the DS as a discrete optimization. Furthermore, the DS problem is typically over constrained in CFMU-regulated airports [9, 11]. A novel approach is proposed in this paper to deal with this feature of the problem. It is based in managing hard constraints for avoiding unfeasible departure sequences, and to optimize amongst possible sequences using a hierarchical objective function [12] that incorporates performance metrics and soft constraints.

2 Problem Setting

The DS problem may be seen as establishing a Departure Plan (DP) subject to various inputs and constraints for maximizing a series of objective functions. The DP comprises the sequence of aircrafts which have to take-off from a (or several) runway(s), and the time epochs at which the events of each aircraft's departure process shall occur. Let us refer to standard EUROCONTROL's nomenclature, which can be found in [13]. The departure process is seen as a series of departure milestones, whose definition can be found in [13]. The sequence timing consists in assigning a time epoch to all relevant departure milestones.

As previously mentioned, because of the features of the DP, finding the optimal DP is an optimization problem with a discrete component (sequencing) and a continuous one (milestones timing). Such Nonlinear mixed Discrete – Continuous Optimization (NDCO) is a highly complex problem, for which no well-established approaches exist. Without referring to the DMAN framework, the most popular way for solving such problems is of linearizing the continuous part (constraints and objectives). This allows recasting the NDCO as a Mixed Integer Linear Program (MILP). A solution can be generally achieved for the MILP, by combining the very efficient Linear Programming solution methods, to somehow more involved Integer Programming techniques.

In the DMAN context, only reference [14] follows this approach, whereas for scheduling landing aircraft, this is a widely used approach. This discrepancy arises mainly because the determination of milestones timing is less critical in DMAN applications, when compared to the determination of the sequence. In fact, references [9], [10] argue that most objective functions in DMAN applications are monotone w.r.t. the timing, i.e. earliest is always better. Because of this assumption, the milestones timing is uniquely determined by the sequence. This allows one to avoid optimizing the continuous part of the problem, implying that only the sequence in which the aircraft depart determines the optimality of a solution. As such, the majority of departure planning problems are approached with combinatorial optimization algorithms that consider only the discrete part of the problem.

Furthermore, the DS problem is typically over constrained in CFMU-regulated airports [9,

11]. This implies that one has to face the situation in which no sequence exists that complies to all the constraints. In order to provide at least one DP, the problem shall be stated allowing the solutions to violate, when necessary, at least one of the constraints. This is usually accomplished by dividing the constraints in:

- **Hard constraints**, that shall never be violated, and to which any feasible DP shall comply.
- **Soft constraints**, that should not be violated as far as practical, but can be violated if no compliant DP exists.

We refer herein to the operational concept developed by the SICTA consortium under contract by ENAV, the Italian ANSP. We briefly recall its basic requirements

- **Planning Horizon.** The planning horizon length shall be traded off depending on various factors, such as the computational burden, the maximum number of runways, aircrafts, and gates to be handled, the capacity and traffic load of the airport, and on the accuracy of the data necessary to completely define the DP. Typical values are around 20 – 50 minutes ahead in time.
- **Update Rate**. Departure planning is a highly dynamic environment. Thus, the DP shall be regularly updated reflecting the change over time of input data, constraints, and/or objective functions. Typical values for update rate are in the order of 1 minute.
- **Stability and traceability of solutions**. It is of upmost importance that the Air Traffic Controller (ATCo) is able to understand the rationale behind the modification of the DP due to a change in the situation, as well as to agree with the optimal DP proposed by the algorithm. Indeed, if these conditions are not satisfied, the ATCo will probably avoid using the tool.
- **Tool authority**. The DMAN tool is assumed to have no planning authority on departing aircraft after the aircraft starts the off-block procedure, that is, the DMAN plans up to the AOBT. This is because the developed DMAN concept does not foresee any data update on the aircraft status after it goes off-block, mainly to avoid any additional workload for the ATCo. Furthermore, the DMAN is assumed not to be in charge of determining the runway (RWY) in use and the ICP / SID for departing aircraft, mainly because their selection depends on variables that are out of the DMAN scope, such as noise abatement procedures, ATFM regulations, and so on. Moreover, the DMAN is assumed to have no authority on arriving flights. As such, the DMAN ultimate function is to plan TSAT (see [13] for the TSAT definition) in order to allow for a realistic establishment of the optimal departure sequence with a minimum surface occupancy by the aircraft with motors idle.

The following observations are drawn on the features that the DS algorithm shall possess.

- Allow relaxation of (some) constraints. In particular, validation of other DS algorithms has shown that traffic flow restrictions must be treated as soft constraints [6]–[9], [11].
- Formulate the problem as a combinatorial optimization using objective functions that are monotone w.r.t. the timing, i.e. earliest is always better.
- Heuristics must be used for solving the problem. In fact, even for a combinatorial optimization, the problem complexity considering the order of magnitude of the number of aircraft to be planned in a meaningful time horizon (that determines the problem dimension) is not compatible with the update rates needed in DMAN applications
- Iterative algorithms are not desirable. The fixed update rate is not compatible, in general, with iterative algorithms. Indeed, these algorithms have no guarantee of convergence in fixed time.
- Random global optimization methods shall be avoided. These algorithms are generally unable to produce same solution under same input, and the variation of the solution may be disproportioned from input variation. In DMAN applications, the ATCo would thus be unable to understand the rationale for the DP modification, probably disregarding DMAN suggestions.

2.1 Choice of the Algorithm Typology

The algorithm for determining the DP shall thus be capable of dealing with the following tasks:

- Optimize the DP in a Sliding Time Horizon, considering all aircraft starting from TOBT – X minutes up to AOBT, and updating the solution every Y minutes. Typical values for X are 20 – 50 minutes and Y is less than 1 minute.
- In each time horizon, formulate the problem as a combinatorial optimization. Use an exact algorithm (e.g. enumeration-based), modified by introduction of problem - specific heuristics to reduce the computational burden.
- Enforce equity requirements and solution traceability implicitly, by limiting sequence deviations from the nominal one. This further allows simplifying the problem solution, with related computational benefits.
- Differentiate constraints between hard and soft ones, where the CTOT shall be treated as a soft one.

Among the various departure planning algorithms analyzed, the following two techniques have been selected for the DS algorithm design:

- **Exhaustive Evaluation**, which consists in evaluating all the possible departure sequences, that is, in determining the timing for all departure sequences. This technique is used by NLR's OPS (see [11] for details).
- **Sliding sub – horizon technique**, which is a problem specific heuristic consisting in obtaining a complete sequence of aircraft in the time horizon by optimizing a smaller sequence several times, thus allowing to reduce the number of possible sequences. This technique is used both by NLR's OPS [11] and by DLR's DMAN [9].

As far as the exhaustive evaluation is concerned, it provides a significant advantage in terms of algorithm simplicity, with resulting benefits in terms of algorithm design and development times. Furthermore, it is also an exact algorithm, which explores the whole design space providing thus the globally optimal solution. Unfortunately, this can be done only for problems of very small size. Indeed, when the planning task comprises N aircraft, the number of possible sequences is N!, given by all possible permutations of the n aircraft, implying an almost exponential increase in the number of possible sequences with n. For instance, if N = 5, there are 120 possible sequences, that increase to over 3.5 millions in case N = 10. As such, no more than O(5) aircraft can be reasonably analyzed in seconds by conventional computers. Exhaustive evaluation techniques are thus not appropriate by themselves, but need to be augmented with additional techniques that allow reducing the number of sequences to be analyzed, that is, for the current application framework, the sliding sub – horizon technique.

The way sliding sub – horizon techniques do not explore the whole design space is easily understandable, which is desirable for traceability requirements. In this case the algorithm optimizes a sequence of n aircrafts out of the N that are present in the whole time horizon, with $n < N$, and freezes the first $k < n$ aircrafts in the sequence, then slides the sub-horizon of k places. The sliding sub-horizon slides from the first aircraft in the sequence to the last one. Therefore, an aircraft can be anticipated in the sequence for no more than n places, whereas it can be delayed up to the end of the sequence. The algorithm thus does not explore the portion of design space made by sequences in which at least one aircraft is anticipated for at least n+1 places w.r.t. the initial sequence. This feature also provides some inherent equity to the solution explored by the algorithm, since deviations from the nominal sequence are limited (even though only one-way). On the other hand, the initial sequence provided to the algorithm assumes a notable significance, since the sub-optimal solution will heavily depend on its choice.

3 Algorithm Design

3.1 Reference Departure and Arrival Timelines

The reference timelines of aircraft served by the airport have been devised based on EUROCONTROL's milestones approach (ref. [13]): in response to the intended TOBT, the

ATC issues a TSAT, which is the primary DMAN output. We also assume, as ref. [13] does, that the start-up and off-block phases occur simultaneously. The TOBT, provided to DMAN via CDM, is considered as a constraint, which dictates the earliest time in which the TSAT can by issued by DMAN. The time necessary to the aircraft to take-off from TSAT is broken down in three parts:

- **Estimated Time To Runway (ETTR)**, is the time necessary to the aircraft to reach the Runway Holding Position (RHP), or to enter in the runway queue. This time comprises a portion of the estimated taxi out time (EXOT).
- **Runway Holding Time (RHT)**, is the time the aircraft will spend at the runway holding position either because it arrives at the RHP before the runway is available for take-off, or because there is a departing queue.
- **Runway Occupancy Time, Departure (ROTD)**, is the time necessary to the aircraft for taking off, once the line-up clearance has been issued.

As a consequence of this reference timeline, the following timing relationship exists between the TTOT and the TSAT.

$$TTOT = TSAT + ETTR + RHT + ROTD \quad (1)$$

The reference arrival timeline foresees two events:

- **Estimated Landing Time (ELDT)**, provided to DMAN via CDM, is the estimated time epoch in which an aircraft will touchdown on the runway.
- **ELDT + Runway Occupancy Time, Arrival (ROTA),** is the time epoch in which the landed aircraft will exit the runway, passing (or arriving to in case a clearance is required to enter the taxiway) the runway exit point.

As a consequence of this reference timeline, the following timing relationship exists between the Time at which the runway becomes physically available and the ELDT:

$$\textit{Time at which RWY becomes available} = ELDT + ROTA \quad (2)$$

3.2 Hard Constraints Models

The algorithm philosophy is to optimize the departure sequence based on the TTOT of the aircraft in the sequence. The optimal TTOT of each aircraft in the sequence is always the earliest take off time that is consistent with all the hard constraints. From the so obtained optimal TTOT sequence, the algorithm issues the TSAT sequence, based on Eq.(1).

The hard constraints taken into account are:

- Runway occupancy times by arriving aircraft.
- ATC separations between arriving and departing flights.
- Wake Vortex separations between arriving and departing flights.
- ATC separations between consecutive departing flights.
- Wake Vortex separations between consecutive departing flights.
- Earliest time at which each departing aircraft can take-off. This is obtained from the TOBT and the timing relationship of Eq.(1).
- CTOT. As previously mentioned, the CTOT is considered as a soft constraint, rather than a hard one. Nevertheless, we argue that CTOT violations might occur due to the effects of other constraints that do not allow an aircraft to take off inside its slot, but not because the aircraft is anticipated too much by the ATC. Hence, the time epoch at which the slot begins is treated as a hard constraint only on the earliest possible TTOT.

The effects of the hard constraints on the TTOT are to inhibit for a certain time interval the aircraft take-off from the runway. The strategy designed to handle all the constraints effects on the single aircraft TTOT foresees two phases. First, some constraints are not directly referred to the TTOT (e.g. TOBT is referred to the start-up time. As such, the ROTD, RHT, and ETTR shall be taken into account for determining the earliest possible TTOT). All such constraints are expressed in terms of take-off times exploiting Eq.(1). Second, since the effect of the hard constraints is to identify the time interval(s) in which the TTOT can occur, all the constraints are combined using interval theory. More precisely, the overall effect is obtained by making the union of all the finite

and infinite time intervals in which the take-off is inhibited due to the various constraints. The time intervals available for the TTOT are then determined simply by making the complement of this union with the set of positive real numbers.

The strategy designed to handle the hard constraints allows determining the optimal TTOT simply as the earliest epoch within the available take-off time interval, with related computational benefits. Furthermore, all the departure sequences are feasible. The only effect of the hard constraints is thus to determine the timing of a certain departure sequence, that is, the time epochs at which the milestones of the various aircraft in the sequence can occur at best. Thus, all sequences can be evaluated, even though many of them will be clearly far from being optimal.

Summarizing, the logical flow of actions to be performed by the algorithm, given a departure sequence, is:

- Evaluate all hard constraints for all the aircraft in the sequence to determine the time intervals available for each TTOT
- Fix the TTOT of all the aircraft in the sequence as the earliest time epochs in the relevant time intervals.
- Compute the TSAT of all the aircraft in the sequence using Eq.(1).
- Evaluate the sequence optimality.

The next logical step in the algorithm design is to devise the objective function, combining both soft constraints and preference functions.

3.3 Objective Functions Models

Given the diversity of the nature and of the importance of the various objectives and constraints concurring to determining the overall objective function, the option of using a vector objective function was discarded. This indeed would have been either intractable for the problem at hand, or would have required to be combined into a scalar objective function by applying weighting factors (e.g. [11]), which results in the optimal solution heavily depending on a number of tunable parameters with little physical meaning. The approach designed to tackle the DS problem is to consider a hierarchical objective function [12], made of several scalar objective functions, the "layers," arranged in order of importance. The hierarchy works by evaluating all solutions by the top layer, and passing to the next layer only those solutions with the maximum score based on the first layer's objective function. A tolerance can also be defined on the scalar objective functions, creating a "gate" on each layer. The gate can be established to pass to next layers not only solutions with the maximum score, but also the ones with a near maximum score. This allows more solutions to be evaluated by the deepest layers, providing a more balanced trade-off between top and bottom objectives.

The top layer of the objective function is the compliance to the CTOT, in terms of TTOT within the slot. The scalar objective function of this layer is the opposite of the number of aircraft that violate the slot in the sequence. Thus, only the sequences with the minimum possible slot violations among all the sequences are passed to the next layer. In this way, the CTOT compliance, which is a constraint, is relaxed only when no solutions exist allowing to comply with all slots. Moreover, to avoid issuing a TSAT for a TTOT very near to the slot limits, a 30 seconds tolerance is enforced.

If more solutions with the same number of missed slots (possibly zero) exist, then they are passed to the second layer. The second layer of the objective function is concerned with the maximization of the runway throughput. This is implemented as the opposite of the root-mean-square value of the TTOT of all aircraft in the sequence. In this way, the algorithm is able of maximizing the runway throughput in a continuous manner in the departing sequence.

If more solutions with the same runway throughput exist, then they are passed to the third layer. The third layer of the objective function is concerned with providing some equity to the optimal solution. This is done considering the waiting times of the departing aircraft, that is, the difference between the time at which an aircraft estimates to be ready (TOBT) and the time at which it is cleared for starting to move (TSAT). This is implemented

as the opposite of the root-mean-square value of the difference between the TSAT and the TOBT of all aircraft in the sequence. In this way, the algorithm is able of minimizing the waiting times in a continuous manner in the departing sequence.

The fourth (and bottom) layer considers adherence of the actual situation with the planned one. This is done taking as the objective function the opposite of the root-mean-square value of the difference between the ETOT and the TTOT of all aircraft in the sequence. In this way, the algorithm is able of minimizing deviations from the planned schedule in a continuous manner in the departing sequence.

3.4 Algorithm Logic

As outlined in section 2.2, the algorithm chosen to solve the DS problem is an exhaustive evaluation of the possible sequences applying a sliding sub-horizon technique. An example run of the algorithm is here discussed for describing the algorithm optimization logic. Let us consider a planning horizon comprising N aircraft. The complete sequence will thus be a sequence of length N, and all possible permutations of the N aircraft yield N! possible sequences. The sliding sub-horizon technique works instead on a smaller subsequence of length n < N. When the sliding sub-horizon technique is first applied to the complete sequence, it considers the first n aircraft of the working sequence. Let us call this subsequence the working subsequence. At this point, the algorithm exhaustively evaluates this subsequence. First, it determines all the possible n! subsequences. Then, computes the timing of each subsequence by projecting the hard constraints. Once the timing is computed, it applies the hierarchical objective function. In particular, it searches for the subset of all the n! subsequences that have the minimum number of missed slots, and then applies the other layers, depending on how many solutions pass through each layer's gate. At the end of this phase, the algorithm determines the optimal subsequence. The first n aircraft of the working sequence are rearranged based on the determined optimal subsequence, even though only the first k < n aircraft are frozen in the optimal sequence, the output sequence. Then, the sub-horizon slides of k places on the complete sequence and the algorithm iterates until all the complete sequence is optimized.

As outlined in section 2.2, this algorithm is only sub-optimal. Being sub-optimal, the quality of its results depends on the initial sequence on which the algorithm is applied. In particular, the output sequence will not consider sequences in which any aircraft is anticipated more than n places with respect to its position in the initial sequence. Hence, the choice of the initial sequence is of great significance to obtain adequate results. Because of the emphasis on slot compliance, the initial sequence is chosen in such a way to maximize the number of aircraft that adhere to their slot. In addition, since there are cases in which the aircraft cannot match its slot due to problems before the taxi phase, the initial sequence takes explicitly into account if an aircraft can match its slot based on its TOBT. Clearly, such initial sequence is expected to deliver benefits in terms of slot compliance, but at the expenses of the other objectives, most notably the runway throughput.

4 Algorithm Evaluation

4.1 Benchmark case description

The benchmark case on which the DMAN validation is performed considers real traffic data from Milan Malpensa (LIMC) airport, recorded on July 13, 2009. The traffic sample extends over a one-day period, and comprises 283 departing flights, of which 134 from RWY 35R and 149 from RWY 35L, and 300 arriving flights, of which 135 on RWY 35R and 165 on RWY 35L.

All the four possible RWY in use combinations (e.g. 35L for arrivals / 35R for departures) are used in the period under analysis. Given ICAO runway dependency constraints and LIMC published ICPs (see [15]–[17]), runways 35L and 35R may be treated as independent within the DMAN scope.

In order to implement the DS algorithm for LIMC airport, models shall be developed for computing separation constraints on Take-Off and the Estimated Time To Runway (ETTR. These are described in the following subsections.

An equivalent separation model has been implemented in DMAN for enforcing separations between consecutive departing flights, which depends only on the engine types of the lead and follower aircraft (AC), and is specified in Table 1.

Table 1. Equivalent Model of ATC Separations between Consecutive Departing Flights

Lead AC \ Follower AC	Jet	Turboprop
Jet	1 min	1 min
Turboprop	2 min	1 min

ATC separations between arriving and departing flights have been fixed based on the traffic sample data and on operational practices, assuming that no take-off can take place less than 60 seconds before a landing occurs. Wake vortex separations have been fixed according to ICAO standards [16]. Runway occupancy times by arriving and departing aircraft are assumed to be equal for all arriving AC, by specifying a ROTA = ROTD = 30 s. The RHT is set constant and equal to 90 s. for enforcing runway pressure.

In order for DMAN to be able to compute TOBT-relative constraints on the earliest possible TTOT, a taxi time has been developed for LIMC. More precisely, a model has been provided to DMAN for computing the ETTR for each flight, that is, the time necessary to reach the RHP from the stand after engine start-up. The ETTR model has been obtained by statistical analysis of the taxi times of the traffic sample, following the approach used in [18]. Given the low number of aircraft available in the traffic sample, an ETTR model for each stand would have resulted in ETTR values of low significance, being inferred from few taxi times' samples. Thus, the model developed uses ETTRs that are valid for group of stands rather than for individual stands, defined based on the stand type (e.g. push-back, self-maneuvering, etc.) and on the taxiways for arriving at the runways.

4.2 Numerical results

The following results have been obtained optimizing a subsequence of 7 aircraft, freezing the first 3, i.e. n=7, k=3, which turned out to be a good compromise between flexibility and computational efficiency. The DP in each planning horizon was obtained with an algorithm running time of about 1 s. on a standard personal computer, in face of a O(1 min.) required refresh rate.

A series of performance metrics are shown, comparing their value on the real traffic sample with the corresponding values computed on the DMAN-optimized DP. Results suggest that the DMAN performances agree with the expected effects of a valid DS algorithm. More specifically, the hourly airport departure capacity remains substantially unchanged by the DS algorithm, as expected. Indeed, actual ATC departure procedures are already capable to fully exploit the airport capacity, but letting many aircraft queue at the RHP, with related pollution, safety and economical disadvantages. Also CTOT compliance remains unchanged and full CTOT compliance is guaranteed both in the real traffic sample and in the DMAN sequence.

The time spent on the airport surface per flight is more regular for the DMAN sequence, ranging from ~6 to ~23 minutes, and does not exhibit the high peak values of 40 and higher minutes which characterize the real traffic sample. Consequently, the total time spent on the airport surface is significantly smaller for the DMAN sequence than for the real traffic. In particular, the DS algorithm allows saving 10.3 hours of time with active engines over the real traffic sample in the analyzed period. Considering that in the real traffic case the total time with active engines amounts to ~75 hours, these results suggest that DMAN allows saving around 15 % of the overall time the aircraft spend with active engines. Thus, the DS algorithm has the potential to deliver a significant improvement in terms of fuel

savings, with related cost, pollution and noise benefits.

The runway queuing time per flight in the DS algorithm is always equal to the RHT, which, in this case, is equal to 1.5 minutes. This is true for all flights except the first ones departing from RWY 35R and 35L, whose RHT is null. Results show that in real traffic data runway queuing times as high as 30 minutes exist, which are avoided by DMAN by strategic planning of the start-up approval times. This allows also gaining benefits from a safety point of view, having less AC maneuvering simultaneously on the airport surface. The RWY Pressure in the DMAN sequence is much more regular than in the real traffic sample, with one or two AC in the runway queue for the majority of time, instead of having peak values of 5 or even six aircraft in the queue. These results demonstrate the capability of the DS algorithm to account for a non null queue at the runway, and thus to deliver a sufficient RWY pressure for optimal exploitation of the runway.

At last, the waiting times at gate per flight in the DMAN sequence are below 5 minutes in more than 90% of the flights. The remaining 10% is mostly made of regulated flights, which are commanded to wait at the stand in order not to arrive at the runway earlier than their slot.

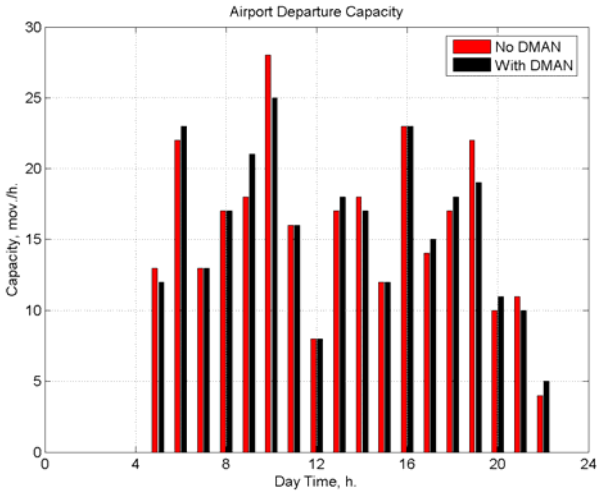


Fig. 1 Hourly Airport Departure Capacity

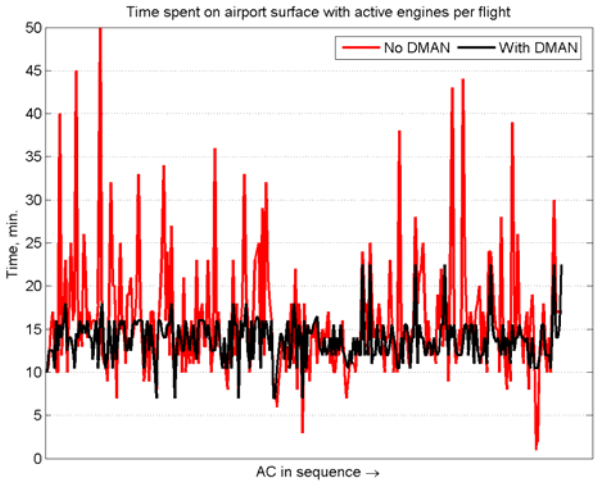


Fig. 2 Time spent on airport surface with active engines per flight

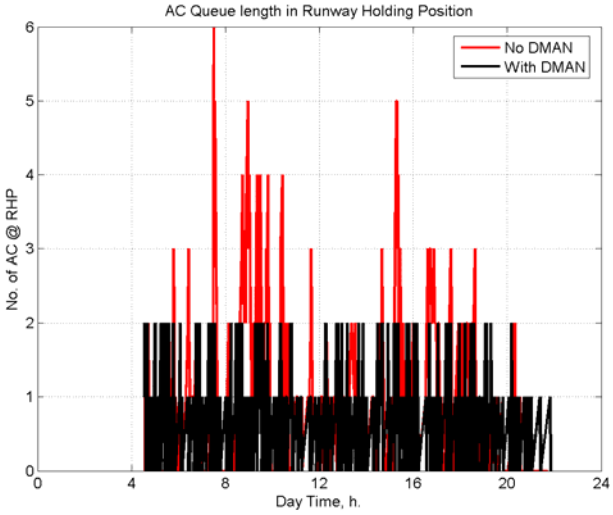


Fig. 3 Runway queue length

5 Conclusion

The present paper has described the proof-of-concept of a Departure Sequencing algorithm. To this end, LIMC was chosen as the airport for validating the algorithm, and a real traffic sample was considered, concerning an entire day period in which a significant traffic load was recorded. The algorithm has been adapted to the benchmark case chosen for validation, considering LIMC airport layout and rules, and real traffic data from LIMC airport covering a whole day period.

DMAN validation is performed investigating if, in the considered realistic case, the algorithm is capable of delivering the benefits expected by a DMAN tool. The DS algorithm capabilities

are quantified by introducing a set of performance metrics. Results of the departure sequence optimization validate the DS algorithm, by demonstrating that it is capable of delivering the benefits for which it has been designed on a real application case of an Italian hub.

More precisely, results suggest that DMAN allows saving around 15 % of the overall time the aircraft spend with active engines. Thus, the DS algorithm has the potential to deliver a significant improvement in terms of fuel savings, with related cost, pollution and noise benefits. Moreover, the DMAN strategic planning of the start-up approval times is capable of avoiding runway queuing times as high as 30 minutes which are instead present in real traffic data, in face of a waiting time at the stand that for over the 90% of the flights is below 5 minutes. This yields also benefits from a safety point of view, having less aircraft maneuvering simultaneously on the airport surface. All these benefits are obtained while maintaining full exploitation of the airport departure capacity, capability of letting regulated flights match their slots, and adequate runway pressure having one or two aircraft waiting at the runway holding position in most cases.

Acknowledgements

This research was developed under a research grant issued by the SICTA consortium, within a research project led by ENAV. The author wishes to thank all SICTA and ENAV personnel involved in the activities.

References

[1] Huck, V., Blajeva, T. “DMAN Operational Concept Development”, Draft Document, EUROCONTROL, Version 0.4, 2006.

[2] Haschke, K., Tuinstra, E. “DMAN integration in Airport CDM & A-SMGCS,” EUROCONTROL 2008

[3] Huck, V., Lagios, E., “The EUROCONTROL DMAN Prototype - Description of DMAN in the A-CDM context,” Version 0.4, 2010

[4] Haschke, K., Tuinstra, E. “Generic Operational Concept for Timely Delivery of Aircraft for Take-Off Enabled by DMAN Integration in Airport CDM and A-SMGCS,” Draft Document, EUROCONTROL, Version 07A, 2009.

[5] Huck, V., Lagios, E., “Airport CDM DMAN Evaluation at Brussels Airport Zaventem,” EUROCONTROL, Version 1, 2008.

[6] Böhme et al., “Tactical and Pre-Tactical Departure Planning,” *5th ATM R&D Symposium*, DLR Braunschweig, Germany, 11.–13. Oct. 2005.

[7] Tuinstra E. “NLR Departure Planner Tool – Scope, requirements, and component descriptions for Departure Planner Controller working position,” NLR-CR-2004-321, NLR, Amsterdam, 2004.

[8] Gupta, G., Malik, W., and Jung, Y. C. “A Mixed Integer Linear Program for Airport Departure Scheduling,” *AIAA ATIO Conference*, AIAA-2009-6933. 2009.

[9] Böhme, “Tactical Departure Management With The Eurocontrol / Dlr Dman,” *ATM seminar*, 2005.

[10] Anagnostakis et al., “Runway Operations Planning and Control: Sequencing and Scheduling”, *AIAA Journal of Aircraft*, Vol. 38, No. 6, 2001.

[11] Tuinstra, “Outbound Punctuality Sequencing by Collaborative Departure Planning”. *6thUSA / Europe ATM R&D Seminar*, Baltimore, 2005.

[12] Lau, H. C., M. Sim, K. M. Teo. “Vehicle routing problem with time windows and a limited number of vehicles,” European Journal of Operations Research, Vol.148, pp.559–568, 2003.

[13] EUROCONTROL, “Airport CDM Implementation – The Manual”, Version 3.1, 2010.

[14] Cooper, et al.,” Use of the departure enhanced planning and runway/taxiway assignment system (DEPARTS) for optimal departure scheduling at busy airports,” *Proceedings of the 21st Digital Avionics Systems Conference*, 2002

[15] ENAV, AIP-Italia, AD 2 LIMC, June 2009.

[16] ICAO PANS “Air Traffic Management”, Doc 4444 ATM/501, 14th edition, 2001

[17] ICAO Doc 9643- AN-Conf/11-IP/3, “The Manual of Simultaneous Operations on Parallel or Near-parallel Instrument Runways”, 2004

[18] Clewlow, R., Simaiakis, I., Balakrishnan, H., “Impact of Arrivals on Departure Taxi Operations at Airports,” *AIAA Guidance, Navigation, and Control Conference*, AIAA 2010-7698, 2010, Toronto, Canada

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

A Method for Improving Alerting Performance in Runway Incursion Alerting Systems

A. Sammut
University of Malta, Malta

D. Zammit-Mangion
University of Malta, Malta
Canfield University, United Kingdom

Keywords: *runway conflict detection, Kalman filtering, uncertainty propagation, early detection, false and missed warning.*

Abstract

A solution to the problem of runway incursions is long overdue. The known technique of detecting runway incursions through the use of the concept of a protected zone in which no other entity must exists to ensure the safety of the aircraft in take-off or landing has been shown to be valid. The challenge, however, is to ensure an acceptably low probability of false alerts caused by uncertainties, whilst still maintaining system effectiveness through a reasonable level of successful conflict detections. In fact, the selection of the protected zone width is critical in controlling the false and missed detection rates. This paper addresses two techniques applied in this context. The first is intended at reducing the effect of data uncertainty through a filtration process. The second involves a careful modulation of the protected zone width to achieve the desired false and missed detection levels.

1 Introduction

Runway incursions have been a major concern in the aviation industry for several years and this has resulted in numerous initiatives to mitigate the risk of incursion, within both Europe and the United States. These range from awareness campaigns and improved signage to air traffic control alerting tools that are used in order to advise the air traffic controller of a runway incursion. Whilst these initiatives have undoubtedly provided improved levels of safety and contributed to the reduction of the numbers of incidents, up to now, they have not provided satisfactory mitigation to the problem. This is evidenced by the fact that, to date, the United States National Transportation Safety Board still keeps the topic of runway incursions in its top ten most wanted safety improvements list.

There have been a number of programmes dealing with the design of systems for airfield surveillance with the capability of detecting conflicting situations on-board the aircraft and generating a corresponding alert in the cockpit [1-3]. One of the biggest challenges, however, is to reliably detect a runway conflict, given the uncertainties inherent to the ownship and traffic data. Although the general technique to runway conflict detection available in the literature has been shown to be valid, this paper addresses a study carried out to determine the impact data uncertainty has on the positioning of the alerting threshold, in order to maintaining an acceptable level of false detections and subsequent alerts.

2 Runway Conflict Detection

When an aircraft is cleared to make use of a runway for take-off or landing, the designated runway is essentially reserved for sole use by that aircraft. Any other aircraft or vehicle entering this reserved area results in a runway incursion. This is similar to the concept of a 'protected zone' or no transgression zone (NTZ) surrounding an aircraft whilst airborne, which defines a region into which no other entity must enter to ensure safety of the flight. In a similar way, the protected zone of an aircraft during runway manoeuvres is, actually, the entirety of the runway, extended in length and width such that it encapsulates traffic approaching the runway for landing and those approaching the runway through taxiways (Figure 1).



Figure 1 - The protected zone described around the active runway which defines the region reserved for the ownship whilst performing a take-off or landing manoeuvre. (Not to scale).

Whilst this technique functions well with accurate aircraft positional data, it is susceptible to the generation of false alerts when positional errors such as those introduced by positional reporting uncertainties are present. For example, an aircraft holding short of a runway just before the hold short bar could trigger an alert if it transmits its position as being beyond the bar. This would constitute the generation of a false alert.

2.1 Balancing False and Missed Detections

The trade-off between the probability of false detection *P*(FD) and that of a missed detection *P*(MD) requires very careful consideration. With the intended role of such a conflict detection system being that of a safety-net function, the trade-off is biased towards maintaining a low false detection rate. The effect of false detections is most serious during take-off. Falsely alerting of a runway conflict can lead to an unnecessary rejected take-off (RTO) that could have major or hazardous consequences. This is particularly so if the RTO is initiated at high speed (close to V_1) where, apart from the risk of brake fire and tyre failure, there is only a 50% chance of bringing the aircraft to a halt within the distance allowed. In field limited runways, therefore, a high speed RTO could lead to an overrun. Consequently the *P*(FD) during the latter part of the take-off run (typically above 80 or 100*kts*) must be maintained below 10^{-7} to be in compliance with the philosophy of AMC-25.1309 in CS-25 [4]. An unnecessary low speed RTO, although not as serious as a high speed RTO, is disruptive and is of nuisance since it requires the aircraft to taxi back to the start of the runway, join the queue of aircraft waiting to depart and allow sufficient time for the brakes to cool. Such an occurrence would invariably affect scheduling and induce delays.

False alerting during landing is also undesirable. However, unlike in take-off, where an unnecessary abort will have a relatively higher impact on safety, the main effect of unnecessary go-arounds is of a nuisance level, as this would usually only result in a delay in the flight and the inconvenience of needing to, where relevant, join an airborne queue to land. Nevertheless, unnecessary alerts would still lead to loss of confidence in the system with the possibility of crews disregarding the alert when a runway conflict really occurs. A further consideration is that, in certain circumstances, an unnecessary go-around may actually compromise safety and even lead to an accident. The presence of obstacles and high terrain in the vicinity of the airfield, bad weather (in particular with the presence of thunderstorm cells), conditions of low fuel and other contributing effects may combine to

significantly raise the risk of accident following a missed approach.

To achieve a low false detection rate, the conflict boundaries need to be carefully placed so that they can handle error in state data without causing false detections at a rate higher than acceptable. However, the positioning of the boundary also affects the missed detection rate. Within the context of runway conflict detection, this results in a degradation in the alerting performance of the system, where, as the aircraft in conflict proceeds within the protected zone, a missed alert is expressed as a delay in detection. This will affect the effectiveness of the system in successfully mitigating conflicts. This however will not mean that the role of the system is not achieved. Indeed, even if, through late detection the system would be able to lead to the successful resolution of the conflict on only 50% of events, this will translate to a 100% improvement in the mitigation of the risk of runway collisions. Therefore, the trade-off between false and missed detections is that of reducing the false detection rate to an acceptably low level at the expense of late detection. This ensures that the improvement in safety introduced by the system is not compromised by false alerts and their associated unnecessary manoeuvres.

As a means to mitigate this problem two techniques have been developed. One is based on filtering the aircraft data using a Kalman filter to significantly reduce the uncertainty of position reports. This alone significantly improves the detection performance without compromising the false detection rate. The second involves withdrawing the alerting threshold sufficiently such that an acceptable false detection rate due to the residual state uncertainty is achieved. A prediction technique has also been developed, to detect whether or not an aircraft approaching the hold-short bar will not stop before crossing it. This technique, based on the use of speed and acceleration profiles, further improves the early warning of an incursion. It is relevant to note that in runway conflicts, even a one second difference can differentiate between a collision or a successful avoidance of the conflict. This underlines the impact these techniques will have on the effectiveness of the runway collision avoidance system.

3 Reducing the Impact of State Uncertainty

The data sources being considered in this study are those widely available on today's aircraft fleet. Global Positioning System (GPS) is therefore being proposed as the source of positional and velocity data for the ownship. Additionally, ADS-B is being proposed as the source of traffic data, which is expected to be mandated on all large transport aircraft by the year 2020 [5]. ADS-B in itself does not compute any state information but solely encapsulates the data obtained from the aircraft systems, primarily GPS, and transmits them over a digital link. Consequently, the uncertainties expected in ADS-B data are similar to those expected in GPS data.

The accuracy with which a GPS receiver can determine its position depends on many factors, including: errors in the satellite atomic clocks, orbital errors, ionosphere and troposphere induced propagation delays, multipath reflections and noise introduced by the receiving unit. These result in an error in the range measurement between each satellite and the user, known as the user-equivalent range error (UERE). In general, the errors from the different sources will have different statistical properties. For instance, propagation delays and orbital errors tend to vary slowly over time and appear as a bias, whilst errors induced by the receiver, vary much more rapidly. The effect of all errors over time can, however, be considered as having an approximate Gaussian distribution. This results in a scatter of the position fix. Additionally, the shape of the scatter varies depending on the relative geometry of the satellites with respect to the receiver's position when performing multilateration. This effect is known as Dilution of Precision (DOP). Satellites that are orthogonal to each other result

in the lowest DOP with the scatter in position being spherical, whilst satellites at obtuse or acute angles result in a higher DOP and an elliptically shaped scatter.

Modern GPS receivers determine the velocity through carrier phase measurements. When the carrier signal transmitted by a satellite reaches a GPS receiver, the relative motion between the satellite and the receiver causes a Doppler shift of the arrival carrier signal at the receiver side. This Doppler shift is a measure of rate of change of the range, and therefore velocity, between the GPS satellite and the receiver. This method of velocity measurement is insensitive to atmospheric disturbances. However, as in the case of the position solution, velocity measurements suffer from uncertainties due to phase measurement noise, multipath effects and inaccurate knowledge of satellite ephemeris and velocity. This results in the measured speed being corrupted with a Gaussian noise. Unlike the position solution, the error in Doppler-derived speed is almost unaffected by satellite geometry, as long as the minimum of three satellites are in view.

Traditionally, filtering of signals is performed through frequency domain analysis, by designing a filter capable of reducing or suppressing the frequency components making up the noise content in the signal. Although this technique is known to be very useful in several applications, it fails to provide adequate filtering in multi-variable signals, such as aircraft states. Alternatively, through the knowledge of the interaction between the variables and their expected noise characteristics, a model-based filter can generate enhanced state estimates by amalgamating measurements with model predictions. Such a filter is commonly known as an estimator. In the context of this application, both the ownship and traffic data are GPS-derived (and therefore exhibit noise having an approximate Gaussian distribution) and can be described by an approximately linear model using Newton's Laws of motion. The Kalman filter (estimator), therefore, offers an ideal solution to data enhancement. Specifically, it is known to be capable of providing good performance in improving state estimates, synchronisation of data streams and dealing with data outages.

The purpose of the Kalman filter is primarily that of enhancing the aircraft state through estimation. To achieve this, the Kalman filter requires a definition of the system model describing the aircraft dynamics through a set of difference equations relating the aircraft states. Using this model, the Kalman filter estimates or predicts the state of the system at the next time step. These predictions, together with measurements taken from the system described by a measurement model, are used by the filter in a weighted computation to determine the filtered output. The weighting applied between the estimated states and the measured states, controls the confidence the filter has in the two and therefore controls the response of the filter. This weighting effectively tunes the filter.

The secondary purpose of the Kalman filter in this application is that of synchronising ownship and traffic data streams. Whilst GPS-derived ownship data could be expected to be available from on-board aircraft systems at a minimum 5Hz rate, ADS-B traffic reports are available at only approximately 1.5Hz with reports from different aircraft being unsynchronised in time. Consequently, it is preferred to upsample and synchronise traffic data with that of the ownship, allowing the conflict detection algorithm to operate on a fixed 5Hz scheduler. This is possible through the estimation capability of the Kalman filter, which, during consecutive ADS-B reports (measurements), makes use of the aircraft model to provide an estimate to the traffic state.

ADS-B is also prone to data outages, with only 95% of the traffic reports arriving at the ownship within the stipulated update period. When measurements are not updated, the Kalman filter estimation capability is used to ‘coast’ over the missing data, through the use of the aircraft model to estimate the states. Naturally, this can only be performed for a short

period of time, beyond which state estimates based on the model alone are considerer unreliable in the highly dynamic environment of the runway.

3.1 Design of the Kalman Filter

The design of a Kalman filter requires the description of the system and measurement models using discrete state-space equations. The system is modelled by the following equation:

$$\mathbf{x}(k) = \mathbf{A}\mathbf{x}(k-1) + \mathbf{w}(k-1) \qquad (1.1)$$

where:

x is the state vector of an aircraft,

A is the state transition matrix relating the states at the previous time step $k-1$ to the state at the current step k (taken as constant),

w is a random variable representing the process (system) noise assumed to be white with Gaussian distribution.

Aircraft dynamics are modelled as a point mass using the standard Newton's equations of motion to describe the aircraft's position, speed and acceleration during the time interval ΔT. For constant acceleration a:

$$\begin{aligned} s(k) &= s(k-1) + v(k-1)\Delta T + a(k-1)\frac{\Delta T^2}{2} \\ v(k) &= v(k-1) + a(k-1)\Delta T \end{aligned} \qquad (1.2)$$

These equations approximate the aircraft dynamics as having a constant acceleration and whilst they are adequate when ΔT is small, they do not perform well when ΔT is significantly large to allow the acceleration to change appreciably during the time period. This is particularly the case when the aircraft performs rapid manoeuvres. To improve the model and, therefore, improve the tracking capability of the filter, the rate of change of acceleration, or jerk j must be included. This, being the third derivative of position, can be expressed by adding higher order terms to each of the Newton's equations of motion, resulting in Eq. (1.2) being modified as follows:

$$\begin{aligned} s(k) &= s(k-1) + v(k-1)\Delta T + a(k-1)\frac{\Delta T^2}{2} + j(k-1)\frac{\Delta T^3}{6} \\ v(k) &= v(k-1) + a(k-1)\Delta T + j(k-1)\frac{\Delta T^2}{2} \\ a(k) &= a(k-1) + j(k-1)\Delta T \end{aligned} \qquad (1.3)$$

As aircraft states need to be resolved into the X and Y directions, with the X-axis aligned with the length of the runway and the Y-axis along the width of the runway. Six states, therefore, are needed to define the state vector $\mathbf{x}=(p_x,p_y,v_x,v_y,a_x,a_y,j_x,j_y)$ where (p_x,p_y) are the coordinates of the aircraft position, (v_x,v_y) are the speeds, (a_x,a_y) are the accelerations and (j_x,j_y) are the jerk in the X and Y directions respectively. Eq. (1.1) can therefore be rewritten as:

$$\begin{bmatrix} p_x(k) \\ p_y(k) \\ v_x(k) \\ v_y(k) \\ a_x(k) \\ a_y(k) \\ j_x(k) \\ j_y(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta T & 0 & \frac{\Delta T^2}{2} & 0 & \frac{\Delta T^3}{6} & 0 \\ 0 & 1 & 0 & \Delta T & 0 & \frac{\Delta T^2}{2} & 0 & \frac{\Delta T^3}{6} \\ 0 & 0 & 1 & 0 & \Delta T & 0 & \frac{\Delta T^2}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 & \Delta T & 0 & \frac{\Delta T^2}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 & \Delta T & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \Delta T \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_x(k-1) \\ p_y(k-1) \\ v_x(k-1) \\ v_y(k-1) \\ a_x(k-1) \\ a_y(k-1) \\ j_x(k-1) \\ j_y(k-1) \end{bmatrix} + \begin{bmatrix} w_{p_x}(k-1) \\ w_{p_y}(k-1) \\ w_{v_x}(k-1) \\ w_{v_y}(k-1) \\ w_{a_x}(k-1) \\ w_{a_y}(k-1) \\ w_{j_x}(k-1) \\ w_{j_y}(k-1) \end{bmatrix} \qquad (1.4)$$

The process noise **w** represents the errors introduced due to incomplete modelling of the aircraft dynamics. In fact, in the form described by Eq. (1.3) the jerk j is assumed constant over the time period ΔT. This assumption is valid for aircraft accelerating or decelerating in a smooth manner, such as during taxi and take-off, where the acceleration over the short period ΔT changes slowly (and can therefore be assumed to be linear). However, in highly dynamic manoeuvres such as during a landing flare and touchdown, the assumption of constant j introduces an error. This could be compensated for by adding additional states to model the higher order derivatives, such as the rate of change of jerk, but the added complication and increased number of states does not justify the benefits in this application. Any omitted higher order derivatives of position effectively manifest themselves as process noise. Although,

this noise component is not strictly Gaussian, it is small and is therefore not expected to significantly impact the performance of the Kalman filter.

Besides the modelling errors, the Kalman filter is also required to cater for noise in the measured states. The measurement model is, therefore, given by the following state-space equation, which relates the measurements to the states, in the presence of measurement noise:

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \boldsymbol{\mu}(k) \tag{1.5}$$

where:

z is the measurement vector,

H is the measurement matrix relating the states to the measurements,

μ is a random variable representing the measurement noise assumed to be white with Gaussian distribution.

However not all of the states can be measured. In fact, the measurements available are those of position and ground speed, for both the ownship and traffic. Hence, the measurement vector could be defined by $z=(p_{xm},p_{ym},v_{xm},v_{ym})^T$ where (p_{xm},p_{ym}) and (v_{xm},v_{ym}) are the measured position and velocities in the *XY*-plane respectively. Equation (1.5) can therefore be rewritten as:

$$\begin{bmatrix} p_{xm}(k) \\ p_{ym}(k) \\ v_{xm}(k) \\ v_{ym}(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x(k) \\ p_y(k) \\ v_x(k) \\ v_y(k) \\ a_x(k) \\ a_y(k) \\ j_x(k) \\ j_y(k) \end{bmatrix} + \begin{bmatrix} \mu_{p_{xm}}(k) \\ \mu_{p_{ym}}(k) \\ \mu_{v_{xm}}(k) \\ \mu_{v_{ym}}(k) \end{bmatrix} \tag{1.6}$$

The random variables **w** and **μ** in Eqs. (1.1) and (1.5) are assumed to be independent of each other as there is no correlation between the process noise and the measurement noise. **w** and **μ** are therefore white in nature, having normal distributions such that:

$$\begin{aligned} P(\mathbf{w}) &\sim N(0,\mathbf{Q}) \\ P(\boldsymbol{\mu}) &\sim N(0,\mathbf{R}) \end{aligned} \tag{1.7}$$

where **Q** is the process noise covariance matrix and **R** is the measurement noise covariance matrix.

3.2 Tuning the Measurement Noise Covariance Matrix

The noise covariance matrix **R** describes the expected magnitude of the Gaussian noise content present in the measurements and is used by the Kalman filter to determine the relative weighting to be applied between the measurements and the model based estimates. The choice of **R**, therefore, has a large impact on the Kalman filter performance. Setting the elements in **R** larger than the actual measurement noise has the effect of making the Kalman filter less responsive to the changes in the measurements, whilst making it rely more on the model predictions. Selecting **R** smaller than the actual measurement noise would make the filter rely more on the measurements making the output of the filter noisier.

The diagonal entries of **R** are the variances of the measured variables whilst the off-diagonal entries are their covariances, describing the way the errors in measurements vary with respect to each other. The errors contained in GPS-derived position and speed are independent due to the fact that GPS speed, measured using Doppler techniques, is measured independently of the receiver's position on Earth and therefore is independent of the accuracy with which the position is determined. The covariances describing the error relation between velocity and position measurements are therefore taken to be zero.

However, the errors contained in the *X* and *Y* components of a GPS position fix are not fully independent, as they depend on the satellite geometry (described by the DOP) with which the fix was determined. At high DOP values,

where the position fix error distribution is more elliptical, the covariance between the error in the *X* and *Y* directions is significant. In contrast, when good satellite coverage is available[1], the approximately spherical error distribution (due to the low DOP value) allows the approximation that the covariance of the errors in the *X* and *Y* directions are zero.

Similarly, the errors contained in ground speed in the *X* and *Y* directions are not independent. This dependency is caused when resolving Doppler GPS speed in either direction, according to the aircraft's heading with respect to the local reference frame. Since this dependency varies with aircraft heading, the covariances of position and velocity cannot be assigned a static value. In fact, within the runway environment, aircraft are typically travelling either along the runway and therefore along the *X*-axis or entering the runway laterally along the *Y*-axis. In both these circumstances, the velocity vector is aligned with either of the axes, therefore allowing the assumption that the covariances are zero.

For this application, therefore, **R** is taken to be a diagonal matrix with the elements being the variances of the measured positions and speeds. However, it is evident that the variances, particularly in position, are time variant due to being dependent on the atmospheric conditions and satellite geometric configuration. By taking the maximum expected variance as the entries in **R**, it is assured that the filter responds adequately during the worst case scenario when the actual measurement noise is close to the maximum. However, in situations when the actual measurement noise is lower than the noise statistics described by **R**, the filter does not take full advantage of the good accuracy of the measurements and still applies the same weighing between model estimates and measurements. Although in such situations the Kalman filter output is non-optimal, the output state noise will not exceed that which occurs when the measurement noise is at the maximum. The filter could be improved further by making it adaptive to the measurement noise statistics using online tuning techniques [6].

From a study performed on quantised GPS-derived data to simulate the effect of ADS-B, traffic positional data could be expected to be reported with a typical maximum error of 9.2*m* dRMS and velocity data with 0.12*m/s* in either of the *X* and *Y* directions. Using these values, **R** becomes a diagonal matrix with:

$$\sigma^2_{p_{xm}} = \sigma^2_{p_{ym}} = 6.5^2 m^2$$
$$\sigma^2_{v_{xm}} = \sigma^2_{v_{ym}} = 0.12^2 m^2/s^2$$

3.3 Tuning the Process Noise Covariance Matrix

The choice of process noise covariance matrix **Q** also has a large impact on the filter's performance. However, quantifying the process noise could be difficult and is often subject to trial-and-error techniques when attempting to obtain the best estimate for **Q**. The noise characteristics of each state in the process are assumed to be uncorrelated, thus making the off-diagonal elements equal to zero. The diagonal entries are then the variance of the noise of each state in the process. With small values for the elements of **Q**, the filter is slow to respond to changes in the states and relies more on predictions rather than on the measurements. On the other hand, large values of **Q** indicate that the system model is less precise and the filter becomes more dependent on measurements through the increase in the filter gains. This has the effect of reducing the steady state error in the output of the filter at the cost of increased sensitivity to measurement noise. Additionally, the weighting of the elements in **Q** makes the process more sensitive to changes in one state over another. The smaller the process noise variance for a particular state, the larger the confidence the filter has in that particular

[1] Since this work deals with the open environment of the runway, the assumption of good satellite coverage, unshaded by buildings or other obstacles, is valid.

state when producing the state estimates. The typical approach is to assign all diagonal elements in **Q** with an initial value of the same weighting and then fine-tune the values manually until an adequate filtered output is achieved. The initial value is typically chosen depending on how well the system model represents the actual system dynamics. The performance of the filter for various values of **Q** is quantified by comparing the noise characteristic before and after filtering.

The dynamics modelled by the system equation (Eq. (1.4)) describe fairly well the motion of an aircraft whilst on the ground either taxiing or during take-off. In fact, despite the variation in acceleration being a quadratic function of ground speed during take-off, the system model still succeeds in estimating the states well. Following a manual tuning process of trial-and-error, setting the diagonal entries of **Q** to 0.01, is found to provide the best compromise between the filter's estimation capability and measurement tracking.

In contrast, the system model does not perform so well when the aircraft is airborne, owing to the higher number of degrees of freedom and therefore larger effect of disturbances, limiting the validity of the assumption that the jerk is constant over the period ΔT. This modelling error manifests itself as having a larger process noise in the system and therefore impacts the filter's reliability in estimation. This can be compensated for by penalizing the model by setting the elements of **Q** larger. With the largest of the model deviations expected in position and following a manual tuning process, setting $\sigma_{p_y} = 1m$, with the remainder of the diagonal entries set to 0.01, was found to give an adequate filter response with no evident steady state error.

3.4 Filter Performance

To evaluate filter performance and to allow tuning of the filter, two typical aircraft paths were simulated; one with the aircraft on the ground and one airborne. A B747-200 6-DOF model [7] was used to simulate the aircraft path. Model outputs were augmented with noise to simulate the errors in ADS-B. The purpose of these tests was to evaluate how well the filter can be expected to perform during the various phases of flight and to quantify the variation in the filter's output with respect to the actual aircraft path. This was done through a comparison of the state noise characteristic before and after filtration. The paths simulated were:

- Simulation 1 - Ground Movement Scenario: Taxi from the stand for approximately 300*m* towards the runway, commencing take-off and accelerating down the runway for 1900*m* before becoming airborne.
- Simulation 2 - Landing Scenario: Final approach (9*km* prior to the threshold) in light turbulence, touch down and deceleration to taxi speed on the runway.

Figure 2 shows the residual error in the filtered states obtained for a typical run using Simulation 1 for traffic data at 1.5*Hz* and with the Kalman filter updated at 5*Hz*. In this form, the filter acts as an event-based filter, where, during the time between ADS-B reports (events), the states are estimated (through dead reckoning) and then corrected when the next ADS-B report is received. Figure 3 shows similar results for Simulation 2.

Monte Carlo simulation conducted on the filter with over a 100 runs of Simulation 1 resulted in an average residual noise after filtering of approximately 1.1*m* in both p_x and p_y, 0.27*m/s* in v_x and 0.30*m/s* in v_y. The improvement in position is significant when compared to the simulated noise content for ADS-B of 6.5*m* RMS. The error in speed is larger than the error in the data itself, due to the uncertainty introduced by the estimation in between measurements. Similarly in Simulation 2, a residual RMS noise of 1.3*m* in p_x, 2.8*m* in p_y, 0.27*m/s* in v_x and 0.09*m/s* in v_y was achieved.

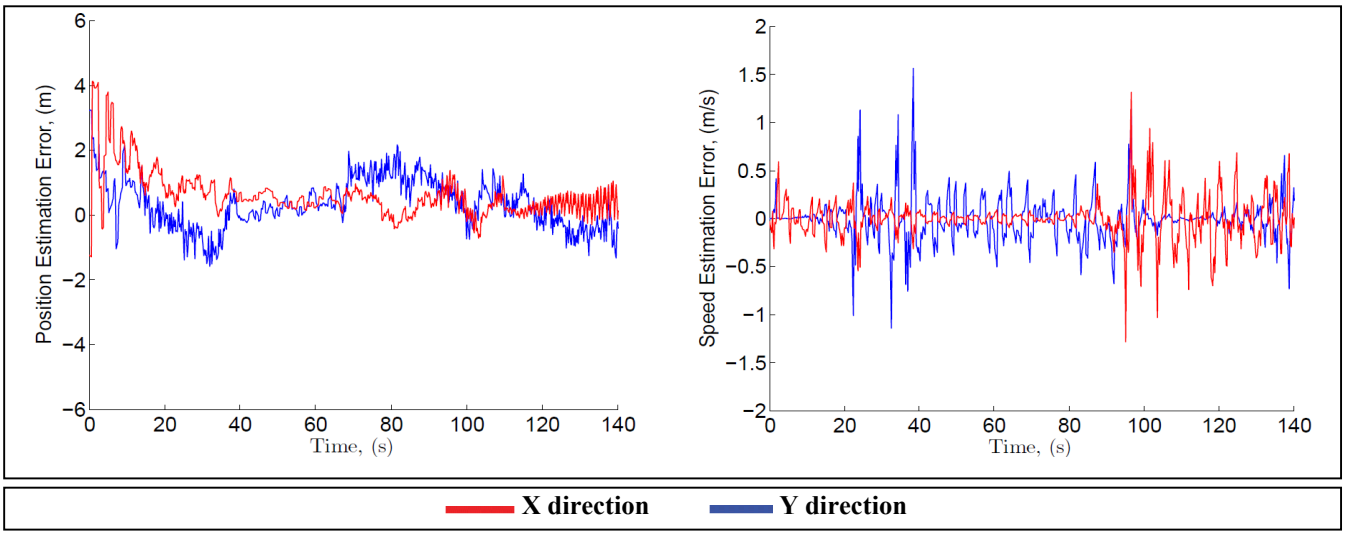


Figure 2 - Filter estimation error in position and ground speed during ground movement and take-off (Simulation 1).

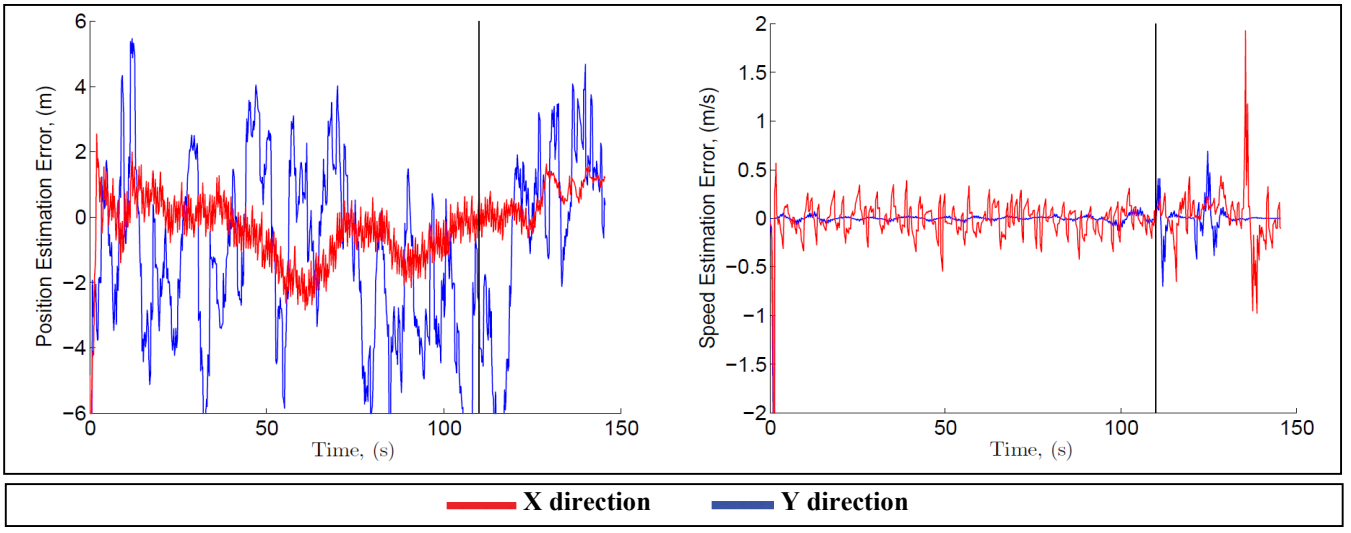


Figure 3 - Filter estimation error in position and ground speed during landing (Simulation 2)
(Vertical line indicates the braking initiation point).

4 Tuning the Conflict Detection Boundary

The width W of the protected zone affects the sensitivity of the detection algorithm to traffic intruding laterally into the protected zone (along the Y axis). When the runway is cleared for use by an aircraft in take-off or landing, procedure dictates that no traffic entity is to enter the runway and must therefore hold-short at the runway stop-bars. Choosing $W/2$ to be the distance S_S of the hold-short bars from the centreline of the runway is the theoretically ideal solution as it correctly maps the actual protection zone generated by procedure. This is a conservative approach, since when an aircraft crosses the hold-short bar, there still remains a buffer distance (the distance of the hold-short bar from the runway shoulder) before the aircraft actually enters the runway. However, choosing the zone to extend exactly to the hold-short bars may result in a high false detection rate due to the residual uncertainties in the filtered traffic positional states. False detections are expected to be generated by aircraft correctly holding at the hold-short bar but are erroneously reported as being beyond the bar.

The lateral distance $S_\perp$ between the reported traffic position and the runway centreline

exhibits noise having a Gaussian distribution, defined by:

$$S_{\perp} \sim N(s_{\perp}, \sigma_{p_y}^2) \qquad (1.8)$$

where $s_{\perp}$ is the mean perpendicular distance and $\sigma_{p_y}^2$ is the variance of the filtered traffic position in the Y direction.

Therefore, when aircraft stop exactly at the hold short bar, there is a 50% expectancy that the reported position of the aircraft falls beyond the threshold, erroneously triggering a conflict. This is inadequate and in order to reduce this probability, the threshold (width) of the protected zone needs to be displaced away from the hold-short bar as show in Figure 4. With the threshold set to $S_s - 5\sigma_{p_y}$, a false detection rate P(FD) of 2.9x10^{-7} for aircraft positioned exactly at the hold-short bar can be expected, being in compliance with the AMC-25.1309 philosophy previously explained. However in reality, aircraft tend to stop before even reaching the hold-short bar, which means that the shift in the detection boundary of $5\sigma_{p_y}$ is conservative and a lower false detection rate would actually be achieved.

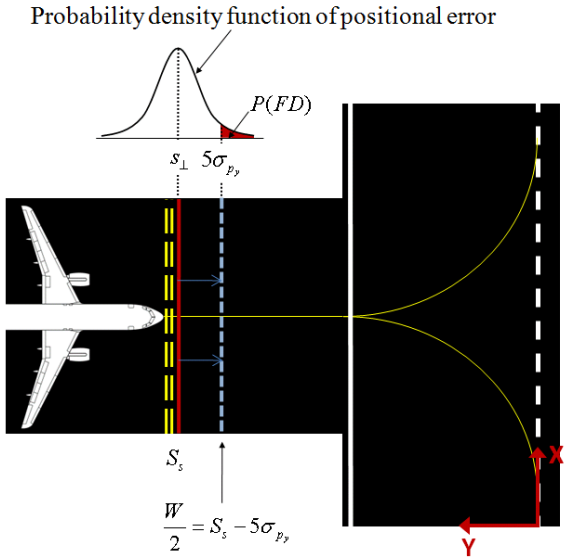


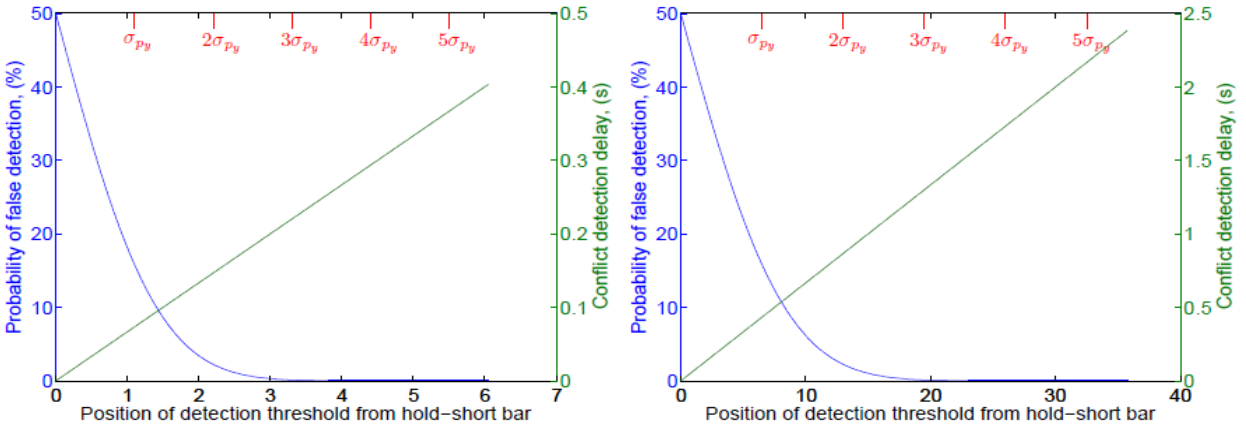
Figure 4 - Shifting in the protected zone detection boundary to maintain the level of false detections at 2.9x10^{-7}.

Shifting the detection boundary also impacts the missed detection rate. This has two effects in this application. The first is that aircraft stopping just over the hold-short bar will be outside the detection boundary and will therefore not be detected. This could be considered as an advantage, as in reality, although the aircraft will have violated procedure and entered the protected zone, it will pose no risk to the ownship in take-off or landing. In fact, a rejected take-off or landing, in this case, would be unnecessary and triggering an alert in such circumstances could be more detrimental to safety. In this respect, therefore, shifting of the detection boundary is considered advantageous.

The second effect of shifting the boundary is that of effectively delaying the alert in the classical runway incursion case where an aircraft continues onto the runway. Introducing a delay in the alerting of a runway incursion is detrimental to the effectiveness of mitigating runway conflicts. Indeed, early warning is fundamental in this application where a few seconds could make the difference in the outcome of the conflict. The delay introduced by the threshold shift is best expressed as a percentage of the leeway available from when the aircraft crosses the hold-short bar to when it arrives at the runway shoulder, which is effectively a measure of the early warning given before the physical conflict actually occurs. To quantify this numerically, a typical airfield layout can be considered, where the perpendicular distance between the hold-short bar and the runway shoulder is 50m [8]. Using a σ_{p_y} of 1.1m after Kalman filtering, the $5\sigma_{p_y}$ margin by which the protected zone boundary would be shifted is 5.5m. This gives an 11% decrease in early warning from when an aircraft is detected to have crosses the hold-short bar until it arrives at the runway shoulder. In terms of time delay, if the aircraft is assumed to be taxiing at a speed of 15m/s (30kt), the delay introduced will be of the order of 0.4s. From this analysis, the critical role of the Kalman filter becomes evident. In fact, in the case where ADS-B data is not filtered σ_{p_y} is approximately

6.5*m*, the margin by which the protected zone width would need to be reduced to ensure the same false detection rate is approximately 5-fold (32*m*). This would delay early warning by approximately 65% (time delay of 2*s*), rather than the 11% obtained with the use of the Kalman filter.

Using SPSS statistical software package, the effect of shifting the protected zone boundary on false detection was estimated for the range of shifts between $0\sigma_{p_y}$ and $6\sigma_{p_y}$. The delay introduced by the shift was also estimated, assuming the aircraft approaches the hold short bar at 15*m/s*. Figure 5 shows the result of this analysis. Figure 5a is applicable to the case were the traffic data is filtered, whilst Figure 5b is for the case where raw ADS-B measurements are used. The critical role of the Kalman filter is again made evident by these figures.



(a) When using filtered position estimates with $\sigma_{p_y} = 1.1m$.

(b) When using raw ADS-B position estimates with $\sigma_{p_y} = 6.5m$.

Figure 5 - Trade-off between the false detection rate and the delay in conflict detection, introduced by shifting the protected zone (conflict) boundary position from the hold-short bar (closer towards the runway).

4.1 Extending Early Detection

Basing the conflict detection algorithm on the presence or otherwise of traffic in the protected zone is a straight forward, effective and reliable technique for detecting runway conflicts. However, in critical situations where the conflict dynamics are marginal with little time to react, earlier detection would be beneficial towards successful conflict mitigation. A typical context is one where the ownship will be in the advanced stage of the take-off run and an intruder enters the runway such that there is insufficient distance remaining for the ownship to safely clear the conflict or to stop in time (cannot-go/cannot-stop situation). In these circumstances, early detection could result in the ownship being able to stop or, at the very least, reduce the impact speed in the case of a collision. Early detection can be achieved when the intruder approaching the hold-short bar is travelling at a relatively high taxi speed, as it will be evident that the aircraft will not stop before the hold-short bar, well before it crosses it. Indeed early detection is most useful in these cases, as at high taxi speeds, the time from crossing the hold-short bar to entering the runway will be short (approximately 3*s* for an aircraft travelling at 15*m/s* and a hold short bar positioned at 50*m* from the runway shoulder). In circumstances where the aircraft approaches the hold-short bar at low speeds, or is stationary at the hold-short bar and starts moving towards the runway, it will not be possible to predict the intentions of the aircraft with confidence any significant time before it crosses the bar. Fortunately the long time (in the order of 20*s*) until the aircraft enters the runway reduces the need for early (pre-emptive) alerting in these cases.

To take advantage of this concept of pre-emptive alerting, a prediction (or look-ahead) of the states of the aircraft approaching the hold-short bar needs to be performed. The technique is that of monitoring the aircraft speed profile as it approaches the hold-short bar and determining whether it can be expected to stop before the bar or otherwise. The challenge here is how to ascertain with confidence that the aircraft approaching the hold-short bar does not have the intention to stop. This is necessary as to keep the level of false detections at an acceptably low rate.

The aircraft deceleration d could be expressed in terms of partial derivatives given by:

$$d = \frac{\partial V_g}{\partial t} = \left(\frac{\partial s}{\partial t}\right)\left(\frac{\partial V_g}{\partial s}\right) = V_g\left(\frac{\partial V_g}{\partial s}\right) \quad (1.9)$$

where V_g is the ground speed and s is the distance travelled. At low speeds, aircraft deceleration can be assumed to be primarily affected by braking force and therefore constant with respect to V_g. The distance S_B an aircraft or vehicle will travel from an initial velocity V_B until it comes to rest will then be:

$$S_B = \int_{V_B}^{0} \frac{V_g}{d}\delta V = -\frac{V_B^2}{2d} \quad (1.10)$$

Since what is of interest is determining whether an aircraft approaching the hold-short bar can be expected to stop in time or otherwise, a speed profile defining the maximum braking that can reasonably be expected of an aircraft with the intention to stop at the hold-short bar is identified. This is shown in Figure 6.

Aircraft beyond this limit can be expected to not have the intention to stop before reaching the bar, allowing early detection of the conflict. Conversely, for aircraft in the region below this limit, it cannot be ascertained whether the aircraft will be stopping or otherwise and therefore no early warning of a conflict can be provided in such cases. The value chosen for the maximum expected deceleration d_{max} controls the sensitivity of the prediction and needs to be set to provide an adequate balance between false and missed detections.

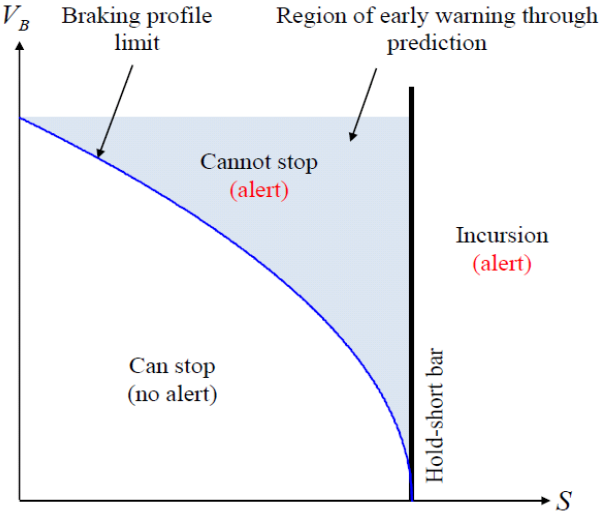


Figure 6 - The early detection braking profile limit. Aircraft above the limit cannot stop before reaching the hold-short bar and therefore an alert can be triggered earlier.

Large values of d_{max}, close to the aircraft's maximum braking capability would essentially delay the generation of an early warning. This, in effect, would ensure that aircraft approaching the hold-short bar, realising only at the latest instance that a runway incursion was about to be caused, thereby braking heavily to stop in time, would correctly not generate a false alert. However, in the normal incursion case, where aircraft would be proceeding without the intention to stop, the early warning effect is compromised as shown in Figure 7.

In this context, it is reasonable to capture aircraft with a somewhat higher deceleration rate than what would normally be expected when approaching the hold-short bar, whilst it is deemed unnecessary to cater for the event of crews becoming aware of causing a runway incursion so late as to necessitate emergency braking to stop before the hold-short bar. Although the rate of occurrence of such an event cannot be quantified statistically, the author is of the opinion that the risk of false alarms generated by such events should be acceptable.

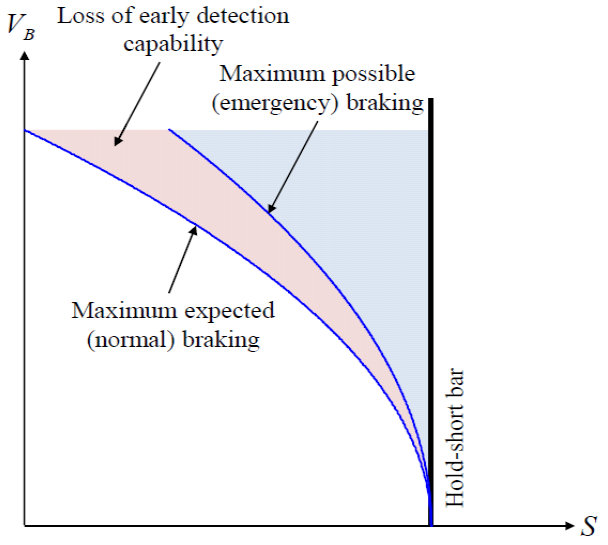


Figure 7 - Graph showing the effect of choice of maximum braking profile on early detection.

In this context, therefore, it is appropriate to identify the maximum braking d_{max} that would be expected in normal braking behaviour on approaching a hold-short bar and then add an extra leeway to mitigate the risk of false detections. In an experiment conducted by NASA to model aircraft trajectories on the airport surface [9], the maximum decelerations measured were of the order of $-0.25 m/s^2$ when travelling in a straight line. These results were consistent across different types of aircraft, including narrow, wide-bodied and regional aircraft. In fact, the experiments also showed that pilots typically apply hardest braking when decelerating from higher velocities (of the order $-0.25 m/s^2$ when at $15 m/s$), rather than when decelerating from low velocities. Therefore, setting d_{max} of $-0.5 m/s^2$ should ensure an adequate representation of the typical maximum deceleration expected when approaching a hold-short bar. This value was consequently use in this work.

4.1.1 Uncertainties due to Speed and Position Errors

The reliable prediction of whether an aircraft can stop before crossing the hold-short bar requires an analysis of the uncertainties involved in performing the calculation. This is dependent on both the uncertainty in the aircraft speed V_B as it approaches the hold-short bar, as well as on the uncertainty in its position. In fact, due to the dependency of the braking distance on the aircraft velocity, the uncertainty with which the braking distance can be determined is velocity dependent. These two effects can be depicted graphically in Figure 8.

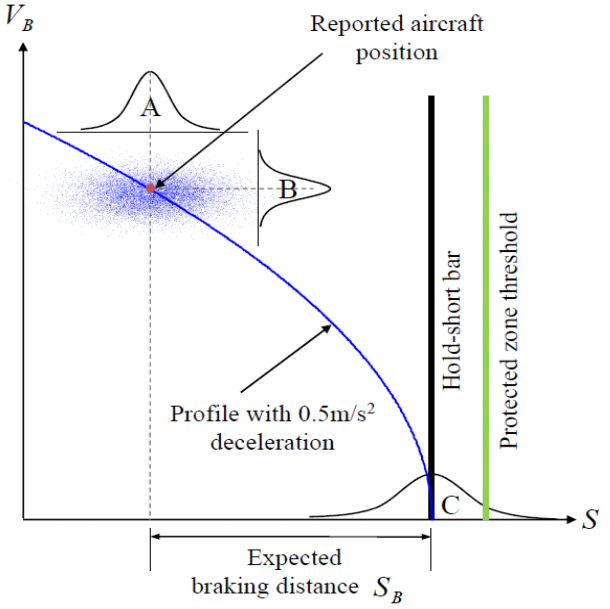


Figure 8 - The effect of uncertainties on the estimate of the braking distance. A,B and C are the probability density functions of the velocity error, positional error and stopping point respectively.

The extent of the effect of this uncertainty can be determined through analyses. Referring to Figure 8, the worst case effect occurs when the reported aircraft position is on the limit of acceptable braking that will allow it to stop at the hold-short bar. Due to uncertainties in actual aircraft position and velocity, a probability density function exists about the reported state, as shown in the figure. Approximately half the combinations of actual speed and position will clearly allow the aircraft to stop in time. However, the remainder of the combinations will result in the aircraft overrunning the hold-short bar. This is graphically expressed by the probability density function C. In order to limit the effect of false warnings, the displacement of the protected zone threshold away from the hold-short bar is advantageous, as can be seen in the figure. What is necessary here is to ensure that the threshold is beyond the 5σ tail of distribution C (Figure 8), in order to reduce the

probability of false detections to a level of the order of 10^{-7}. Clearly, the probability of an aircraft actually reporting its state to be on the limit braking profile is also low and this further reduces the overall probability of false detections due to such an event.

The shifting of the threshold away from the hold-short bar also has the effect of displacing the braking profile limit to the right as shown in Figure 9. In this context, it is then necessary to ensure that all but only a few instances in 10^7 occurrences of the density function fall within the displaced profile to ensure the target low false detection rate. The overall effect of this consideration, is once again that, of reducing the effectiveness of the early warning concept. This can be seen in Figure 9 and comparing it with Figure 8.

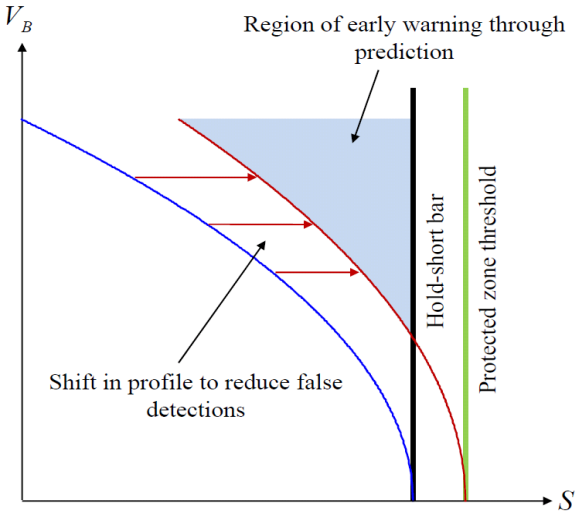


Figure 9 - Shifting of the braking limit to reduce false detections. This also has the effect of reducing the effectiveness of the early warning concept.

Following the above discussion, it is relevant to quantify the extent of the displacement in the braking profile required to ensure satisfactory operation. Referring to Figure 8, the expected braking distance of the aircraft is given by:

$$S_B = \frac{V_B^2}{2\,|\,d_{max}\,|} \qquad (1.11)$$

where V_B is the velocity at which braking starts and d_{max} is the maximum expected braking deceleration of -0.5m/s^2.

In the general case, assuming the taxiway is not perpendicular to the runway, V_B needs to be considered as the vector sum of the components in the X and Y directions. Hence,

$$S_B = \frac{V_{B_x}^2 + V_{B_y}^2}{2\,|\,d_{max}\,|} \qquad (1.12)$$

where V_{B_x} and V_{B_y} have Gaussian distributed errors with variances $\sigma_{v_x}^2$ and $\sigma_{v_y}^2$ respectively. However, the actual distance travelled by which the aircraft can be expected to come to rest from the reported position must also include the error in the reported position:

$$S_B' = \frac{V_{B_x}^2 + V_{B_y}^2}{2\,|\,d_{max}\,|} + \sqrt{\epsilon_{dx}^2 + \epsilon_{dy}^2} \qquad (1.13)$$

where ϵ_{dx} is the error in the X direction, being Gaussian distributed with σ_{p_x} and ϵ_{dy} is the error in the Y direction, being normally distributed with σ_{p_y}.

With these considerations, the variance σ_B^2 of the actual braking distance S_B' (i.e. the variance of distribution C in Figure 8) can be expressed by applying a first-order Taylor series approximation of the statistical moments [10]:

$$\sigma_B^2 = \left(\frac{\partial S_B'}{\partial V_{B_x}}\sigma_{v_x}\right)^2 + \left(\frac{\partial S_B'}{\partial V_{B_y}}\sigma_{v_y}\right)^2 + \left(\frac{\partial S_B'}{\partial \epsilon_{dx}}\sigma_{p_x}\right)^2 + \left(\frac{\partial S_B'}{\partial \epsilon_{dy}}\sigma_{p_y}\right)^2 \qquad (1.14)$$

Therefore, taking partial derivatives on Eq. (1.13) gives:

$$\sigma_B^2 = \frac{V_{B_x}^2\sigma_{v_x}^2 + V_{B_y}^2\sigma_{v_y}^2}{d_{max}^2} + \frac{\epsilon_{dx}^2\sigma_{p_x}^2 + \epsilon_{dy}^2\sigma_{p_y}^2}{\epsilon_{dx}^2 + \epsilon_{dy}^2} \qquad (1.15)$$

In the cases where the taxiways are perpendicular to the runway, the aircraft approaching the hold-short bar only travels in the Y direction, with its velocity also aligned to this direction. Eq. (1.13) will then be simplified, by setting V_{B_x} and ϵ_{dx} to zero, reducing it to:

$$S_B^{'} = \frac{V_{B_y}^2}{2\,|\,d_{max}\,|} + \epsilon_{dy} \qquad (1.16)$$

and its variance to:

$$\sigma_B^2 = \left(\frac{V_{B_y}\sigma_{v_y}}{d_{max}} \right)^2 + \sigma_{p_y}^2 \qquad (1.17)$$

As previously discussed, maintaining a low false detection rate necessitates displacing the braking profile limit towards the right, with a displacement of $5\sigma_B$ giving a false detection rate of the order of 10^{-7}. However, as described by Eq. (1.15), σ_B is speed dependent, resulting in the displacement of the braking profile also being speed dependent. A simulation was performed to demonstrate the shift in the boundary by $5\sigma_B$ using the uncertainties in position and velocity representative of those in the filtered ADS-B data for a taxiway perpendicular to the runway. Figure 10 shows the result of this simulation.

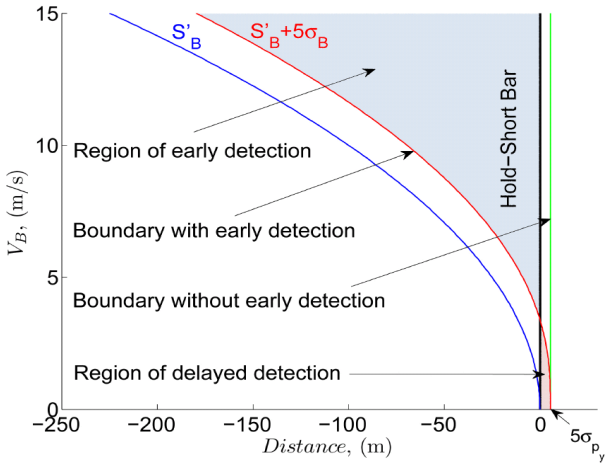


Figure 10 - Simulation of the detection boundaries with and without early detection using $\sigma_{p_x} = \sigma_{p_y} = 1.1m$, $\sigma_{v_x} = 0.27m/s$, $\sigma_{v_y} = 0.30m/s$ and $d_{max} = -0.5m/s^2$.

From this it is evident that as the speed approaches zero and therefore σ_B approaches σ_{p_y}, the shift in the threshold away from the hold-short bar becomes equivalent to that without the early detection. The figure also shows that early detection with a displaced threshold partially mitigates the disadvantage of late detections (missed warnings) brought about by the shift. Fortunately, it recovers the events involving higher taxi speeds, which, in practice form the major threat in the context. Indeed the undetectable area involving lower speeds will be associated with events where, although delayed, the detection of the runway incursion (at the displaced threshold) will still provide sufficient time for reaction before the aircraft arrives at the runway shoulder.

4.1.2 Early Detection Time

What is essentially of interest in early detection is the time from when a conflict is identified to occur to when it actually occurs; that is, how early the conflict is detected. According to the technique described above, a conflict is detected once the aircraft approaching the hold-short bar arrives to a point where it cannot be reasonably expected to stop short of it (that is, it no longer remains within the typical maximum deceleration envelope). In such circumstances the aircraft will be expected to cause a runway incursion and therefore it is reasonable to expect it to be taxiing at a constant speed towards the hold short bar. With this assumption, the early detection time t_B is defined by the distance between the early detection point and the hold-short bar, divided by the aircraft speed. The former can be derived from Figure 11, where the intersection between the braking profile boundary and the aircraft profile is the point where it is identified that a conflict would occur and, therefore, the early detection could be triggered. t_B can, therefore, be defined by Eq. (1.18), where $S_B^{'} + 5\sigma_B$ defines the distance from the hold-short bar at which the early detection can be triggered.

$$t_B = \frac{S_B^{'} + 5\sigma_B}{V_B} \qquad (1.18)$$

Figure 12 shows the relation between t_B and V_B with and without the $5\sigma_B$ leeway allowed to the braking profile boundary. As expected, the leeway delays the point at which it can be identified that a conflict would occur and therefore reduces the early detection time. However, even with the leeway in place, aircraft

travelling at high velocities, in the order of 15*m/s*, are expected to be identified as being in conflict up to approximately 12*s* prior to their crossing of the hold-short bar. As previously mentioned, aircraft travelling slowly in the order of 3.5*m/s*, cannot be detected to cross the hold-short bar earlier.

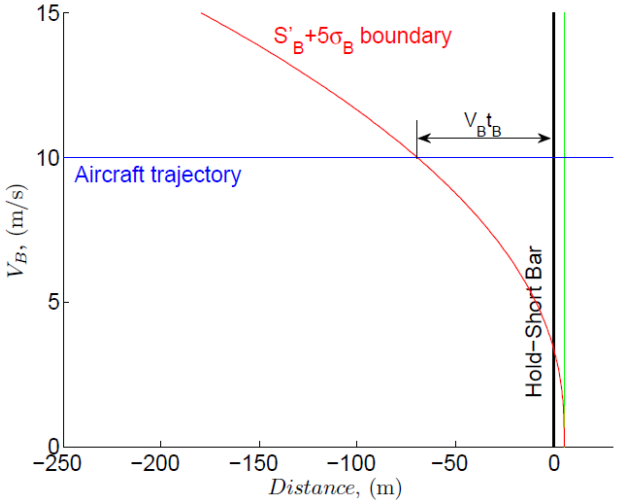


Figure 11- Intersection between the aircraft constant velocity trajectory (in this case at 10*m/s*) and the braking profile boundary. The early detection time is given by the time required for the aircraft to travel from the intersection point to the stop-bar.

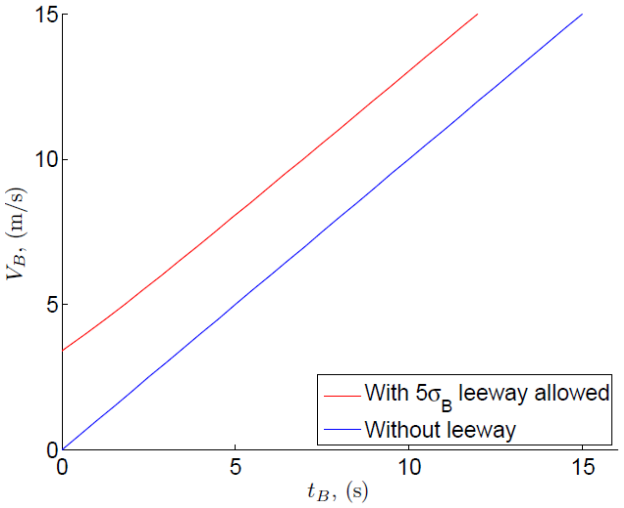


Figure 12- Variation in the early detection time with intruder approach velocity to the hold-short bar, with and without the $5\sigma_B$ leeway allowed to the braking profile boundary with d_{max}=-0.5m/s^2.

5 Discussion and Conclusion

Due to the need for high reliability in the algorithm driving runway conflict alerts, brought about by the safety critical nature of the application, considerable attention has been given to the design of the algorithm so that false and missed detection rates could be controlled. Conflict detection based on the presence or otherwise of traffic within the ownship's protected zone has shown to be a valid technique. However, the selection of the protected zone width is critical in controlling the false and missed detection rates. It has been shown that through state uncertainty propagation, the protected zone width can be tuned to achieve the required false and missed detection rates. In fact, by withdrawing the detection threshold by approximately 5*m* from the hold-short bar, a false detection rate below 10^{-7} can be achieved without introducing a significant delay in conflict detection. Although the technique for tuning the conflict boundary is valid, the actual limit allowed on false detections requires further consideration through consultancy with the stakeholders, including the certification bodies.

Triggering an alert on the presence or otherwise of traffic in the protected zone is effective. However, in critical circumstances, earlier alerting will be beneficial towards successful conflict mitigation. Therefore the concept of conflict prediction was developed, where, through monitoring of the conflicting aircraft's speed profile, it could be determined whether the aircraft will stop before overrunning the hold-short bar. This is achieved through the definition of a typical maximum braking profile, which, once exceeded, indicates that the aircraft will not stop and will consequently overrun the hold-short bar. A leeway is added to this profile to ensure that the desired low false detection rate is not exceeded.

References

[1] Green D., "Runway Safety Monitor Algorithm for Runway Incursion Detection and Alerting," NASA/CR-2002-211416, Lockheed Martin Corporation, 2002

[2] Cassell R., Evers C. and Yang E.Y., "PATHPROX - A Runway Incursion Alerting System," *The IEEE 19th Annual Digital Avionics Systems Conference*, 2000, vol. 2, pp. 7D3/1 - 7D3/6.

[3] Vernaleken C., Urvoy C. and Klingauf U., "Simulator Evaluation of a Novel Surface Movement Awareness and Alerting System (SMAAS) for Runway Incursion Avoidance," *The 26th Congress of International Council of the Aeronautical Sciences*, 2008.

[4] Anon. "Certification Specification for Large Aeroplanes," Airworthiness Code CS-25 Book 1, European Aviation Safety Agency, 2003.

[5] Anon., "Automatic Dependent Surveillance-Broadcast (ADS-B) Out Performance Requirements to Support Air Traffic Control (ATC) service," 14CFR Part 91, United States Federal Aviation Administration, 2010.

[6] Oliva A. P. and Cook M. V., "A Simulation of the Boeing B-747 Aircraft," Cranfield Institute of Technology, 1992.

[7] Myers K. and Tapley B., "Adaptive Sequential Estimation with Unknown Noise Statistics," *IEEE Transactions on Automation Control*, 520-523:21(4):520-523, 1976

[8] Anon, "Air Traffic Management," Technical Report DOC4444, International Civil Aviation Organization, 2001.

[9] Gong C., "Kinematic Airport Surface Trajectory Model Development," The 9th AIAA Aviation, Integration and Operations Conference, 2009.

[10] Coleman H. W., and Steele W. G., "Experimentation, Validation and Uncertainty Analysis for Engineers," 3rd Edition, John Wiley and Sons, Inc., 2009, pp 62-81, Appendix B.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Integration of AMAN and DMAN using Particle Swarm Optimization

C.A. Persiani,
Department of Electronics, Computer Sciences and Systems, University of Bologna, Italy

S. Bagassi
Department of Mechanical, Nuclear, Aerospace and Metallurgical Engineering, University of Bologna, Italy

Keywords: *Arrival Management, Departure Management, Particle Swarm Optimization*

Abstract

The problem of sequencing arrivals and departures on airport runways is one of the most important challenge for current and future ATM. Presently only few airports in Europe use Departure and Arrival management systems: respectively DMAN (Departure Manager) and AMAN (Arrival Manager). The First Come First Served principle is often used to define the sequence of arrivals on a runway and the sequence of departures at the holding points; Air Traffic Controllers try to insert each departure between arrival flights maintaining the appropriate separation. This results in non efficient systems from the perspective of delays, fuel consumption and workload of both pilots and controllers. The goal of this research is to optimize the sequence of runway utilization, in terms of arriving and departing aircraft, that produces the minimum average delay for the airport system. For this purpose a Particle Swarm Optimization Algorithm is designed and implemented over real air traffic data. The algorithm redefines the sequence taking into account constraints derived from a realistic shift of the aircraft position and CFMU flow restrictions. As a result, a delay is assigned to each aircraft arriving or departing from the airport: it corresponds to the delay on the start-up or on the beginning of the approach that makes it possible to optimize the traffic flow on the runways.

1 Introduction

The runways system is recognized as the bottleneck of the whole airspace transportation system; in recent years different researches and applications have been focused on the management of the surface-in and the surface-out phases of the flight. Departure Manager (DMAN) and Arrival Manager (AMAN) have been currently implemented only in few airports in Europe. At some airports, the pre-departure sequence is established using the First Come First Served (FCFS) principle: the ATC gives the start-up or delays it according to the CFMU slot or foreseeable congestion that exceeds a 20 minutes timeframe. This results in a FCFS departure sequence that can lead to long queues at the runway holding points with less than optimum traffic mix in terms of outbound routing (SID allocation) and wake vortex separation. In the same way the arriving sequence is often established according to the FCFS principle with some improvements obtained using vectoring techniques. Moreover a congested situation in the arrivals is often translated in a long and narrow approaching sequence that produces further delay in the departures.

Several research projects studied effective systems for AMAN and DMAN and some SESAR activities deal with the integration of these systems. Actually the main DMAN systems used in Europe are: PDS (Frankfurt, Munich and Zurich), EMMA2 (Prague), LEONARDO (Paris) and G2G RTS1/3 (Malmo).

Some cooperative AMAN/DMAN systems have been developed in recent years; they are based on the master/slave configuration in which the AMAN is used as a master.

Important AMAN systems and coupled AMAN/DMAN systems in Europe are OSYRIS (London, Oslo and Zurich), LEONARDO (Paris and Madrid), IBP (Amsterdam) and MAESTRO in various airports.

On the other hand many researches investigated the so called " Aircraft Sequencing (or Landing) Problem". It is a combinatorial optimization problem whose objective is to find the best sequence of arrivals or departures considering the minimum separation between aircraft of different categories. The use of evolutionary algorithms has proved to give good results in solving Aircraft Sequencing Problem.

The problem was presented first by J. E. Beasley et al. [1] through a mathematical formulation of the problem as a mixed-integer zero–one program. After relaxing binary variables and strengthening the formulation with additional constraints the problem is solved optimally with a linear programming based tree search algorithm. The formulation of the problem is presented for the single runway case and is extended to the multiple runway case.

This work was based on an earlier mixed integer linear programming formulation and a Genetic Algorithm (GA) approach by Abela et al.[2].

Cheng et al. addressed four different genetic search formulations, a broader term incorporating both GAs and Genetic Programming (GP), applied for multiple runways [3].

In [4] Ernst and Krishnamoorthy proposed an exact method based on branch and bound and a heuristic one based on GA: the model is applicable to a mix of takeoffs and landings on the same or on different runways.

In [5], V. Ciesielski and P. Scerri presented a series of experiments on landing data for Sydney Airport on the busiest day of the year to investigate the applicability of Genetic Algorithms to the problem of real time scheduling of aircraft arrival time.

H. Pinol and J.E. Beasley applied two genetic approaches to the ALP: the scatter search and bionomic algorithm for the multiple runway presenting a mathematical formulation with two types of objective functions: a linear and a nonlinear one[6].

The main objective of this research is to design a sequencing algorithm that can be used to manage arrivals and departures on the same runway. For this purpose a combinatorial optimization approach is used to model the problem and a Particle Swarm Optimization (PSO) algorithm is proposed. As Genetic Algorithms, PSO are evolutionary algorithms able to provide valid solutions in a short computing time fitting real time applications. Real air traffic data from majors Italian Airports are used to test the algorithm.

2 Coupling AMAN and DMAN: system architecture

In the master/slave configuration of current cooperative AMAN/DMAN systems, the AMAN works as a master and the DMAN fills up gaps between arrivals provided by the AMAN. In this context coordination between approach controller (which is in charge of the arrivals management) and the tower controller (which is in charge of departures management) needs to be ensured.

The main objective of the coordination is the definition of the so called arrival free interval (AFI) . It is a time slot during which no landings shall take place. AMAN takes into account this constraints and reschedules affected arrivals; DMAN reschedules the departure sequence taking into account the arrival updated sequence.

In a fully AMAN/DMAN integration the arrival and departure sequence (aircraft sequence on a runway) is built taking into account the information coming from both AMAN and DMAN.

The system architecture proposed in this paper comprises four sub-structures: a AMAN, a DMAN, a Vehicle Manager and a Runway Manager (RMAN) .

The DMAN objective is to optimize the departures flow by managing the Aircraft Off-Block time (via the start-up approval time). Departure sequence is based on minimum radar after take off, wake vortex, speed (ac-type), SID/MDI separations. DMAN is linked with the Airport CDM and receives as input the following information:

- Runway Capacity: this parameter depends on the airport characteristics and on the visibility conditions. In Low Visibility Conditions the runway capacity is reduced in accordance to local restriction.
- Target Off-Block Time: it is provided as up to date information by the airlines or ground handlers as part of the A-CDM process.
- Variable Taxi Time: in case the routing and planning function of A-SMGCS (Advanced Surveillance Movement Ground Control System) system is available, this taxi time information can be derived from the integration of the Surface Manager and the DMAN (some SESAR work packages are studying this issue).
- Aircraft type and SID (Standard Instrumental Departure): often aircraft that use diverging SID are positioned successively in the departure sequence. This makes it possible to ensure the minimum separation earlier.

The DMAN operates as follows. Once defined an appropriate Planning Horizon (PH), for example 20 minutes, that is compatible with the airport operations, the DMAN elaborates a so called Requested Departure Sequence (RDS), i.e. the sequence of departures in the CDM Off-Block time order. Between each couple of departures, the appropriate minimum separation is imposed and the related total delay is computed. Then a re-sequencing phase takes place: a genetic algorithm is used to find a new sequence that minimizes the total delay. This algorithm takes into account operational constraints as the CFMU slots and sequence shifting constraints. It is described in the following section. As a result, a Target Take-off Time (TTOT) and a consequent Target Start-Up Approval Time (TSAT) is assigned at each departure.

Arrival sequence is established by applying vectoring, speed adjustments and holdings: the landing sequence is not necessarily established according to the planned sequence. The AMAN objective is to optimize the arrivals flow by managing the Expected Approach Time (EAT). For a given entry point in the Airport Air Space (Terminal Manoeuvring Area or Control Zone) the EAT is the estimated starting time of the final approach. The AMAN takes into account:

- The Runway Capacity
- Estimated Time Over (ETO): the time in which an aircraft is estimated overflying the entry point of the airport air space (it is provided from the Area Control Canter).
- Aircraft type and STAR (STandard ARrival): this data makes it possible to compute the approach duration.

The AMAN operates as follows. Using the same Planning Horizon of the DMAN, the AMAN elaborates the so called Requested Arrival Sequence (RDS), i.e. the sequence of arrivals in the FCFS over the entry point order. Between each couple of arrivals, the appropriate minimum separation is imposed and the related total delay is computed. Then a re-sequencing phase takes place: also in this case a genetic algorithm is used to find a new sequence that minimizes the total delay. This algorithm takes into account operational constraints as unrealistic shifting in the sequence. This constraint corresponds to the well known in literature constraint shift position. As a result a

Target Landing Time (TLT) and a consequent Expected Approach Time (EAT) is assigned at each arrival.

The Vehicle Manger is used to model controls or inspections on the runway. This operations make the runway unserviceable for a defined time interval. Bird Control Unit inspections, friction test or FOD (Foreign Objects Debris) reported require the interruption of the normal operations and consequent delay on arrivals and departures.

The Runway Manager is used to integrate the air traffic from AMAN and DMAN and to define an optimized runway sequence. The Runway Manager operates into the Optimizing Horizon (OH). It represents the last part of the Planning Horizon, i.e. the part of the Planning Horizon that could be really modified without generating excessive workload. For example in a Planning Horizon of 20 minutes, the Optimizing Horizon could be set to the last 10 minutes. In other words, the sequence in the first 10 minutes of the Planning Horizon is not changeable while the aircraft in the last 10 minutes could change their position in the sequence. This happens with an assigned probability that increases as the end of the Planning Horizon is near.

The algorithm used for the runway manager is described in the following section.

The result is the assignment to each aircraft of a runway slot that is translated into a delay in the EAT or in the TSAT.

3 Sequencing Algorithm

The TLT and the TTOT provided by AMAN and DMAN define a runway sequence based on a FCFS criterion. The PSO algorithm redefines this sequence minimizing the global delay. The algorithm works on a population of sequences. A sequence is characterized by a set of aircraft defined by the planning horizon, each aircraft has an Estimated Runway Time (ERT) that represent its slot on the runway.

The algorithm starts generating n_s sequences of aircraft that define an initial population. Each sequence is initialized by randomly generating an ERT for each aircraft. For this purpose four bounds are used. Let us define a_{max} as the maximum delay acceptable for an arrival respect to its TLT and a_{min} the maximum feasible anticipation for the same flight. The ERT is initialized by randomly generating a number between a_{min} and a_{max} . In the same way, let us define d_{max} as the maximum delay acceptable for a departure respect to its TTOT and d_{min} the maximum feasible anticipation. The ERT for that flight is initialized by a randomly generated number between d_{min} and d_{max} .

Then the fitness function value associated with each sequence is calculated. Considering the order of the aircraft defined by the ERT generated, between each couple of aircraft a minimum separation is applied. It basically depends on the wake turbulence category of the aircraft: for example two minutes between two medium aircraft (Airbus 320, Boeing 737, etc) and three minutes for an heavy aircraft (Boeing 747, Airbus 330, etc) preceding a medium one. The application of the minimum separation results in a time shift of each aircraft runway slot. Moreover also the runway inspections requested by the Vehicle Manager define a delay: these are priority operations that require a suspension in the arrivals and departures. In this way a Slot Runway Time (SRT) for each aircraft is identified. The further delay applied to each aircraft due to the minimum separation is multiplied for a cost coefficient that depends on the number of passengers, the type of aircraft and an environmental factor. The fitness function value of a sequence corresponds to the sum of the costs of each aircraft delay.

The initialization phase ends identifying the best sequence of the population, G_{best} and setting the value of the actual best sequence found, P_{best} to the value of G_{best}.

Then the iteration phase starts; it consists in repeating n_t times the following improvement procedure (Eq.1). For each aircraft of a sequence a shift in the ERT is computed as follows and the actual best sequence P_{best} in terms of fitness function is identified. If it represents an improvement respect to the global best solution found, G_{best} is updated. The shift of an ERT is computed by the sum of three components: a component in the "G_{best} direction" dG and a component in the "P_{best} direction" dP and an "inertia factor" IF which is

computed multiplying a weight w by the velocity obtained at the previous iteration (Eq.2 and 3). These components model an attraction through valid solutions represented by G_{best} and P_{best} and make it possible to define a shift in the solution space towards those solutions. The component dG is identified by multiplying two factors (one random number r_1 and one coefficient C_1 comprised between 0 and 1) by the difference (distance in the solution space) between the current ERT and the same aircraft ERT in G_{best} (Eq.4). In the same way the component dP is defined using the coefficients r_2 and C_2 (Eq.5). The weight w depends on the current iteration and on two bounds W_{up} and W_{dw}. Finally the last value of G_{best} represent the best sequence found.

$$ERT_i^{t+1} = ERT_i^t + V_i^{t+1} \quad (1)$$

$$V_i^{t+1} = IF + dP + dG \quad (2)$$

$$IF = wV_i^t \quad (3)$$

$$dG = C_2 r_2 (G_{best\,i}^{\ t} - ERT_i^t) \quad (4)$$

$$dP = C_1 r_1 (P_{best\,i}^{\ t} - ERT_i^t) \quad (5)$$

4 Simulation Results

The sequencing algorithm have been tested using real air traffic data from Milano Linate, a busy airport in the north of Italy with an average of 450 movements per day with only one main runway. Two different days during the summer 2011 have been considered. For each day 5 instances that consider an Optimizing Horizon of one hour have been proposed.

The figure below reports the Interface designed for the simulation management. For each aircraft the following data are reported: the type, the wake turbulence category, the status (Departure or Arrival), the CFMU slot, the ERT, the cost coefficient and the SRT.

Once the PSO algorithm is launched, the SRT is calculated for each flight and the cost associated to the calculated sequence is displayed onto the interface.

The interface is designed also to compare results obtained through PSO versus results obtained through other evolutionary approaches as Genetic Algorithm which are not covered in this paper.

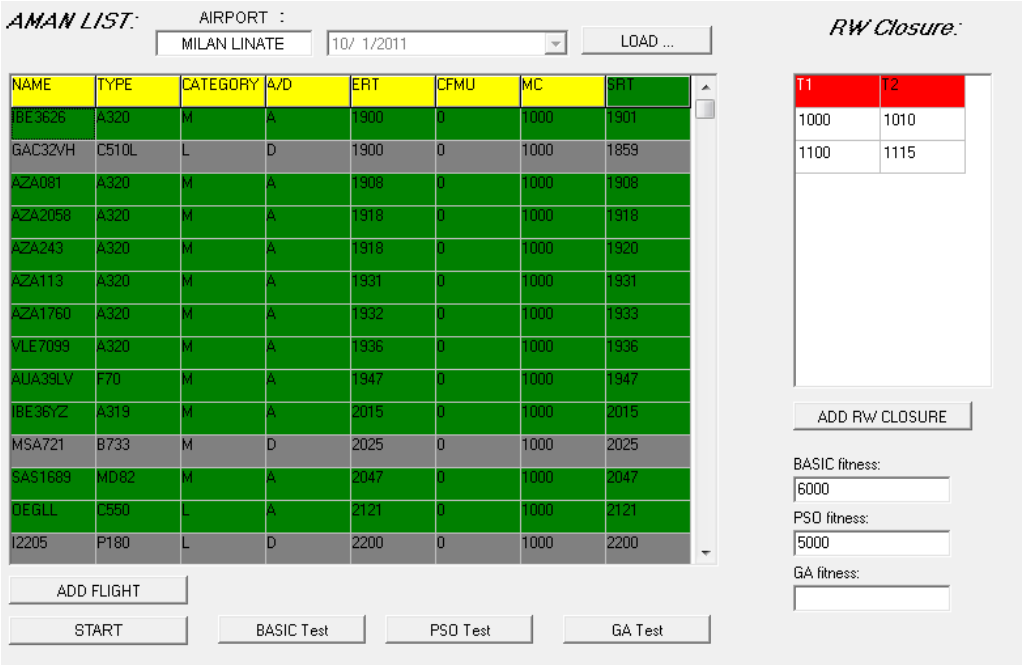


Figure 1 Screen shot of the interface used to manage the simulation. The user can load data and launch the simulation. Once the PSO algorithm is launched, the cost associated to the calculated sequence is displayed onto the interface.

For the simulation, the bounds a_{min} and a_{max} have been set to 2 and 10 minutes respectively; instead the bounds d_{min} and d_{max} have been set to 5 and 10 minutes respectively. The weight bounds W_{up} and W_{dw} have been set to 0.2 and 0.9; the coefficients *C1* and *C2* have been set to 0.6 and 0.7.

The tests have been performed on an Intel Core Due 2.00GHz Processor PC with 2.0 GB RAM.

The Vehicle Manager is simulated by randomly generating a 7 minutes long inspection on the runway during the normal operations.

In the tables below the results of the simulation are reported. For each instance the values of the fitness function (OBJ) in case of FCFS order and in case of resequencing using the PSO algorithm are reported. Moreover the last column reports the CPU time.

Table 1 Simulation results

Time slots	FCFS OBJ	PSO OBJ	CPU
14:00-15:00	5300	1700	4.21
15:00-16:00	2200	1800	4.13
16:00-17:00	6000	3800	4.58
17:00-18:00	1400	1400	3.57
18:00-19:00	2000	1600	3.56

Table 2 Simulation results

Time slots	FCFS OBJ	PSO OBJ	CPU
15:00-16:00	4300	3400	5.32
17:00-18:00	3500	2800	5.21
18:00-19:00	4600	3600	4.2
19:00-20:00	5900	3600	4.91
20:00-21:00	600	500	4.12

It is possible to notice that the proposed re-sequencing algorithm makes it possible to reduce the cost compared with the FCFS approach. Only one instance (from 17:00 to 18:00 of the first group) does not propose any improvement. This is due to the low traffic situation during that time interval. In this case, the FCFS solution remains valid and corresponds to the solution of the PSO.
Moreover, the computation time is compatible with real time applications. In this way a re-sequencing can be applied using an appropriate planning and optimizing horizon in case of consistent traffic update.

5 Conclusion

The runways capacity represents a bottleneck in the air transportation system. AMAN and DMAN integration is considered as a promising way to optimize this capacity. To overcome the actual master/slave configuration, we propose a Particle Swarm Optimization algorithm. Preliminary results show how this algorithm is able to provide valid sequences in a short computing time. In this way it is possible a fast replanning compatible with real time applications.

References

[1] Beasley J. E., Krishnamoorthy M., Sharaiha Y. M. and Abramson D. Scheduling aircraft landings – The static case, Transportation Science, 34, pp 180-197, 2000.

[2] Abela J., Abramson D., Krishnamoorthy M., De Silva A. and Mills G. Computing optimal schedules for landing aircraft , Proceeding 12th National ASOR Conference, Adelaide, Australia, pp 71-90, 1993.

[3] Cheng V.H.L., Crawford L.S. and Menon P.K. Air traffic control using genetic search techniques, IEEE International Conference on Control Applications, 1999.

[4] Ernst A.T. and Krishnamoorthy M. Algorithms for Scheduling Aircraft Landing, CSIRO Mathematical and Information Sciences Private Bag 10, Clayton South MDC, Clayton VIC 3169, Australia. 2001.

[5] Ciesielski V. and Scerri P. Real Time Genetic Scheduling of Aircraft Landing Times, In D.Fogel, editor, Proceeding of the 1998 IEEE International Conference on Evolutionary Computation (ICEC98), IEEE, New York, USA, pp 360-364, 1998.

[6] Pinol H. and Beasley J. E. Scatter Search and Bionomic Algorithms for the Aircraft Landing Problem, European Journal of Operational Research, 127(2), pp 439-462, 2006.

[7] D.W.Hutchinson, S.D.Hill, "Simulation Optimization of the Airline Delay with constraints" Proceeding of the 2001 Winter Simulation Conference, B.A. Peters, J.S Smith,D.J Mederiros and M.W.Rohrer eds.

[8] James V.Hansen "Genetic search methods in air traffic control", Computer and Operations Research 31 (2004) pages 445- 259 .

[9] Hamsa Balakrishnan and Bala Chandran "Scheduling Aircraft Landings under Constrained Position Shifting", AIAA Guidance, Navigation and Control Conference and Exhibit, 21-24 August 2006 Keystone Colorado USA.

[10] Bianco Dell'Olmo and Giordani, "Scheduling Models for Air Traffic Control in Terminal Areas", Journal of Sheduling (2005), Volume 9, Number 3, Pages 223-253 .

[11] Min Wen, "Algorithms of scheduling Aircraft Landing Problem" PhD Master Thesis, Informatics and Mathematical Modelling (2005).

[12] Ludovica Adacher, Dario Pacciarelli, Marco Pranzo, "A Graph-Theoretical Model for the Aircraft Sequencing Problem" Università degli Studi di Roma TRE (2003).

[13] A.T. Ernst, M. Krishnamoorthy and T.H Storer "Heuristic and Exact Algorithms for Scheduling Aircraft Landing", Network (1999) Pages 229-241.

[14] J.Abela, D.Abramson, M. Krishnamoorthy, De Silva, G.Mills "Computing Optimal schedules for Landing Aircraft" Proceeding of the 12th National ASOR conference 1993 Pages 71-90

[15] J.A.D Atkin, E.K.Burke, J.S. Greenwood and D.Reedson "A meta-heuristic approach to aircraft departure scheduling at London Heathrow airport", available at http://fungazi.engr.arizona.edu/caspt/atkin.pdf .

[16] G.L.Wong "The Dynamic Planner: the Sequencer, Scheduler, and Runway Allocator for Air Traffic Control Automation" , Report NASA/TM 2000-209586, NASA Ames Research Center, Moffet Field,CA,USA 2000

CEAS 2011

Future Cockpit Design

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Luminance image analysis for display visibility in cockpits

C. Aghemo, V. R.M. Lo Verso, A. Pellegrino, A. Fossella
Politecnico di Torino, Department of Energetics, TEBE Research Group, corso Duca degli Abruzzi, 24 – 101129 Turin, Italy, www.polito.it/tebe

Keywords: *display visibility; luminance images; display contrast*

Abstract

This work presents the development of an operative methodology for assessing the visibility and readability of different kinds of displays (emissive, reflective, transmissive; analog, digital) within cockpits of automotive vehicles or aircrafts taking into account the influence of different environmental sunlight conditions. The methodology is based on the analysis of high dynamic range luminance images and afterwards on their processing to obtain information about display contrast, lighting uniformity or glare. A procedure and a program in MatLab were specifically developed to automatically analyze and process the HDR images.

The paper describes the procedure and the tool and presents some results which were obtained for different kinds of vehicles.

1 Introduction

The science of Visual Ergonomics focuses on analyzing visibility condition and on improving visual performance. It involves visual system and body postures of the individual, environment lighting, job stress and other factors and it consists of setting up a work space in which clear and comfortable vision occurs [1]. In the aeronautical and automotive fields, Visual Ergonomics is concerned with a more efficient interaction of pilots'/derivers' vision with the task they are performing in the cockpit. In every conditions, all cockpit display visual information must be clearly and immediately perceived by pilots/drivers to allow them carry out their task in safety condition and reducing eye-strain problems. The optic and photometric features of cockpit displays can be highly different due to their construction technology (emissive, reflective, transmissive; analog, digital, see fig. 1) and their dynamic behavior as the environmental lighting conditions change. During flight as well as while driving a car, these conditions vary of several orders of magnitude, ranging from clear sky with sun in different critical positions (characterized by illuminance levels as high as 10^5 lux, see fig. 2), to cloudy sky and to nightlight sky (down to 10^{-4} lux). As the real world natural ambient illumination is both too variable and too complex for every condition to be taken into account in the design of cockpit displays, a few worst-cases scenarios were identified and standardized in literature, with reference to both daylight and nightlight conditions [2] [3]. In particular, for aircrafts each worst-case was

identified through reference illuminance values on the display surfaces and on pilots' eyes.

In this regard, a methodology to assess lighting condition and visibility was developed, based on both objective measurements taken within cockpits and on subjective interviews to pilots/drivers. The methodology as well as the preliminary results which were obtained were published by some of the authors in a previous paper [4]. In particular, three main fields were identified and addressed as contributing issues to evaluate display legibility (fig. 3): the "*man*" field is directly concerned with the workload and task properties; the "*machine*" field involves the display characteristic and position within the cockpit; the "ambient" field involves the physical quantities characterizing the environmental illumination, both internal and external with respect to the cockpit. This paper presents the development of the proposed operative methodology for assessing the visibility of cockpit displays and focuses on objective aspects concerned with display legibility in relation to day ambient worst-cases, while subjective display readability will be addressed in the near future.



Fig. 1 Examples of display lay-out and typologies used in cockpits: emissive digital display (above) and transmissive Head-Up Display (below)-



Fig. 2 Examples of light distribution over the displays for a sun-rear condition (above and in the middle) and for a sun-forward condition (below).

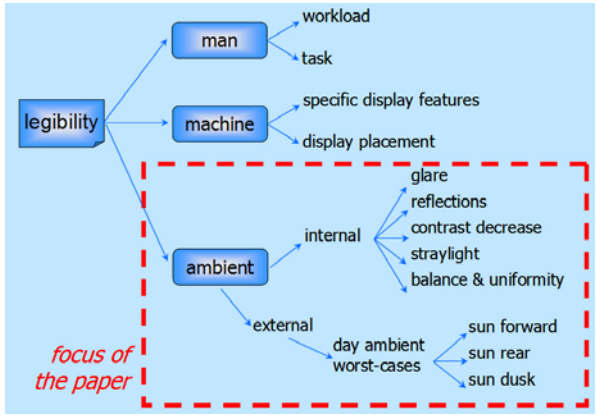


Fig. 4 Factors involved in determining the final display legibility assessment [3]: the specific focus of this paper is highlighted.

In particular, the problem of determining contrast values of whole displays or single indicators (characters, digits, labels) in presence of day ambient conditions with sunlight is addressed. In these scenario, actually, it may often happen that displays are partly in sunlight and partly in shadows (especially for a sun-rear condition), with a consequent decrease of the ideal contrast for legibility due to the veiling luminance which results overlaid to the display. Within this frame, the paper has the objectives of describing:

- a procedure proposed to calculate contrast values for sunlight conditions, through processing High Dynamic Range (HDR) luminance images
- presenting some results of contrast values calculated for different typologies of displays of different vehicles.

2 Development of a tool to automatically process luminance images

The application of the objective part of the methodology to calculate the display contrast values relies on two series of physical measurements: illuminance levels are measured on both display surfaces and on the vertical plane corresponding to pilot/drivers' eyes so as to identify how closely the reference conditions (day-ambient worst-cases) are met. For these conditions, luminance images are taken within the cockpit: an Imaging Luminance Measurement Device ILMD (TechnoTeam LMK 98-3) is used for this purpose, positioned in such a way that the lens results in the same position as the pilot/drivers' eyes.

Different series of HDR luminance images are taken, using different lenses, to measure luminance distribution with regard to legends, characters, symbols within a single display (aiming at determining luminance contrasts), to the whole display (to determine luminance uniformity) and to lay-out of displays (to determine luminance balance). Usually, wide-angle lenses, such as a fish-eye, a 4.5 mm an a 8 mm lenses, are used at first to assess the global distribution of luminances over the lay-out of displays so as to identify which areas or parts of displays are potentially the most critic for pilots for legibility. Narrow-lenses such as 25 mm and 50 mm are then used to focus on the identified areas so as to measure luminance values with a better resolution and calculate contrast, uniformity and balance indices.

The obtained luminance images are analyzed by means of a specifically developed program in MatLab®: starting from a general image of the display, potentially critical areas are identified in terms of macroscopic phenomena such as reflections, glare, shades and straylight, which can result in reduction of visual information visibility and thus in a contrast decrease. After extracting frames of the regions to analyze, the program is used to identify which pixels of the image belong to the target and which ones to the background. Luminance values are attributed accordingly to target and background areas and contrast values are then calculated through different contrast formulae available in literature [4]. The contrast values which are obtained are eventually compared to the reference values prescribed by technical standard and codes [4].

The main problem in processing images taken in sunlight condition is that displays are partly in light and partly in shadows or present areas with a veiling luminance. As a result, target and background pixels tend to have similar luminance values and do not result clearly distinguishable from each others. This makes it difficult to correctly identify target and background pixels for contrast calculations (fig. 4).

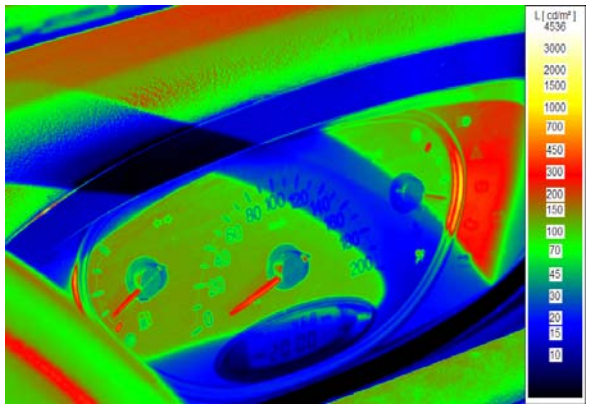


Fig. 4 Luminance image of a display in sun-rear condition: the presence of shadows and veiling luminances alter the ideal target/background contrast.

To overcome this problem, two series of measurements are taken: a first HDR luminance image is captured in controlled ambient lighting conditions for which the contrast can be easily and quickly calculated: this implies that luminance of the display targets and background are uniform over the target and background area and quite different for each other. An easy condition for this purpose is achieved when displays are exposed to a dark lighting environment. Measurement is then repeated in actual day-time conditions (after removing the curtains): a second HDR luminance image is taken, paying attention to keep the lens in the same position so as to have a new image which is perfectly corresponding to the previous one. The program applies the target/background coordinates to the new image and couples them with day-time luminance data.

As an example of the data which can be obtained by using the described procedure, the results of a first series of experiments are presented. These were carried out within cockpits of some city cars which are most largely used in Italy. These vehicles were selected so as to measure contrast indices for a variety of display lay-out (located in front of the driver or in a central position with respect to the dashboard) and typologies (analog or digital, with digits and characters of different colors, size and brightness). The experimental measurements were carried out during four consecutive days (all characterized by a sun shining condition) at the end of July 2011 in the late afternoon, in such a manner that the sun was in the typical positions of the sun-forward and sun-rear conditions.

For each scenario (sun-rear first and then sun-forward), a general luminance image was taken with the ILMD in day conditions using a 8 mm lens: together with a direct observation of the scenario, this allowed the research team identifying the most potentially critical areas in terms of contrast reduction and, thus, of legibility (fig. 5).



Fig. 5 Pictures and HDR luminance image to assess sunlight/daylight distribution over the displays and to detect most critical areas: in this case, it can be noted that all the main displays (digital and analog) are partly in direct sunlight and partly in shadows, while the navigator is in shadows.

In detail, the procedure consisted of the two following steps (figg. 6 - 8):

1. a series of HDR luminance images of the display area was taken in reference controlled lighting conditions with (fig. 6). This condition is obtained reproducing a dark environment within the cockpit by covering the windshield and all other windows with black opaque theatre curtains. The displays were set to the night-mode so as to produce highly distinguished luminances for target and background and a 25 mm lens was used to take the luminance images: these are referred to as '*reference images*'. Afterwards:
   - using the program in MatLab, from the '*reference images*' a frame, comprising a single digit or a part of a display or the entire display, is extracted and analyzed in terms of luminance spatial distribution
   - in order to identify target and background pixels, the frame is processed and a '*reference contrast image*' is generated by calculating for each the contrast value pixel, $C_{image,i}$, through the formula:

$$C_{image,i} = \left(L_{pixel,i}\right) / L_{ave,frame} \qquad (1)$$

   where $L_{pixel,i}$ is the luminance of a single pixel and $L_{ave,frame}$ is the average luminance of the whole frame, both in [cd/m^2].
   All calculated contrast values above a user defined threshold are identified as target, the ones below as background. The process goes through a number of reiterations for the pixels labeled as background only, as long as the target is clearly identified (fig. 7). The coordinated of the pixels labeled as target and as background are stored in the '*target matrix*').
   - the luminance contrast for the reference conditions C_{ref} is eventually calculated through the equation:

$$C_{ref} = \left(L_{t,ref} - L_{b,ref}\right) / L_{b,ref} \qquad (2)$$

where $L_{t,ref}$ and $L_{b,ref}$ are the average luminance of all pixels identified as target and as background respectively in *reference* conditions [cd/m^2]. Since in a dark environment the ambient light does not significantly influence the luminance contrast of the displays, the '*reference*' contrast value which is obtained can be considered as a close approximation of the ideal contrast as intended by the display manufacturer



Fig. 6 '*Reference images*' taken in the darkened cockpit and then used to identify the target/background pixels.

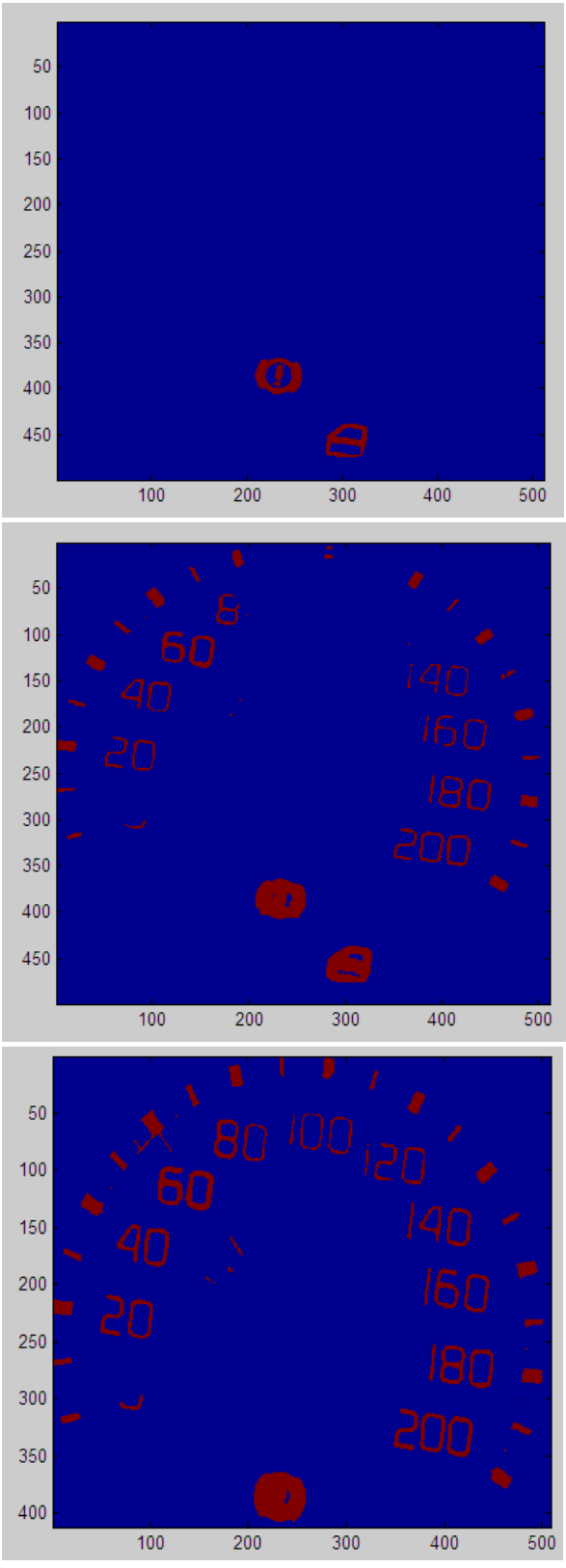


Fig. 7 Identification in the 'reference image' of pixels to be attributed to target and background through a sequence of iterations of contrast calculations. After each iteration, more pixels are attributed to the target, as long as all the digits are correctly delineated.

2. a second series of HDR luminance image is taken in actual day conditions (sun-rear and then sun-forward): these images are referred to as '*day images*' (fig. 8). The lens is kept in the same position as for the '*reference image*', so as to assure that each pixel has the same coordinates in the two images. The same frame which was extracted from the '*reference image*' is extracted from the '*day image*' and analyzed using the program in MatLab: applying the coordinates of target and background pixels stored in the '*target matrix*' (step 1), actual daytime target and background luminance values were found (fig. 9) and the following luminance contrast and uniformity values were calculated through the equations:

$$C_{day,CH} = \left(L_{t,day,CH} - L_{b,day,CH}\right) \Big/ L_{b,day,CH} \qquad (3)$$

$$C_{day,DIS} = \left(L_{t,day,DIS} - L_{b,day,DIS}\right) \Big/ L_{b,day,DIS} \qquad (4)$$

where $L_{t,day,CH}$ and $L_{b,day,CH}$ are the average luminance in *day* conditions of all pixels identified respectively as target/ background for a single character or digit (hence, with regard to a single sub-frame of the considered display); $L_{t,day,DIS}$ and $L_{b,day,DIS}$ are the average luminance in *day* conditions of all pixels identified respectively as target/background with regard to the whole frame (all in [cd/m^2]).

Comparing these contrast values to the ones observed in night conditions allows estimating the contrast reduction due to the daytime boundary conditions and to the presence of reflections, shadows and other phenomena which produce a veiling luminance on the display surfaces.

Steps 1 and 2 were repeated for all the selected frames (and sub-frames) and external scenarios (sun rear and sun forward).

3 Results

Figure 10 summarizes the results which were found for some of the analyzed vehicles.

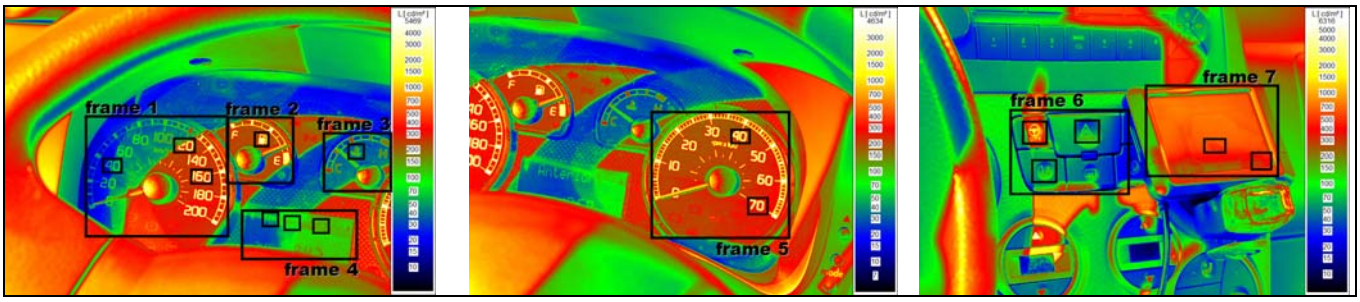


Fig 8 Luminance maps in sun-rear conditions ('day images') of the different displays (lens: 8 mm) and selections of frames and sub-frames relative to single displays or characters/digits within a display.

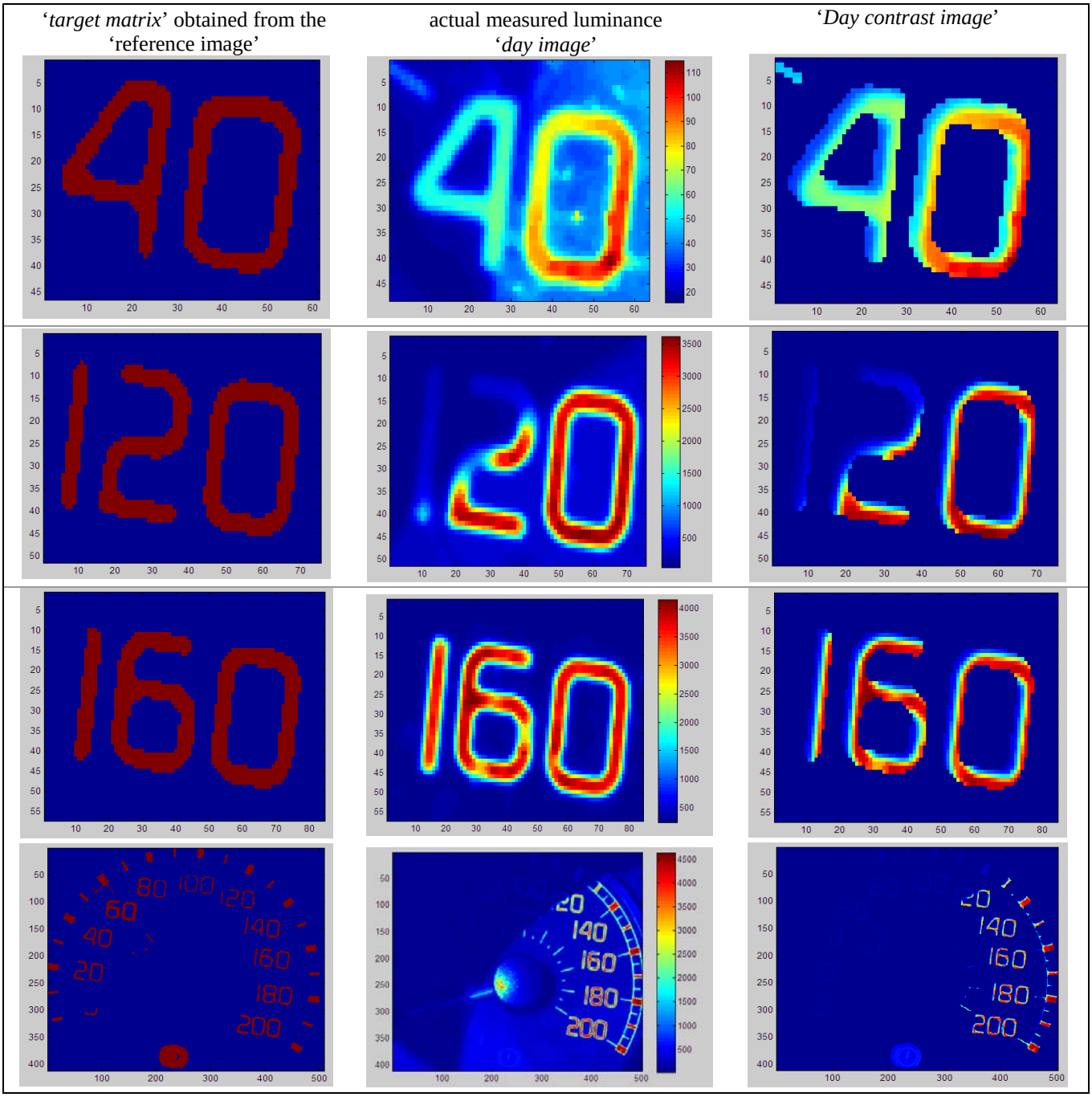
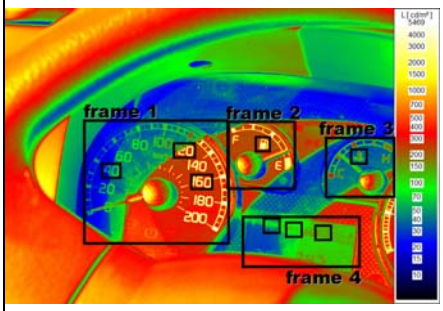
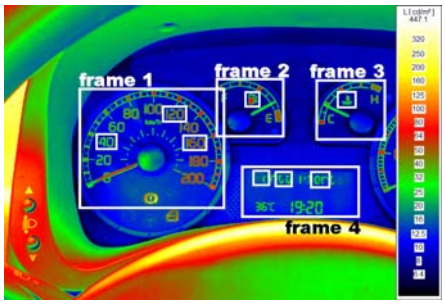
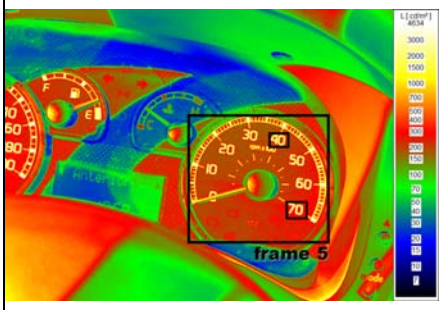
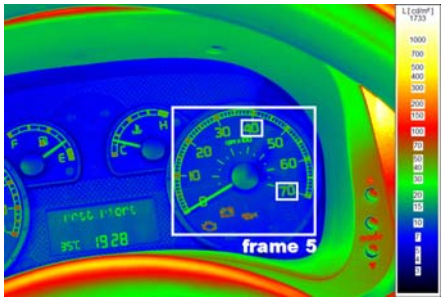
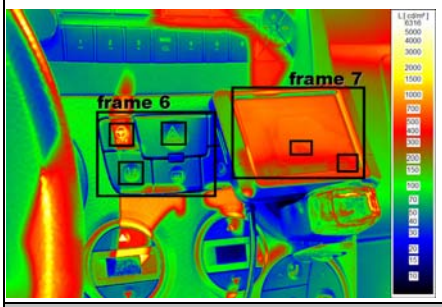
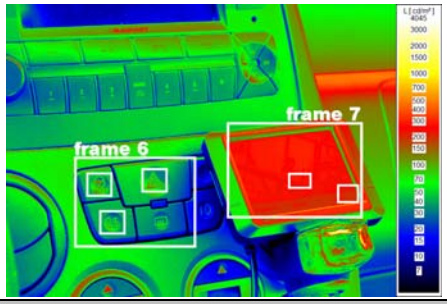


Fig. 9 Calculation of contrast Cday,CH (Eq. 3) for the single characters/digits within the considered display (i.e. sub-frames) and of contrast Cday,DIS (Eq. 4) for the whole display (i.e. frame).

VEHICLE # 1

Boundary condition: SUN REAR	Boundary condition: SUN FRONT	Contrast values	SUN REAR	SUN FRONT	REFERENCE
		$C_{day,Frame\ 1}$	3.78	3.55	23.46
		$C_{day,40}$	1.45	2.64	
		$C_{day,120}$	2.18	3.34	
		$C_{day,160}$	2.33	3.51	
		$C_{day,Frame\ 2}$	4.60	3.85	6.01
		$C_{day,Fuel}$	3.86	4.11	
		$C_{day,Frame\ 3}$	1.89	3.67	5.10
		$C_{day,Oil}$	1.99	2.97	
		$C_{day,Frame\ 4}$	0.91	1.71	6.73
		$C_{day,TE}$	1.17	1.34	
		$C_{day,OR}$	1.05	1.72	
		$C_{day,Frame\ 5}$	2.11	4.13	24.19
		$C_{day,40}$	1.34	2.91	
		$C_{day,70}$	1.39	3.04	
		$C_{day,Frame\ 6}$	4.84	3.61	2.50
		$C_{day,City}$	1.96	1.26	
		$C_{day,FogLight}$	1.21	1.18	
		$C_{day,Emergency}$	1.68	1.49	
		$C_{day,Frame\ 7}$	1.17	1.35	7.07
		$C_{day,car}$	1.07	1.12	
		$C_{day,Position}$	1.05	1.18	

VEHICLE # 2

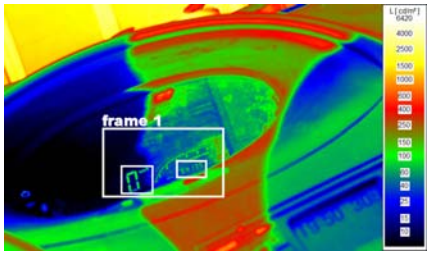
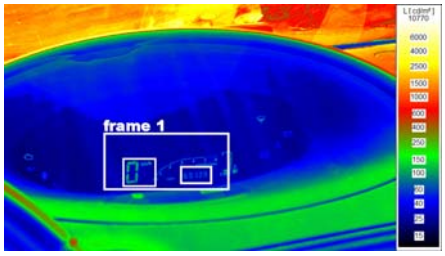
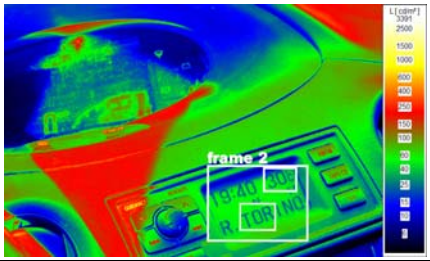
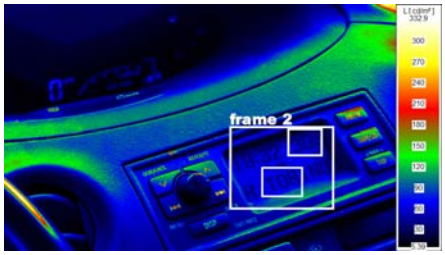
Boundary condition: SUN REAR	Boundary condition: SUN FRONT	Contrast values	SUN REAR	SUN FRONT	REFERENCE
		$C_{day,Frame\ 1}$	/	1.78	6.93
		$C_{day,0\ km/h}$	5.35	2.06	
		$C_{day,69139}$	1.56	1.19	
		$C_{day,Frame\ 2}$	1.89	1.78	44.58
		$C_{day,30°C}$	1.56	2.06	
		$C_{day,TOR}$	1.55	1.19	

Fig. 10 Summary of contrast values calculated for two of the analyzed cars.

It is interesting to point out how the elaboration of different series of luminance images allowed, as final result, calculating a contrast value, that is single number to assess the display legibility within the cockpit for pilots or drivers. This number can be compared to minimum contrast values prescribed in literature. In particular, some requirements taken from military standards can be used as reference [3] [5] [6]:

- luminance contrast C for alphanumeric information (single symbols or legends) in *analog* display:
  C > 8 [3]
  C > 11 – 21 [5]
- luminance contrast C for alphanumeric information (single symbols or legends) in *digital* display:
  C > 4 [6]

In general terms, it can be highlighted that contrast values which were found for the analyzed displays are quite low in sunlight conditions, for both analog and digital displays, compared to the values found within the darkened cockpit ('reference images'). This allows estimating the contrast reduction due to specular reflections and veiling luminance concerned with sunlight. With few exceptions (some digits of the digital display of the 'vehicle # 2'), the contrast values are lower than the required values. Furthermore, especially for the case of analog displays, the contrast values are lower in sun-rear conditions than they are for sun-forward conditions: the sun hitting directly the display area produces shadows and veiling luminances which result in contrast decrease. This effect is les noticeable for the case of digital displays.

4 Conclusion and future work

The topic of assessing the visibility of displays in aircrafts or automotive vehicle cockpits in sunlight conditions was addressed. In particular, the visibility was analyzed in terms of contrast values over the display areas and within a single display through the analysis of HDR luminance images taken in real cockpits. To overcome the difficulty of calculating the values of contrast indices in presence of veiling luminances over the displays (due to direct sunlight or to specular reflections), a procedure was proposed and a program in Matlab was developed to automatically process the luminance images and calculated the contrast values. This allowed the research team calculating the contrast values for, in principle, any typology of display and display lay-out within a cockpit under critical sunlight conditions.

A first series of results was presented, based on luminance measurements taken in vehicle cockpits. This allowed testing the procedure in-the-filed. In the next future, measurement campaigns will be carried out within cockpits of aircrafts and helicopters, both under real sun and sky conditions or within the Sky Light Simulator (SLS), an innovative ambient lighting facility achieved by Alenia Aeronautica (a leader Company in the aeronautical design, integration, testing and production) in 2006 [7] [8]: this facility is able to reproduce the photometric characteristics of daylight and sunlight (sun rear/forward), dusk/dawn transition and nightlight (starred moon/starred moonless sky) and guarantees the constancy and repeatability of reference boundary conditions. It can house real aircraft hence allowing pilots to give subjective assessment of cockpit display visibility and readability. The research will be extended so as to include subjective appraisals from pilots about the legibility of the displayed information.

As far as the objective measurements are concerned, the parameters shown in fig. 3 will be further developed: in particular, glare indices will be implemented into the program in Matlab in order to quantify the resulting annoyance for pilots and to correlate it with the subjective judgments expressed by pilots.

Acknowledgements

This work was carried out within the Regional Project E.54 "Development of a methodology for visual ergonomics within aircraft cockpits for the human-machine

interface", funded by the 'Regione Piemonte', Italy.

References

[1] Hanshel J., *Visual Ergonomics Handbook*, Taylor & Francis, Boca Raton, Florida, USA 2005

[2] DEF STAN 00-970, "Crew Stations General Requirements Guidelines for the Design of Crewstation" Lighting and Displays, Part 1, Section 4, Issue 2, Leaflet 64, 1999

[3] NATO STANAG N°3224, "Aircrew Station Lighting (Non-NVGs)", 1999

[4] Aghemo C., Pellegrino A., Lo Verso V.R.M., Petrigni C. and Rossi L., "Evaluation of display legibility through luminance mapping: an application to aircraft cockpits", 11th European Lighting Conference "Lux Europa 2009", Istanbul (Turkey), September 9-11, 2009, Turkish National Committee on Illumination, pp. 197-206, ISBN 978-975-561-352-9

[5] DOT/FAA/CT-01/08, "Human Factors Design Guide Update: A Revision to Chapter 8: Computer-Human interface Guidelines", 2001

[6] NATO STANAG N°3224, "Aircrew Station Lighting (Non-NVGs)", 1999

[7] Aghemo C., Fabbri M., Lo Verso V.R.M., Pellegrino A., Russo A. "A skylight simulator for assessing cockpit display and ambient visibility for day time and night time conditions". Proc. of the CIE International Conference, 26th edition, Beijing, China, July 4-11, 2007, publication CIE # 178:2007, ISBN: 978 3 901 906 59 6

[8] Russo A., Fabbri M. "Ambient lighting simulation for human-vehicle integration purposes: the Alenia Aeronautica Sky Light Simulator", North Atlantic Treaty Organization Human Factors and Medical Conference, Crete, Greece, 2007

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Using Constraints Analysis for future cockpit designs

Neville A. Stanton and Katie Plant
Transportation Research Group, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

Keywords: *constraints, cockpit design, cognitive work analysis*

Abstract

The aim of this paper is to report on early forays into developing new concepts for glass cockpit architecture harmonisation in helicopters, that would permit aircraft to operate in all weather conditions, to fly closer together at lower risk and to run on schedule most of the time. Low visibility in the critical phases of the flight near and on the ground is one of the most disruptive factors in European aviation today (at major airports almost 50% of arrival delays are due to low cloud and poor visibility). The aim of the ALICIA consortium project is to tackle these difficult issues using novel Human Factors methodologies. This paper presents some of the explorations to date using Cognitive Work Analysis as one of the primary investigative methods to design innovative cockpits.

1 Introduction

The work reported within this paper is the early investigations into the constraints acting on the piloting of helicopters in all weather conditions. The research is in the preliminary phases involving experience test pilots from a variety of organizations. So far, the research team has interviewed test pilots using Cognitive Work Analysis. Other sources of information have been access to simulators (including the opportunity to fly a helicopter in one of the simulators to get a participant-observer perspective) and access to flight manuals and operational procedure manuals. Eliciting knowledge from test pilots using Cognitive Work Analysis allows them to illustrate the constraints graphically. Whilst the project is still in the early stages of exploration, the methods used have enabled the research team to develop some representations of the domain and constraints.

2 Cognitive work analysis

Complex socio-technical systems are systems (such as helicopter flight operations) made up of numerous interacting parts, both human and non-human, operating in dynamic, ambiguous and safety critical domains. The complexity embodied in these systems present significant challenges for modelling and analysis. Cognitive Work Analysis is a structured framework specifically developed for considering the development and analysis of these complex socio-technical systems (Jenkins et al, 2009; Rasmussen et al, 1994; Vicente,

1999). The framework leads the analyst to consider the environment in which the tasks take place, and the effect of the imposed constraints on the way work can be conducted. The framework guides the analyst through the process of answering the question of why the system exists, what activities can be conducted within the domain as well as how these activities can be achieved and who can perform them, identifying the competencies required. The particular attraction of Cognitive Work Analysis for the ALICIA project is the claim that it is a formative approach to the analysis of complex socio-technical systems. Rather than analysing what a system currently does (descriptive modelling) or should do (normative modelling), the analysis offers a framework of methods that allow for the in-depth analysis of the properties of the work domain and the workers themselves, therefore defining a set of boundaries that shape activity within system. The approach leads the analysis into describing how the system could perform (formative modelling) given its' constraints. The analysis comprises five phases; work domain analysis, control task analysis, strategies analysis, social organisation and co-operation analysis and worker competencies analysis.
Vicente (1999) offers a description of the framework including abstraction hierarchies, decision-ladders, information flow maps and the Skills-Rules-Knowledge (SRK) framework. The technique is divided into five phases each focusing on different constraint sets, the phases names are listed on the left-hand side in Figure two. In the centre of the figure, is some guidance on the possible acquisition methods for these phases. The different forms of representation are listed on the right-hand side of Figure two.

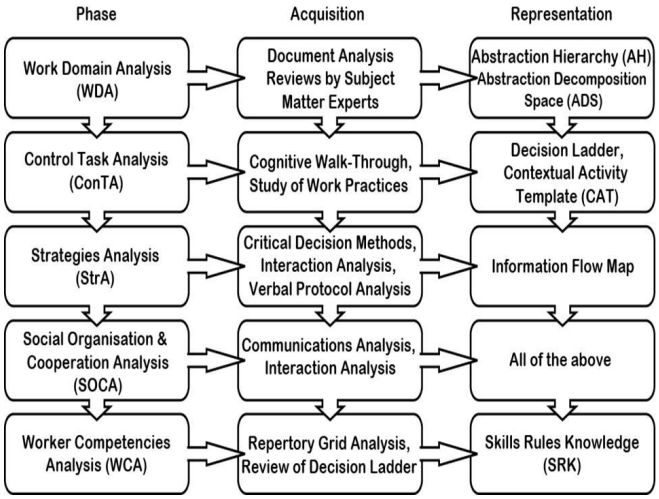


Figure 2. The five phases of CWA (Jenkins et al, 2009).

The CWA process is often criticised for being complex and time consuming. To address these concerns, and in an attempt to provide some level of guidance and expedite the documentation process, the HFI-DTC (Human Factors Integration Defence Technology Centre; www.hfidtc.com) has developed a CWA software tool. Built to run on the Microsoft .net framework, the tool uses a familiar windows-based interface and interfaces directly with 'Microsoft Office' applications such as 'Word'. The tool presents templates to allow documents to be created that describe each of the five phases described in the CWA framework, providing a structure for those unfamiliar with the technique (Jenkins et al, 2009). The software tool allows data to be passed between these phases, expediting the documentation process, facilitating updates and changes. Thus, the tool provides structure to the analysis process and markedly expedites the documentation and presentation of the analysis results. Figure three shows the phases, methods and representation methods used in the present paper. These phases were selected because of the focus on defining system functions and allocation of those functions to roles.

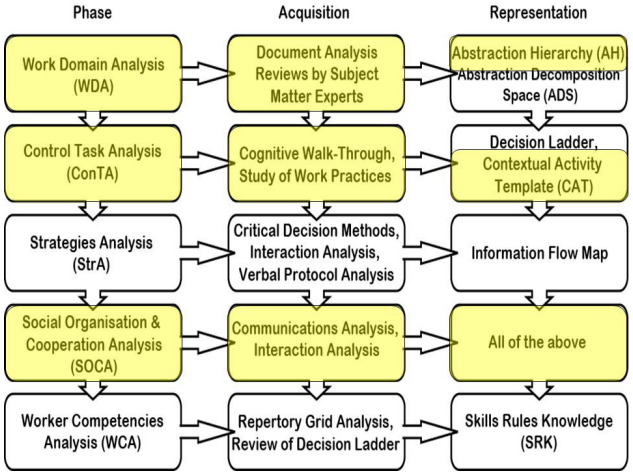


Figure 3. The phases of CWA addressed in the present paper (shaded).

The remainder of the paper presents the analysis of helicopter flying through WDA, CAT and SOCA-CAT. This analysis builds on the information gathered through interviews, review of documentation, simulators, participant observation and the rich pictures. The intention is to develop a generic description of activity across all phases of flight and it is assumed that differences between helicopter types are primarily differences in kind and relate to the ways in which different tasks are assigned to different actors or different contextual demands influence task performance in different ways.

3 Work Domain Analysis (WDA)

WDA identifies the constraints on workers' behaviour that are imposed by the purposive and physical context, or problem space, in which workers operate (Naikar, 2006a). WDA is conducted at the functional, rather than behavioural level; it is used to define the environment within which the activity is conducted. WDA identifies a fundamental set of constraints on the actions of any system component, thus providing a solid foundation for subsequent phases. WDA is typically presented in the form of an Abstraction Hierarchy (see figure four). The abstraction component of the diagram models the same system at a number of levels of abstraction; at the highest level the overall functional purpose of the system is considered, at the lowest level the individual components within the system are described.

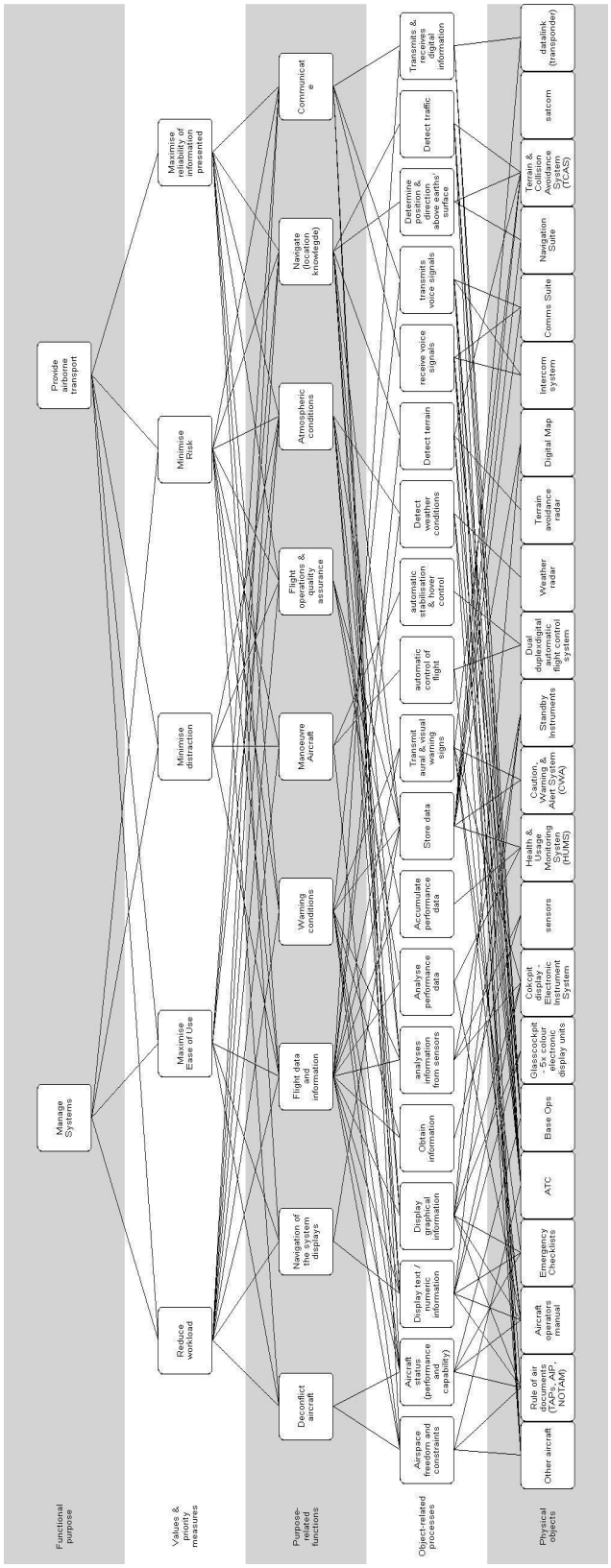


Figure 4. Work Domain Analysis showing the five levels and the means-ends-links.

Generally, five levels of abstraction are used: Functional Purposes (the purposes of the work system and the external constraints on its operation), Values and Priority Measures (the criteria that the work system uses for measuring its progress towards the functional purposes), Purpose-Related Functions (the general functions of the work system that are necessary for achieving the functional purposes), Object-Related Processes (the functional capabilities and limitations of physical objects in the work system that enable the purpose-related functions), and Physical Objects (the physical objects in the work system that afford the object related processes).

Figure four shows the domain relevant analyses in terms of Functional Purposes (e.g., manage systems and provide airborne transport), Values and Priority Measures (e.g., reduce workload, minimise distraction, minimise risk), Purpose-Related Functions (e.g., deconflict aircraft, warning conditions, manoeuvre aircraft, navigate aircraft, communicate), Object-Related Processes (e.g., display information, automatic control of flight transmit voice signals, detect traffic), and Physical Objects (e.g., emergency checklists, standby instruments, weather radar, digital map, TCAS.).

4 Control Task Analysis (ConTA)

The second phase of the CWA framework, Control Task Analysis (ConTA), allows the requirements associated with known, recurring classes of situations to be identified. Naikar (2006b) comments that, by identifying the activity that is necessary to achieve the objectives of a system with a given set of physical resources, Control Task Analysis complements WDA. The phase identifies what needs to be done independently of how it can be conducted or who can conduct it. According to Sanderson (2003), control tasks emerge from work situations and they transform inputs (e.g. current state, targets, etc.) into outputs (decisions, control actions, etc.). Returning to the definition by Rasmussen et al (1994), the ConTA phase can be considered as both, 'activity analysis in work domain terms' and 'activity analysis in decision making terms'. Typically, the work situations are shown along the horizontal axis and the work functions are shown along the vertical axis of the Contextual Activity Template (CAT). The circles indicate the work functions with the bars showing the extent of the table in which the activity typically occurs. The dotted boxes around each circle indicate all of the work situations in which a work function can occur (as opposed to must occur); thus, capturing the constraints of the system. Figure five shows the CAT for the purpose related functions taken from the abstraction hierarchy (see Figure four). The situations were derived with the assistance of the subject matter experts and the development of the process model reported in section one of this report. The situations follow the main phases of a flight from pre-flight to post shut-down. These phases (called situations in CWA parlance) are listed as follows: Preflight, Take-off, Transit, Operational, Dealing with malfunctions, Landing, Shut-down and Post shut-down.

5 Social Organisation and Cooperation Analysis (SOCA)

Social Organisation & Cooperation Analysis (SOCA) addresses the constraints imposed by organisational structures or specific agent roles and definitions. SOCA investigates the division of task between the resources and looks at how the team communicates and cooperates. The objective is to determine how the social and technical factors in a socio-technical system can work together in a way that enhances the performance of the system as a whole. SOCA recognises that organisational structures in many systems are generated on-line and in real time by multiple, cooperating agents responding to the local context (e.g. Beuscart, 2005). In the words of socio-technical theory, this would be a demonstration of the autonomy

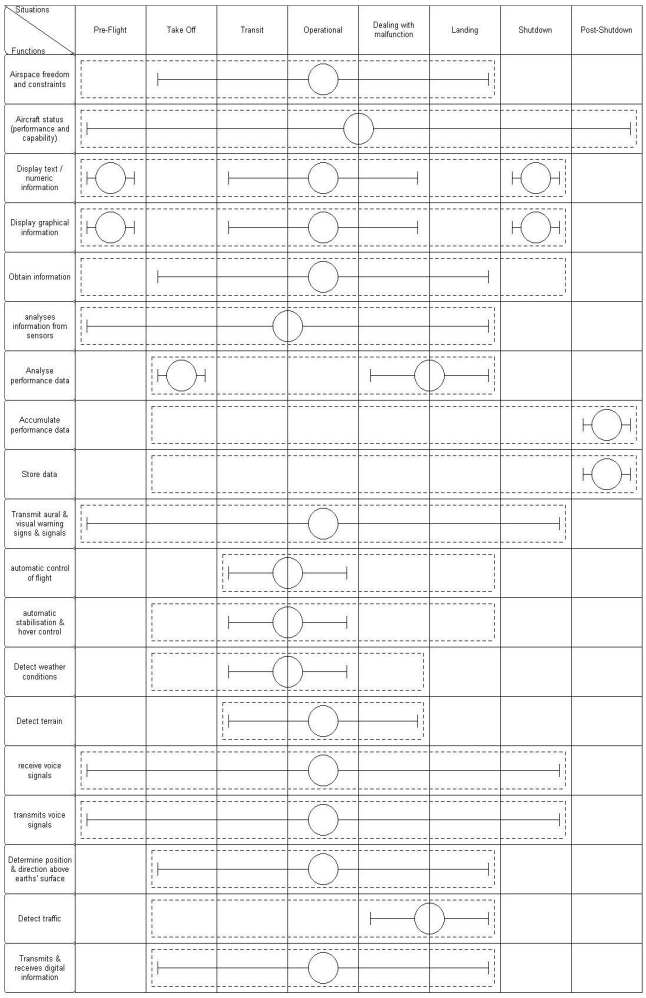


Figure 5. Contextual Activity Template for object-related processes

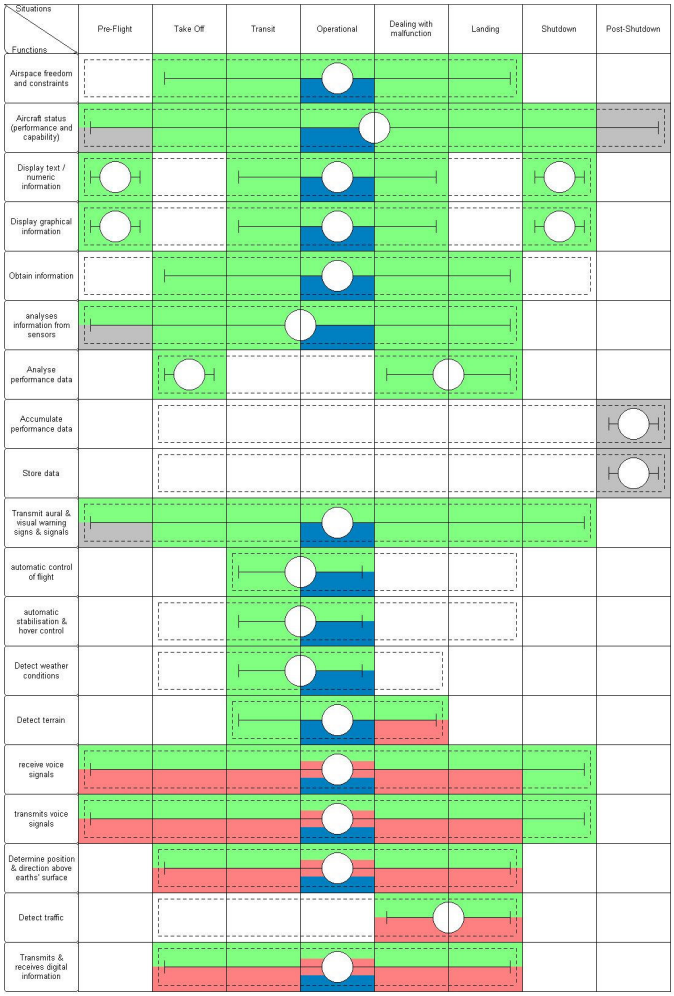


Figure 6. SOCA-CAT for phases of flight

granted to groups and the freedom members of a group have to regulate their own internal states, relating themselves to the wider system. It is not necessarily concerned with planning upfront the nature of organisational structures that should be adopted in different situations. It is instead concerned with identifying the set of possibilities for work allocation, distribution and social organisation. SOCA explicitly aims to support flexibility and adaptation in organisations (the socio-technical principle of 'equifinality'; Bertalanffy, 1950) by developing designs that are tailored to the requirements of the various possibilities (the socio-technical principle of 'multi-functionality; Cherns, 1987). The first stage of the process is to define the key agent roles in the system: engineer (grey), pilot and co-pilot (green), operations (blue), and air traffic control (pink).

6 Conclusions

Cognitive Work Analysis has particular applicability to inform the design of systems for helicopter operations in this project, as they allow the modelling of 'first of a kind' systems. The Work Domain Analysis phase in particular, leads the designer to focus, in a structured way, on the reason for developing the system. The framework encourages the designer to work based on the constraints of the system, whilst focusing on the key criteria by which the system will be evaluated. The process of considering the functions in abstract terms allows for creative thinking and problem solving. This encourages the designer to consider the need they are addressing, rather than jumping straight in to solving the problem. The different phases and tools within the CWA framework can be

applied throughout the design life cycle. CWA applications have previously been used for purposes ranging from the design of novel systems to the analysis of operational systems. It is contended that CWA is ideally designed to cope with the levels of flexibility within complex sociotechnical systems, such as helicopter flight operations and thus hold much promise for the ALICIA project.

7 Acknowledgments

This project is supporting by funding from the European Union through the ALICIA project consortium led by Augusta Westlands.

8 References

1. Bertalanffy, L. V. (1950). The theory of open systems in physics and biology, Science, 111, 23-29.
2. Beuscart, J. M. (2005). Napster users between community and clientele: The formation and regulation of a sociotechnical group, Sociologie du Travail, 47, S1-S16.
3. Cherns, A. B. (1987). Principles of Sociotechnical Design Revisited." Human Relations, 40, (3), 153-162.
4. Checkland, P. (1981). Systems Thinking, Systems Practice. Chichester: John Wiley & Sons.
5. Checkland, P. and Scholes, J. (1990). Soft systems methodology in action. Chichester, England: John Wiley & Sons.
6. Horan, P. (2002). A new and flexible graphic organiser for IS learning: The Rich Picture. Proceedings of Informing Science conference & IT Education Conference, Cork, Ireland, (pp. 133-138). Informing Science Institute.
7. Jenkins, D. P., Stanton, N. A., Salmon, P. M. & Walker, G. H. (2009). Cognitive Work Analysis: Coping. with Complexity. Aldershot: Ashgate.
8. Naikar, N. (2006a). Beyond interface design: Further applications of cognitive work analysis. International Journal of Industrial Ergonomics, 36, 423–438
9. Naikar, N. (2006b). An Examination of the Key Concepts of the Five Phases of Cognitive Work Analysis with Examples from a Familiar System. Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting , Santa Monica, CA: HFES, 447-451
10. Sanderson, P. M., 2003, Cognitive Work Analysis across the system life-cycle: Achievements, challenges, and prospects in aviation. In P. Pfister & G. Edkins (Eds.) Aviation Resource Management (Vol 3). Aldershot, UK Ashgate.
11. Rasmussen, J., Pejtersen, A. & Goodstein, L. P., 1994, Cognitive systems engineering, New York: Wiley
12. Vicente, K. J. (1999). Cognitive work analysis: Toward safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum Associates.

Enhancing Safety using FLYSAFE's Next Generation-Integrated Surveillance System (NG-ISS), an Operational Simulation Assessment

W.F.J.A. Rouwhorst, N. de Gelder and T.H. Verhoogt
National Aerospace Laboratory NLR, The Netherlands

Keywords: *keywords list (runway incursion, integrated surveillance, data link, human factors experiment)*

Abstract

The EU project FLYSAFE held its operational simulation assessment in the period February – April of 2009. The experiment was conducted on the full flight simulator GRACE (Generic Research Aircraft Cockpit Environment) facility of NLR coupled with NLR's ATC simulator (NARSIM). The experiment looked into various human factor aspects of the Next Generation Integrated Surveillance System (NG-ISS), like integrated display content, crew workload, situation awareness (SA), head-down times, taxi times, taxi speeds and operational flight safety. This paper will deal with some novel on-ground aspects, like an Airport Moving Map and runway incursion prevention and data link. It will be shown that the NG-ISS contributed to increase operational flight safety, with improved SA and no significant (subjective) workload increase.

1 Introduction

The Integrated Project FLYSAFE (Airborne integrated systems for safety improvement, flight hazard protection and all weather operations), was a first response to the "European Aeronautics, A Vision for 2020" [1], a vision that was implemented by the Strategic Research Agendas (SRAs) [2,3,4] of ACARE, the Advisory Council for Aeronautics Research in Europe, and co-funded by the European Commission (EC) under the 6th Framework Programme.

2 FLYSAFE project and main objectives

2.1 Project FLYSAFE

The project was executed from February 2005 until August 2009, and was lead was by Thales Avionics (Toulouse, France). The consortium consisted of 36 partners in total, located in 14 countries. Furthermore, a special External Expert Advisory Group (EEAG) was broadening the number and scope of stakeholders. See more project details on website: www.eu-flysafe.org

2.2 FLYSAFE main objectives

The main project aim was to contribute to the safety challenges as defined in the Vision 2020 and via the SRA translated into the main safety goal: "*To reduce the number of accidents by 80% in the year 2020*", taking into account increasing (tripling) traffic over that same period, with the year 2000 as the baseline. FLYSAFE translated above goal into various high- and sub-level goals, of which we list here only the two first-level objectives:

1 To develop, validate and test an innovative, efficient and competitive on-board integrated surveillance system, based on European resources, and prove that it increases safety;
2 To develop, validate and test ground weather means (Weather Information Management System (WIMS)) to provide aircraft with weather safety related information and prove that they increase safety (not addressed in this paper).

As a consequence FLYSAFE addressed the design, development, implementation and evaluation of an innovative on-board integrated safety system, with:

- New systems and functions, including advanced warning and alert prioritisation, and enhanced Human-Machine Interfaces (HMI)
- Improved situation awareness
- New sensors and sensor fusion techniques
- Improved on-board weather hazard detection and data fusion with novel ground-based Weather Information Management System (WIMS) products, uplinked to the aircraft via new or existing data communication means

Resulting in a new overall concept depicted in Fig. 1.

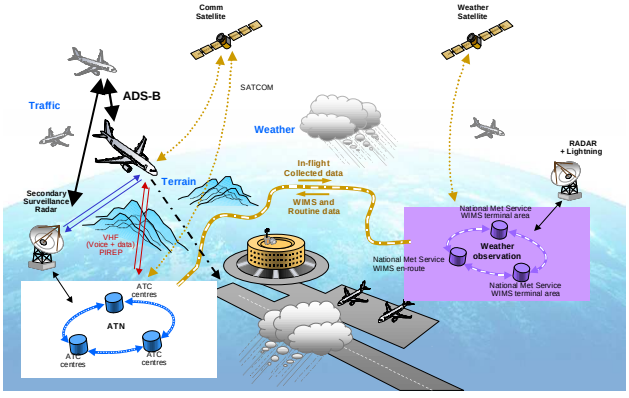


Fig. 1 FLYSAFE overall concept

3 Next Generation-Integrated Surveillance System (NG-ISS)

3.1 NG-ISS concept and functions

The new cockpit safety system concept, based on the NG-ISS concept rationale, see Fig. 2, should prevent the crew from making mistakes via its novel pro-active logics, and also support the crew after having made certain mistakes. The various NG-ISS system functions were based on an Operational Concept Description [5] and various safety studies of FAA and EASA (then JAA), see [6,7].

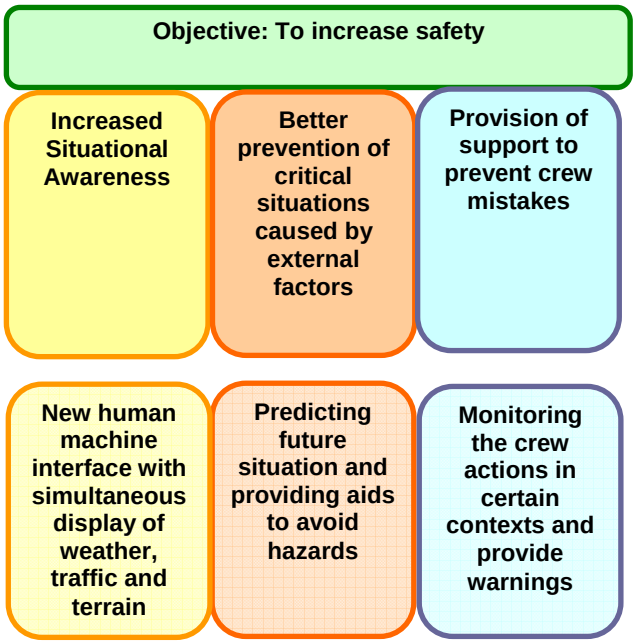


Fig. 2 NG-ISS concept rationale

It consisted of runway incursion prevention logic, weather information (fused and non-fused), Intelligent Crew Support (ICS) matters, tactical and strategic Conflict Detection (and resolution) logic to avoid more timely bad weather situations and traffic conflicts. Even certain applications were incorporated, that anticipated the future ATM related SESAR and NextGen programs.
In addition a novel Cockpit Display System (CDS) and HMI, like Airport Moving Map (AMM) on the Navigation Display (ND) with detailed airport (surface/marking) information, ATC-uplinked taxi route presentation, stop bars and other traffic, uplinked and fused weather information, was developed. All functions, as well as the Primary Flight Display (PFD)

adaptations, were controlled via special menu-driven controls, and by using a software driven Data link Control and Display Unit (DCDU).

The finally developed Next Generation Integrated Surveillance System (NG-ISS) concept can be found in Fig. 3.

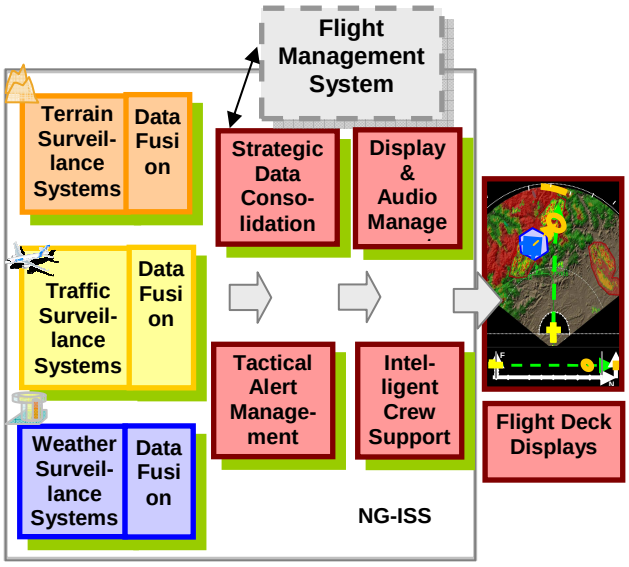


Fig. 3 FLYSAFE NG-ISS concept

The Strategic Data Consolidation (SDC) module was coupled to the Flight Management System (FMS) for flight plan exchange to the ground (i.e. to ATC). The NG-ISS also made use of tactical functions and today's safety nets solutions, like EGPWS and TCAS although the latter was enhanced (i.e. merged) with ADS-B functionalities.

3.2 On-ground crew support

Based on the most recent (2010) annual reports, from EASA [8] and the FAA (see: www.faa.gov/airports/runway_safety), it can be concluded that since about a decade major improvements have been achieved, but that the runway incursion safety problem is still far from being under control. Runway incursions continue to occur roughly 80 times a month in the USA and about 60 times a month in Europe, while it also seems not to be significantly reducing despite the international efforts made, see [9,10,11]. It is expected that some of the novel technologies developed for the FLYSAFE NG-ISS concept will contribute further to the prevention of runway incursion events. But it is also recognized that in the future novel ATC type of solutions, like Advanced Surface Movement Guidance & Control Systems (A-SMGCS), will be provided in conjunction.

This paper will focus on the NG-ISS functions set up for improving situation awareness of crews during on-ground (taxi) operations and in preventing runway incursions.

3.3 On-ground crew support functions

The taxi support functions consisted of the Intelligent Crew Support (ICS), Surface Movement Awareness and Alerting System (SMAAS) and Airport Moving Map (AMM).

The ICS system was programmed to assist the crew in dedicated use cases, like for support in:

- Approaching/crossing taxiway edges,
- Stop bar violation,
- Wrong runway detection,
- Inadvertent take-off from a taxi way,
- Take-off without an ATC clearance, and
- Incorrect altimeter settings.

The SMAAS functions, like Operational Awareness, Clearance Awareness, and Surface Movement Alerting functions, aimed to prevent runway incursions [12,13,14].

The AMM provided position awareness when navigating at an airport. It also presented various SMAAS aspects, such as open/closed/active runway presentations, FMS-selected runway indications, taxi traffic collision indications and alerting, wrong runway heading annunciation and also taxi/runway clearance awareness related aspects, see Fig. 4.

The ICS and SMAAS taxi support functions required Controller Pilot Data Link Communications (CPDLC) with ATC, for instance for ATC-uplinked clearance messages, or ATC-uplinked taxi routes.

Finally the Runway Collision Avoidance Function (RCAF) function for prevention of aircraft conflicts during take-off and landing was developed. Its HMI was also presented on the AMM, see Fig. 5. During the take-off roll the RCAF would alert "ABORT, ABORT",

both aurally and in red warning labels on the PFD, for any particular runway incursion that occurred below a speed of 100 KIAS.



Fig. 4 AMM on ND with ATC-cleared taxi route presentation (in green) and other aircraft

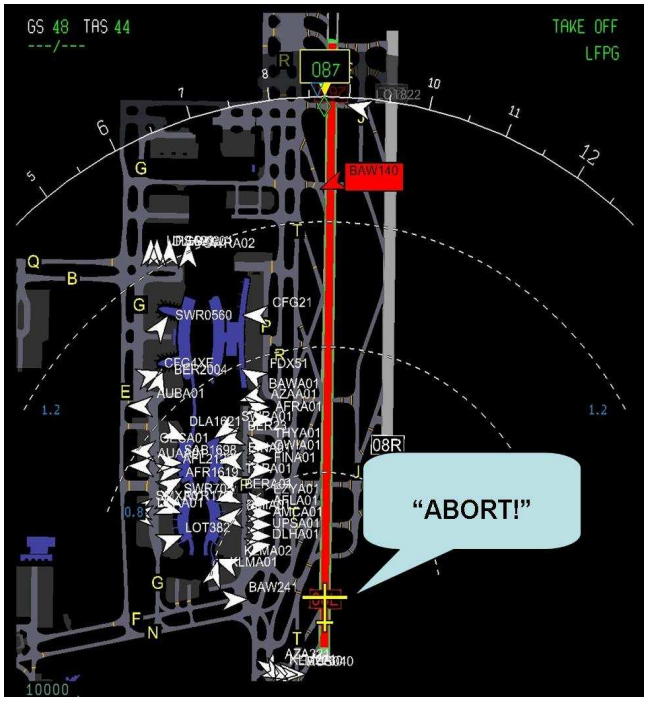


Fig. 5 RCAF runway incursion prevention alert presentation on AMM

When braking and decelerating this would be accompanied with an aural distance-to-target callout in the format "2000", "1800", etc. (meters). Furthermore, the intruder aircraft as well as the runway would be colored red.

When approaching to the landing runway the RCAF would alert for any runway obstructing aircraft using the warning "GO AROUND TRAFFIC" both aurally and in red alert labels presented onto the PFD. It is noted that this RCAF alert was given independently from any landing clearance provided by ATC (as issued via R/T), so it had no internal system coupling to it. RCAF used the ADS-B traffic information to derive the critical position of the intruder on the runway. The intruder and its flight identification were shown in red on the ND, when the alert triggered, see [15,16,17] for more RCAF concept details.

3.4 Concept integration and testing

From 2006 till 2008, partial integration, verification, validation and part-task evaluations of the foreseen novel functional improvements took place at various partners' sites, like at TU Darmstadt (Germany), BAE Systems, Rochester (UK), GTD Sistemas De Información (Barcelona, Spain), Skysoft (Portugal), and the University of Malta in conjunction with Cranfield University, see [12-17]. In parallel, a large overall pre-integration took place of all major NG-ISS functions and the associated CDS at Thales Avionics (Toulouse, France). After its ending, in mid 2008, the full NG-ISS prototype set-up moved over to the National Aerospace Laboratory NLR, Amsterdam, the Netherlands, for final integration testing, from July 2008 until January 2009.

4 NLR simulator facilities

4.1 GRACE

The NLR flight simulator is a full motion Generic Research Aircraft Cockpit Environment (GRACE). It consisted of an advanced Airbus A320-alike cockpit, with six Liquid Crystal Displays (LCDs), a Research FMS (one MCDU per pilot), two Key Cursor Control Units (KCCU) to control display features on the ND and two fold-away Jeppesen Electronic Flight Bags (EFB) for all charted matters.

Furthermore, two separate Data link Control and Display Units (DCDU) could be used for data link message control via a touch screen. See Fig. 6 and 7. The visual system used had no full cross-cockpit view.

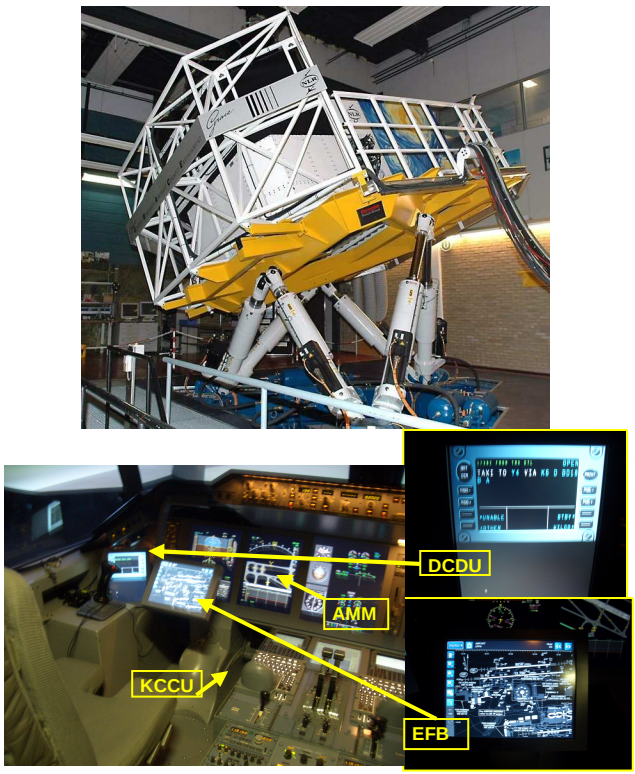


Fig. 6 GRACE simulator (top) and cockpit (bottom)



Fig. 7 DCDU with an ATC-cleared and crew-accepted taxi route instruction

4.2 NARSIM

The NLR ATC Research Simulator (NARSIM) consisted of Tower, Approach and Radar simulators. NARSIM Tower had a 360 degrees visual to show the airport surrounding and the traffic flying and taxiing around Paris Charles de Gaulle and Innsbruck airport (see Fig. 8).

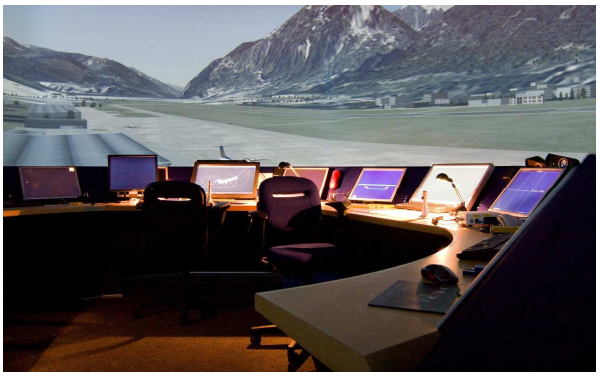


Fig. 8 NARSIM Tower incorporates positions for Runway and Ground Controllers at Innsbruck

The Controller Working Positions (CWPs) were adapted for the airport lay-outs, with controls to set the stop bars and with CPDLC for data link communications (i.e. taxi, line-up and take-off clearances).

NARSIM Approach comprised the approach and departure controller positions. The approach position was representing the TMA of Paris Charles de Gaulle. Innsbruck Approach was positioned inside the Tower simulator, which is often the case at smaller airports.

NARSIM Radar comprised several controller positions in order to have at least one position per simulated sector. NARSIM Radar was set up with ground-stations supporting all sectors between the two airports. Furthermore, CPDLC interfaces were developed in order to support future Air Traffic Management applications, e.g. the downlink of trajectories that were proposed by the SDC-function.

The whole ATC environment aimed at generating more realism from a pilot's perspective. Hence, so-called pseudo-pilots were also involved to operate surrounding aircraft.

5 Description of the Experiment

5.1 General

The FLYSAFE Main Task Experiment (MTE) simulator trials were performed at NLR's GRACE and NARSIM simulation facilities in the February-April 2009 period. Nine crews, consisting of pilots of project partners and external airlines, participated for four days each. This resulted in a total of 295 runs over four sub-experiments. The crews were adequately trained for the NG-ISS functions through laptop-based Computer Based training, a personal briefing and dedicated function training on GRACE. Furthermore, a retired controller trained the ATC participants on operational aspects like CPDLC.

The crews operated an A320-like aircraft inside the TMAs and on the ground of Paris Charles de Gaulle (LFPG) and Innsbruck (LOWI), and during en-route flights between these airports.

5.2 On-ground sub-experiment

During this sub-experiment, all measurement runs were performed having the crews wearing an ASL501 Eye-point of Gaze head/eye tracker system via which workload indications were derived afterwards from pupil activity using NLR's GazeProc and Dualgaze software tools. A minimum fixation time of 100 ms of the pupil is regarded representative for looking at a display element. Hence, the GRACE cockpit front surface was divided into certain Areas of Interest (AOIs) mapped to the novel cockpit displays, but also to the outside visual view, pedestal, autopilot control panel, EFB, etc. For this analysis the recorded number of fixations, fixation times and blinking rates were used, both for the Pilot Flying (PF) and the Pilot Non-Flying (PNF).

Each measurement run ended with a short debrief and a computer-based questionnaire on topics like Situation Awareness, Overall Flight Safety and Overall Workload, to subjectively establish these aspects individually. It contained various, known rating scales, see [18,19,20], as well as self-created 6-point rating scales.

A real comparison (within-groups) design experiment was created. Hence, a full repeat of similar runs was done per crew-group. For certain scenarios the four Even Crews, numbered 2, 4, 6, 8 (named EC in a combined way) all had the FLYSAFE NG-ISS system functions and new display formats available, while the Odd Crews (1, 3, 5, 7, and similarly named OC), had not and as such were used as a baseline. For the other scenarios the Odd Crews used the NG-ISS functions. There were four scenario categories, as shown in the Table 1 below.

	NG-ISS ON	NG-ISS OFF
Reference taxiing and normal take-off (E08) / normal landing (E012)	EC - (E08) OC - (E12)	OC - (E08) EC - (E12)
Taxi conflict and normal take-off	EC - (E07) OC - (E11)	OC - (E07) EC - (E11)
Taxi conflict & Take-off with Runway Incursion	EC - (E06) OC - (E10)	OC - (E06) EC - (E10)
Landing with a Runway Incursion, (go-around expected)	EC - (E09) OC - (E13)	OC - (E09) EC - (E13)

Table 1 Taxi scenarios division over crew groups, with scenario number in parentheses

Scenario E13 was executed at LOWI, all others at LFPG.

When operating near or at LFPG, NG-ISS active implied that all new AMM presentation features, taxi display (control) features, data link communications (DCDU), SMAAS presentation, ICS and RCAF alerts, traffic and traffic conflict presentation were available. Obviously, these functions were not made available without NG-ISS. All LFPG runs were

executed in low visibility in the range of 200 to 300m RVR.

When flying to LOWI communication was performed by radio telephony (R/T), and the visibility was marginal but would not prevent a landing. Also the on-board AMM was not available for Innsbruck, but the traffic information was presented on the ND, all of which was known to the pilots. In this setting, NG-ISS active did imply that an aural RCAF alert would be provided jointly with the PFD alert labels.

5.3 Procedures

Cockpit procedures

During the full MTE, having the role of PF also implied the role of Pilot Taxiing, which sometimes differed from certain airline procedures where the Captain would always be the Pilot Taxiing, irrespective of who was (or had been) PF during that flight. However, it was intended to distribute operating tasks equally over both crew members by allowing them both to be exposed to the novel NG-ISS features in the role of PF as well as in the role of PNF.

The only real novelty being the CPDLC aspects, crews had to apply standard or normal crew operational procedures for all other aspects and events.

ATC and CPDLC procedures

At LFPG, being a complex airport, it was not always clear when the transition of R/T to CPDLC and vice versa should actually take place. So a new procedure was introduced. In this new procedure a pilot would always contact the controller by R/T when he changed frequency and the controller informed him (via R/T) to continue either by CPDLC or via R/T.

The start-up controller, ground controller and runway controller (take-off only) were using data link. The ground controller used a list of pre-scripted taxi routes to minimize the workload. The runway controller (giving landing clearances) and the approach/departure controller were using R/T, which was also used in case of route deviation by the pilot or in case of non-normal operation.

6 Main Results of the On-ground Sub-experiment

The results of the on-ground sub-experiment address the following safety aspects: deviations from the cleared taxi route, changes in head-up/head-down times, perceived situation awareness/workload/safety and prevention of runway incursions. Furthermore, the performance of taxi operations is presented in terms of taxi times and taxi speeds. The question is whether the (anticipated) safety increase results in less efficient taxi operations.

6.1 Taxi routes

The results of the taxi operations with respect to the number of taxi route deviations are presented in Table 2. Human errors were present in most reference (i.e. NG-ISS OFF) scenarios and strongly reduced in the NG-ISS scenarios. The number of taxi errors as percentage of the number of runs was reduced from 56% to 12% with the NG-ISS present.

Scenario No	**Nr of Ref Runs; NG-ISS OFF (Number of Deviations in Ref runs)**	**Nr of FLYSAFE Runs; (NG-ISS ON) (Number of Deviations in FLYSAFE runs)**
E06	4 (5)	6 (1)
E07	5 (3)	5 (0)
E08	4 (2)	4 (0)
E10	5 (1)	7 (0)
E11	5 (4)	4 (0)
E12	4 (0)	7 (3)
Total	**27 (15)**	**33 (4)**
	56%	**12%**

Table 2 Taxi scenario NG-ISS (on / off) results

As an example Fig. 9 presents taxi scenario E06 (in green) implying "Taxi to 08L, via G, F, U, UC, WB and WA" and Fig. 10 shows the executed taxi routes. Note that all runs were

performed in low visibility and by pilots who were not really familiar with the airport lay-out.

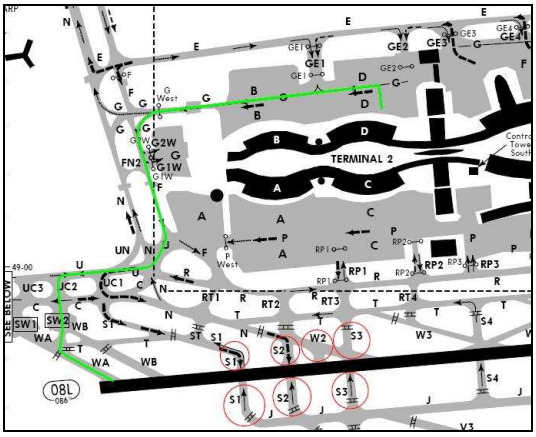
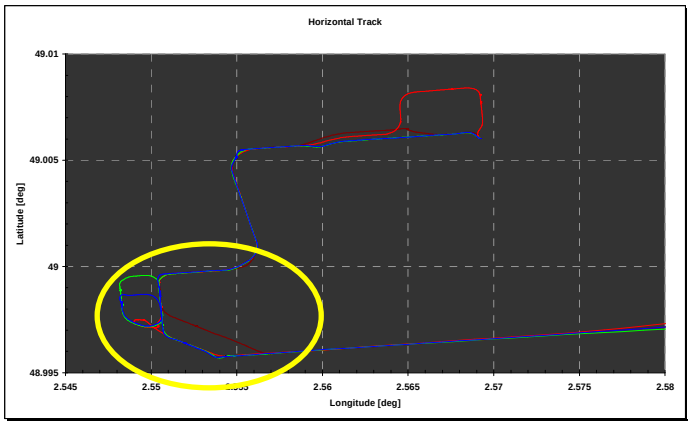
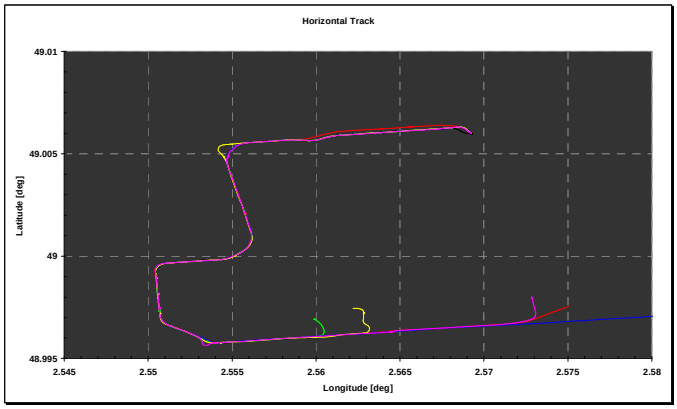


Fig. 9 Paris Charles de Gaulle Taxi Scenario (E06)



a) NG-ISS OFF results



b) NG-ISS ON results

Fig. 10 LFPG Scenario E06 results for NG-ISS OFF (a) and NG-ISS ON (b)

As can be seen, five deviations of the nominally instructed route occurred having the NG-ISS switched off versus only one when having it on. The latter was the result when a crew was turning from taxiway G onto F just after the start of the taxi run. The error was immediately recognized by the flight crew. However, they could not make the normal turn from G onto F anymore

The other taxi route deviations with NG-ISS did occur in scenario E12. In this scenario even two deviations occurred in a single run. This was a run with an AMM display but without a taxi route presentation on it (due to a simulation error). The third deviation was, according to the flight crew, caused by display confusion between the own-aircraft symbol and the cursor symbol on the AMM.

6.2 Taxi times and taxi speeds

An important question is whether functions designed to improve one aspect of flight operations (e.g. safety) do have a negative side effect. For example, do the taxi functions reduce the performance of the taxi operations? For this reason the taxi times and taxi speeds were analyzed. Taxi time depends on the route, speed, visibility and interruptions during taxiing. The interruptions in the taxi scenarios consisted of short-duration taxi conflicts and crews being lost on the airport. The taxi time was examined between two fixed points at the airport, i.e. starting to taxi at the gate pushback position and passing a fixed point near but well before the runway holding position for taxi-out and similar fixed points for taxi-in. The results of the taxi times and taxi speeds are presented in Table 3 and Fig. 11 to 13.

	NG-ISS OFF	NG-ISS ON
Mean Taxi Time (s)	547	449
Mean Taxi Speeds (kts)	9.9	12.0
Max Taxi Speeds (kts)	17.1	19.9

Table 3 Taxi times and speeds over all scenarios

These results show that taxiing took on average about 98 seconds less with NG-ISS. More general this was an improvement of about 18% in taxi times. The main reasons assumed were the improved airport position awareness

and awareness of other traffic in case of conflicts; resulting in less taxiing errors, less stops to orientate, and therefore increased average taxi speeds. An increase of 2.1 knots was measured for the average taxi speed, and 2.8 knots for the maximum speed.

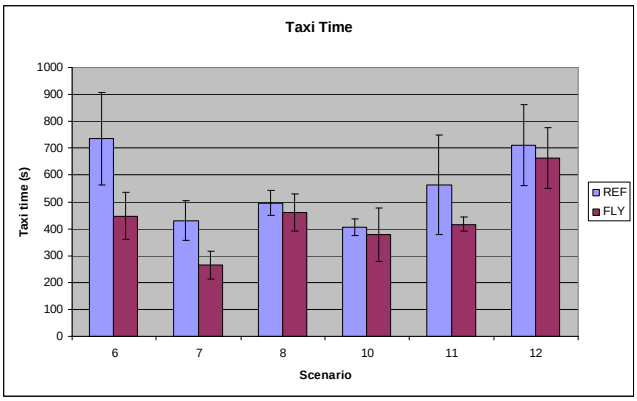


Fig. 11 Mean taxi time results for NG-ISS ON (Fly) and OFF (Ref) per scenario

Statistical analysis showed that the taxi times are indeed significantly lower for taxi operations with NG-ISS in low visibility (t-test, p=0.03007).

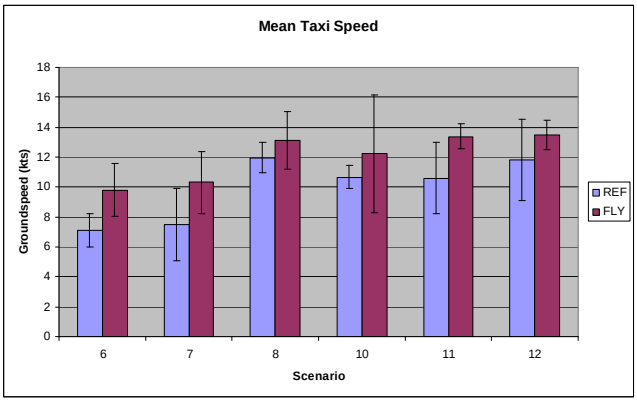


Fig. 12 Mean taxi speed results for NG-ISS ON (Fly) and OFF (Ref) per scenario

And taxi speeds are significantly higher for taxi operations with NG-ISS (t-test, p=0.00716). However, it was not statistically proven that the maximum taxi speed was also higher with NG-ISS (t-test, p=0.07626). Although the maximum speed always had a tendency to increase, irrespective of the scenario, see Fig. 13.

A word of caution is in place. The average and maximum taxi speeds are both rather high for the given low visibility conditions. Taxi speeds in the range of 5-10 knots were mentioned by several pilots as more realistic speeds under these conditions. It is therefore hypothesized that the "flight simulation effect" itself contributed to these higher than expected speeds.

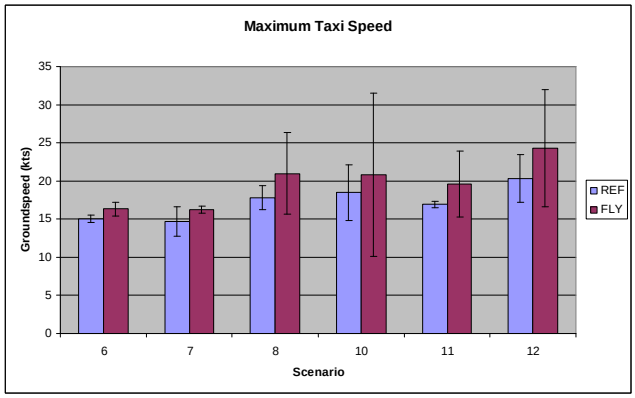


Fig. 13 Maximum taxi speed results for NG-ISS ON (Fly) and OFF (Ref) per scenario

Other flight simulator studies have also measured higher than expected (to real world) taxi speeds. So the absolute values of these results should be treated with caution, but in the comparison made the relative results are valid.

6.3 Head-up and head-down times

The effect on derived total head-up and total head-down time percentages were analyzed. The role of the Pilot (PF vs. PNF) was found highly significant, but unexpectedly not significantly influenced by the NG-ISS system itself.

But when including the various individual areas of interest, some areas in the flight deck showed significant differences when operating with the support of NG-ISS, see Fig. 14. Hence, the derived overall head-down parameter was to global. One should look at the individual head-down areas.

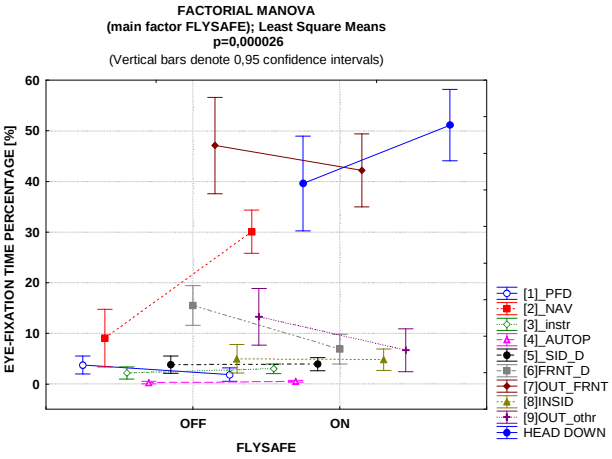


Fig. 14 Eye-fixation time percentages for all Areas of Interest and Head-down time percentages

For each line the point on the left side present the results when NG-ISS was not available and the point on the right side show the results for NG-ISS. The figure clearly reveals that with the NG-ISS system, there was more head-down time (see blue line, right top corner) and more eye-fixation time on the AMM/ND (see [2]_NAV in red) and less on the outside frontal area, EFB (showing the Jeppesen charts) and PFD, see respectively [7]OUT_FRNT, [6]FRNT_D and [1]_PFD. When zooming into these particular AOI effects, the AMM/ND and the EFB results were found to be highly significant (p=0.000000 and p=0.000820); while the head-down contribution was almost significant (p=0.054198). The results proof that with the NG-ISS system, the head-down time significantly increased, with a measured 9%, and there was at least an attention shift from the EFB to the AMM/ND. Given the new AMM this fully makes sense. It furthermore aligns with the fact that the AMM would reduce the need to look at the taxi charts that were presented on the EFB.

6.4 Situation Awareness, Workload and Safety

All crews indicated that the AMM increased SA for the on-ground operations, made traffic awareness very high, and did not increase workload. See Fig. 15 to 19.

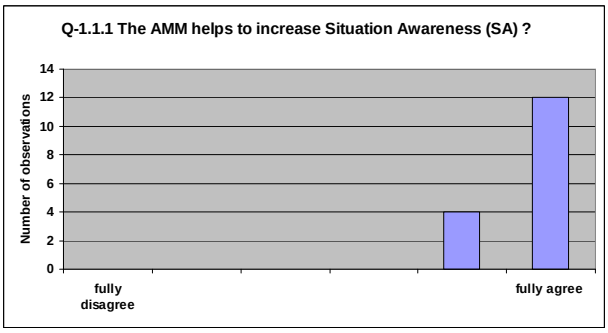


Fig. 15 AMM helps to increase Situation Awareness?

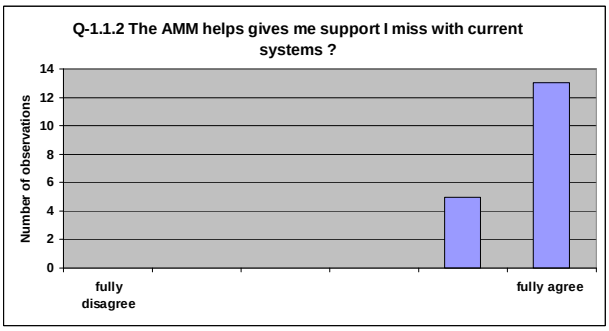


Fig. 16 This AMM helps to give me support I miss with current systems?

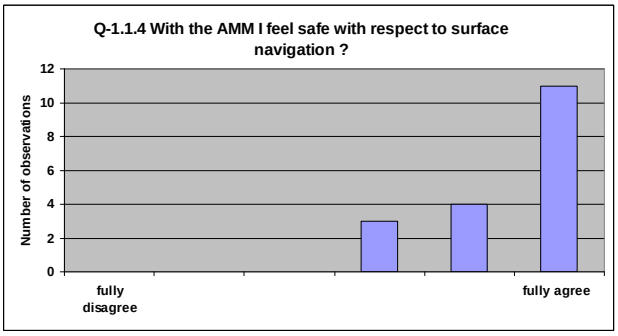


Fig. 17 With the AMM I feel safe with respect to surface navigation?

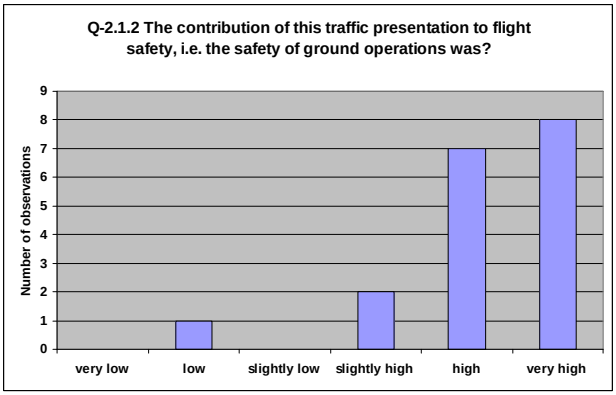


Fig. 18 Contribution of traffic presentation to flight safety for on-ground operations?

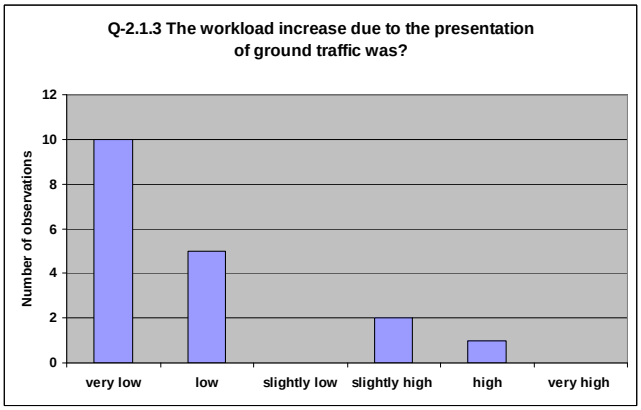


Fig. 19 Workload due to the presentation of surrounding aircraft

When looking at the blink rates and fixation rates, Table 4 below shows that the mean blink rate increased and the mean fixation rate decreased when using the NG-ISS system. This indicates a less demanding visual attention requiring task, so less staring, with the NG-ISS system, which is an indication for decreasing workload.

Blink Rates (number of eye blinks per minute)	**NG-ISS ON**	**NG-ISS OFF**
Mean Value	12.85	9.35
Standard Dev.	9.24	6.80
Fixation Rate (number of eye fixations per minute)	**NG-ISS ON**	**NG-ISS OFF**
Mean Value	140.9	144.5
Standard Dev.	19.0	14.7

Table 4 Blink rates and fixation rates as function of NG-ISS

6.5 Runway incursion prevention

Two types of runway incursions were assessed: incursions during take-off and incursions during approach and landing. See Table 5.

Take-off runs

In low visibility and during the nine reference take-off (TO) runs, the crews were (purposely) never aware of the runway incursions taking place by other traffic. All these runs were 100% incursion events.

Scenario Number	**Number of Reference runs (NG-ISS OFF)**	**Number of runs with NG-ISS ON**
E06 - take-off	4	6
E10 - take-off	5	7
Sub-total-1	**9**	**13**
E09 - approach	4	9
E13 - approach	4	5
Sub-total-2	**8**	**14**
OVERALL TOTAL	**17**	**27**

Table 5 Runway incursion scenarios and NG-ISS

With the NG-ISS system, three out of thirteen TO runs (23.1%) resulted in a continued take-off and 76.9 % into a rejected take-off (RTO). From these latter ten RTO cases, nine were the result of a RCAF alert only, while only one was due to the combination of AMM traffic presentation and RCAF alerting. This indicates the strong reliance on aural alerts during the take-off roll. The visually presented information seems more important for confirmation.

One new HMI issue surfaced: the case where conflicting traffic is visible on the AMM but RCAF did not generate an alert. For some reason the RCAF didn't perform its intended function in three cases, and consequently no alert was presented. Still during these three cases the flight crew carried out the take-off as normal, while conflicting traffic was shown on the AMM. In one run the flight crew just didn't spot the other traffic while in the two other runs they spotted the traffic on the runway but didn't reject the take-off because the other aircraft crossed at the very end of the runway and no RCAF alert was given, hence they judged the system regarded the situation as implicitly safe.

Apart from the system integrity aspects, the (procedural) issue is what to do in these cases, should a procedure be prescribed or should the situation be subject to the captain's (or more precisely pilot in command's) decision? And should the PNF inform the PF at all times during the take-off roll with the displayed and noticed conflict information?

Approach runs

All the approach and landing scenarios included runway incursions. ATC would provide a landing clearance despite the fact that an aircraft still occupied the runway, or infringed a stop bar. The scenarios consisted of 8 reference runs equally divided over LOWI and LFPG, and 9 runs (at LFPG) with the full NG-ISS system and 5 runs (at LOWI) without AMM but with the RCAF system.

In all 14 approach runs where the aircraft was equipped with the NG-ISS system this led to a go-around, which was the correct action because the scenarios should all lead to this action. Two (out of 8) reference runs ended with a 'normal' landing. Despite the involved ATC aspects, these two runs were considered to be real runway incursion occurrences because an aircraft was still prominent on the landing runway.

However, it can not be concluded that AMM/RCAF was always the reason to initiate a go-around. Other factors, accounting for 50% of the go-arounds, were ATC-commanded or visual detection of traffic by the crews themselves.

In about 78% (7 out of 9) of the runs with a timely provided landing clearance, the go-around was initiated based on information from AMM and/or RCAF alerts. In the other two runs the RCAF alert was not audible in the cockpit and traffic was not displayed on the AMM runway, or the AMM was not available at all (e.g. Innsbruck scenario).

Furthermore, in approx 35% of total and 100% of runs without a landing clearance the AMM and RCAF also supported the flight crew in deciding to go-around in case a late landing clearance was expected but not given.

Finally, the questionnaires show that perceived safety was increased through the use of the RCAF system, see Fig. 20.

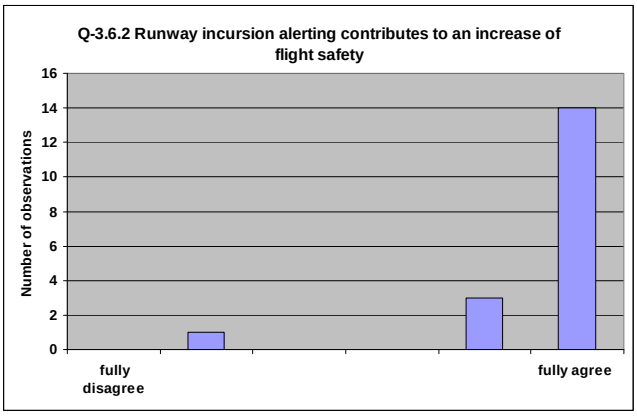


Fig. 20 Runway incursion alerting (of RCAF) contributes to an increase of flight safety

6.6 Some ATC results

The ground controller would select the taxi route to be uplinked to the aircraft via a graphical airport map. This only would work when both sides used the same internal representation of the airport (same database and same projection method) and the same algorithms to display the taxi-route. In current practice this is not possible yet. Ensuring that the same taxi route is presented on the displays of the controller and the pilot required additional unique points to be entered by the controller. This resulted in an unacceptable increase of the controller workload. The solution was to use pre-scripted routes that could be selected by the ground controller. This reduced his workload, but limited his options.

The procedure developed for the transition between R/T and CPDLC was very well accepted by the controllers, and was seen as the correct procedure to ensure that controller and pilot are on the same R/T frequency in case of an emergency.

Sometimes, from an ATC perspective, the aircraft was already "reacting" on the instruction provided by data link while no "Wilco" was received from the pilot. Because the controller could not be sure that the aircraft is reacting to his instruction, or just acting independently, he would contact the aircraft immediately by R/T to stop the started maneuver of the aircraft, leading to confusion and adding extra workload.

ATC (and crews) considered the overall data link process too slow in many situations. This

generally related to the many handling steps to perform, not to the message transmission/reception speed. Data link (message sets and procedures) was considered useful from the controller's point of view, but when a direct response is necessary or in case of an emergency, data link cannot replace R/T. The former makes data link less suitable for the runway and approach/departure positions.

7 Conclusions and Recommendations

The piloted assessment of the NG-ISS concept for the on-ground aspects has shown, based on subjective data, that in an overall sense the NG-ISS functions proved to be very beneficial in terms of improving the crew's situation awareness. Partially based on measured data, it is concluded that crew workload decreased and head-down time increased when working with the airport moving map (AMM) under low visibility conditions. The latter is a concern. Furthermore, strong case-based evidence has been provided that the operational flight safety will be improved using the AMM, especially for runway incursion prevention, due to a lower human error occurrence expectation. In combination with the measured higher relative taxi speeds, hence lowered taxi times, the use of AMM may lead to greater airport efficiency under low visibility in the future.

It is also concluded that the FLYSAFE runway incursion prevention logic (RCAF), solely or in combination with the AMM, improved flight safety and was either the main factor (in case of timely landing clearances) or a contributing factor (in case no landing clearance was given) to initiate a go-around in the given low visibility conditions.

The experimental CPDLC taxi clearance instruction procedure was well accepted by both pilots and controllers. However, the overall judgment of both groups on data link matters was that it was unambiguous and good on the one hand, but generally too slow in comparison to R/T and in operational processes that require immediate actions, on the other hand. Also when data link is used a procedure shall be in place to ensure that crews will start maneuvering after the controller has received "Wilco". Since R/T will always remain necessary for back-up and emergencies, there should be a clear procedure describing how to initiate ATC contact. Hence, the use of data link requires further detailing and optimization. To enable functions like taxi route uplink requires the harmonization between ATM and aircraft concerning the internal representation of an airport and the algorithms used to determine such a route. The pre-scripted routes do not seem to be the way forward in introducing taxi route uplink.

Finally, the involved pilots and ATC experts provided a wealth of detailed and very useful comments to enhance these cockpit novelties even further, of which the most important ones have been reported within the project.

Acknowledgment

The authors wish to acknowledge the efforts made by all relevant project partners and participating test (and pseudo-)pilots, ATC members, airline crews, and the European Commission who co-funded this research via contract No. AIP4-CT-2005-516167.

References

[1] Advisory Council for Aeronautics Research in Europe (ACARE), "*European Aeronautics, A Vision for 2020, Meeting society's needs and winning global leadership*", Report of the Group of Personalities, January 2001.

[2] Advisory Council for Aeronautics Research in Europe (ACARE), "*A Strategic Research Agenda, Volume 1*", October 2002.

[3] Advisory Council for Aeronautics Research in Europe (ACARE), "*A Strategic Research Agenda, Volume 2*", October 2004.

[4] Advisory Council for Aeronautics Research in Europe (ACARE),"2208 Addendum to the Strategic Research Agenda", 2008.

[5] Verhoeven, R.P.M., Rouwhorst, W.F.J.A., "FLYSAFE – WP 111, D 1.1.1 Final Operational Concept Description", 16/09/2005.

[6] Commercial Aviation Safety Team (CAST) and the General Aviation Joint Steering Committee (GAJSC)," Runway Incursion Joint Safety Analysis Team (JSAT), Results & Analysis", August 11th 2000 and December 2002.

[7] Yves Morier (JAA) & Rudi den Hertog (Fokker Services), "Future Aviation Safety Team (FAST), Increased Reliance on Flight Deck Automation", Dedicated presentation for the FLYSAFE Project, March 2005.

[8] EASA, Safety Analysis and Research Department "Annual Safety Review 2010". ISBN 978-92-9210-097-1.

[9] Eurocontrol Runway Safety Office, "European Action Plan for the Prevention of Runway Incursions, Release 1.1, Brussels, 2004.

[10] FAA Offices of the Runway Safety, "FAA Runway Safety Report: Runway Incursion Severity Trends at Towered Airports in the United States 1997-2000", Washington D.C. 2001.

[11] Eurocontrol Runway Safety Office, "European Action Plan for the Prevention of Runway Incursions, Release 1.2, Brussels, 2006.

[12] Vernaleken, Christoph; Mihalic, Lamir; Guettler, Mathias; Klingauf, Uwe, "*A fresh look at runway incursions: onboard surface movement awareness and alerting system based on SVS*", Proceedings of SPIE. Vol. SPIE-6226, pp. 202-217, 2006.

[13] Vernaleken, Christoph; Urvoy, Carole; Klingauf, Uwe, "Simulator Evaluation of a Novel Surface Movement Awareness and Alerting System (SMAAS) for Runway Incursion Avoidance", AIAA-paper-2008-8959-107, 8th *Aviation Technology, Integration and Operations (ATIO) Conference*, 2008.

[14] Vernaleken, Christoph, "Autonomous and Air-Ground Cooperative Onboard Systems for Surface Movement and Incident Prevention", Ph.D. Dissertation (in english), Fachbereich Maschinenbau, Technischen Universitat Darmstadt, Aviatic Verlag GmbH, Oberhaching, Germany, ISBN 978-3-925505-99-7, May 2001.

[15] Sammut, A., Zammit, B., Zammit-Mangion, D., "A Traffic Movement Monitoring Computer Pre-Emptive Runway Conflict Design", Conference Paper, AIAA Paper 2006-6272, *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Keystone, Colorado, USA, August 2006.

[16] Sammut, A., Zammit, B., Zammit-Mangion, D., "A Traffic Surveillance Function and Conflict Detection Method for Runway Manoeuvers", Conference Paper, AIAA Paper-2007-7738-826, *7th AIAA Aviation, Technology, Integration, and Operations Conference (ATIO)*,Proceedings. Belfast, Northern Ireland, 2007.

[17] Sammut, A., Zammit, B., Azzopardi, M.A., Zammit-Mangion, D., Szasz, S., "Design and Pilot Evaluation of a Runway Conflict Detection and Alerting System", Conference Paper, AIAA Paper 2010-7694-358, *AIAA Guidance, Navigation and Control Conference*, Toronto Canada, 2-5 August 2010.

[18] NASA *Human Performance Research Group*. (1987), Task Load Index TLX, version v1. User Manual.

[19] McGuinness, B. and Foy, L. "A subjective measure of SA: the Crew Awareness, Rating Scale (CARS)", Paper presented at the *Human Performance, Situational Awareness and Automation Conference*, Savannah, Georgia, 16-19 Oct 2000.

[20] Zijlstra, F.R.H., "Efficiency in Work Behavior. A Design Approach for Modern Tools", Ph.D. thesis, Delft University of Technology. Delft University Press, Delft, The Netherlands, 1993.

Virtual Reality based HUD (Head Up Display) to simulate 3D conformal symbols in the design of future cockpits

Francesca De Crescenzio, Sara Bagassi, Massimiliano Fantini, Francesca Lucchi
DIEM, II Faculty of Engineering (Forlì - FC), University of Bologna, Italy
Contact Author: francesca.decrescenzio@unibo.it

Keywords: *Virtual Reality, HUD Head Up Display, HMI Human Machine Interface, Cockpit Design*

Abstract

Advantages of using HUDs in aviation have been demonstrated. Therefore, this display technology is widely being applied both in civil and military flight operations. Currently, a considerable interest is being focused on the possibility of displaying conformal symbols, intended as geo-referenced symbols, representing data that are usually displayed onto HDDs (Head Down Displays).

In order to support the design of novel cockpit interfaces based on the use of HUDs for displaying such conformal symbology, in this paper we propose a Virtual Reality simulation system. The aim of this simulator is to provide a tool to early investigate the application of three-dimensional graphics in future HUDs. First concept results of 3D weather data representation in cruise and approach scenarios are presented.

This work has been developed in the framework of the ALICIA (All Conditions Operations and Innovative Cockpit Infrastructure) Project (Seventh Framework Programme of the European Commission).

1 General Introduction

ALICIA is a research and development project funded by European Commission under the Seventh Framework Programme [1]. One of the two main objectives of the project is the development of a new cockpit architecture facilitating the introduction of new technologies and applications. A very challenging issue in this framework is related to the management and the presentation of information.

In fact, the increasing availability of information onboard air vehicles is making the design of the flight deck a very complex issue [2]. Experts of Human Computer Interfaces agree in the need of simple and intuitive interfaces [3]. Nevertheless, the unpredictable behavior of the pilot requires designers to focus on interfaces which ensure that he works under the required safety conditions and increasing the time efficiency.

The HUD concept plays a fundamental role in this context. Displaying data on a transparent layer allows the pilot to simultaneously look at them and the outside world by means of a collimated projection. Commercial HUDs differ in the nature of displayed data and in the interface design. A considerable interest is currently being focused on the possibility of displaying conformal symbology, intended as geo-referenced symbols (as opposed to non-

conformal navigation data), representing data that are usually displayed onto HDDs (Head Down Displays). While being associated with real features in the flight scenarios, such conformal symbols are typically based on two-dimensional graphics. Especially with the aim of enhancing flight safety with regards to airborne obstacle proximity and navigation hazards Synthetic Vision Systems (SVS) are being developed for displaying 3D data on HDDs [4][5]. The implementation of synthetic vision together with the HUD concept is expected to have a big potential in increasing the situation awareness and mitigating the pilot workload [6].

Therefore, the aim of this prototype is to provide a tool to investigate the use of three-dimensional conformal information in HUDs.

A Virtual Reality based HUD simulator has been conceived. It exploits a stereoscopic visualization screen where a double viewport has been created to overlay a computer generated HUD onto a synthetic flight scenario. A library of 3D symbols that represent weather data is implemented in the virtual environment. Finally, simulation results in an approach scenario are presented.

The exploitation of such virtual prototyping system in cockpit design is expected to give advantages in terms of an improved capability of following the evolution of both safety and efficiency requirements [7].

2 VR system architecture

The Virtual Reality system proposed is a visualization environment conceived to support a layered representation of data. The background layer displays the computer-generated scenario, such as the 3D terrain and the other features that replicate the view from the flight deck. This is the external view layer. The HUD layer simulates the transparent display. Conformal symbology is rendered onto the HUD layer. Figure 1 provides the simulator layout.

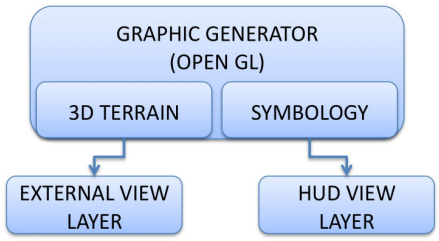


Fig. 1 Lay out of VR graphic generator

2.1 External view layer

The external view layer displays the 3D terrain and the flight environment, depending on the selected scenario. To support the visualization of the scenario, we propose either a large screen stereoscopic display or a multi monitor stereoscopic display. In following Figure 2 such devices are depicted.



Fig. 2 a) Large Screen Stereoscopic Display, b) Multi monitor Stereoscopic Display

The large screen stereoscopic display is equipped with passive stereoscopy while the multi-monitor one is based on active stereoscopic visualization.

The external view layer consists in a visual application that replicates the actual scenario which the pilot would see beyond the HUD layer in a real application.

The terrain surface is built as a quad mesh based on DEM (Digital Elevation Maps) and implements multi-resolution features to adapt the LOD (Level Of Detail) of the surface to the field of view. Terrain rendering is achieved using high resolution satellite pictures.

The airport infrastructure, such as buildings and runways, is also reproduced with an editor

for 3D shapes embedded in the visual application.

The brightness of the external view layer can be adjusted according to the flight condition to be simulated (night time, daytime).

2.2 HUD layer

The HUD visualization layer is parametric. It is not designed for a specific aircraft but is intended to be a general purpose visual simulator adaptable to multiple platforms. The symbology layer is visualized in an OpenGL viewport which is overlapped on the external view display. This results in a simulated symbology display that replicates the *see-through* effect of the real Head up Display.

The lay out of the two graphic viewports is depicted in fig.3.

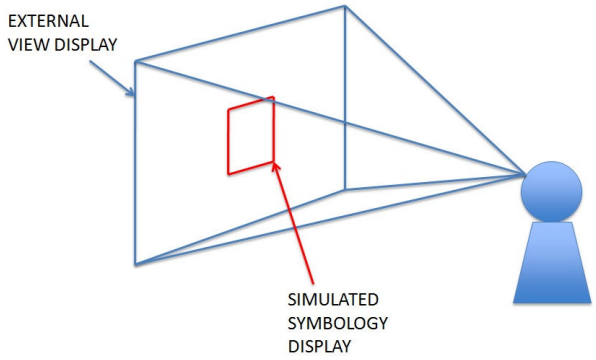


Fig. 3 Lay out of the double viewport display

3 3D Symbology library

The atmospheric hazard is one of the most important external aviation hazards [8]. This is due not only to the physical effect that the atmospheric phenomena can have on the aircraft but also on the operative pilot's capability of reacting to such hazards, e.g. with trajectory changes [9]. Among the various atmospheric hazards which pose risks to aircraft, there are the phenomena related to thunderstorms (CB – Cloud Banks) convective turbulence, icing and also visibility. Usually weather information are available to the pilot onto the head down display. The symbology visualization approach proposed in this paper aims at bringing a set of weather data onto the HUD. Moreover, it supposes to investigate how a three-dimensional visualization of the information related to the weather could provide a useful and unambiguous visual cue to the pilot.

The attributes given to the weather in this paper concern the information coming from radar and satellites. Such input sources are used in defining top and bottom volumes contours, as well as cloud bank height and center. With these data we sketch a 3D conformal volume that is displayed onto the symbology layer.

Cloud banks are split into a top and a bottom part, in order to better define and identify weather distribution. Both parts are rendered as a semi-transparent conformal volume, possibly with an highlighted embedded region, showing any severity increase as a conformal volume as well. In the following fig.4 an example of the CB volumes rendering is shown. It is an approximation to cylinder shapes adopted for the generation of first demonstrative prototypes. Vertical extrusions of plane boundaries extracted from satellite images will be also proposed to increase the realism of simulations.

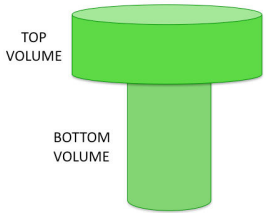


Fig. 4 CB rendering architecture

Further information that has been considered in this work is the specific weather phenomena. To this aim we adapted weather chart symbols to be rendered as 3D geometries. This approach should take advantage of the pilots' previous knowledge and is adopted to reduce the familiarization time.

In a first prototyping phase, we selected a small collection of symbols, with different complexity levels. In following Table 1 it is reported the selected set of weather phenomena, as well as the corresponding weather chart symbol, and the 3D shape proposed to be rendered in the virtual HUD.

	2D Symbol	3D Symbol
Rain	circle	Sphere
Rain shower	triangle	Cone
Thunder	polyline	tubular shape
Turbulence	polyline	tubular shape

Tab. 1 2D – >3D Symbols – correspondence

The circle that usually represents the rain symbol has been converted into a sphere. The rain shower, which is represented with the triangle, corresponds to the vertex down cone. Thunder and turbulence have been modeled as the tubular shapes generated around their representative polylines.

These latter geometries are not symmetrical. Thus in three dimensions a rendering option was arranged in order to keep the plane they lay on perpendicular to the pilot's view.

4 Scenario prototypes

Among the main issues to be addressed in the design of this interface there are:

- Symbology readability, with respect to actual distance of the atmospheric hazard from the flight deck.
- Clutter level.

Thus, it is needed to build an example scenario in order to collect suggestions from a focus group.

The flight here proposed as example refers to an approaching phase to Capodichino airport (Naples, Italy), with a flight heading of 233°. A virtual movie of the approach is made from data recorded during an actual flight and provided by Alitalia. A virtual camera is posed in the flight deck to reproduce what the pilot would see during the approach. Good visibility is set in this first experiment, since it is focused on the communication of the concept based on the 3D weather volumes.

A CB is supposed to be close to the airport: it is approximately centered in: 40.9134 lat, 14.2767 lon. Symbols have been distributed in and around the volume. The plane projection of such distribution is depicted in Fig.5.

		Lat	Lon	Alt m
●	CB Center	40.9134	14.2767	
●	Turbulence	40.93	14.33	3800
●	Thunder	40.93	14.32	1700
●	Rain	40.9034	14.33	3900
●	Rain	40.935	14.33	3600
●	Rain	40.93	14.35	1300
●	Rain	40.93	14.29	1300

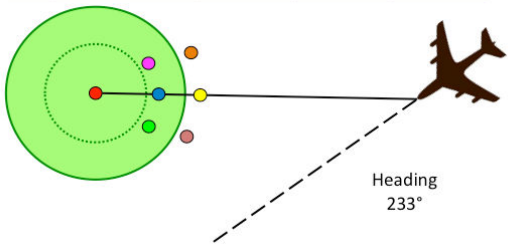
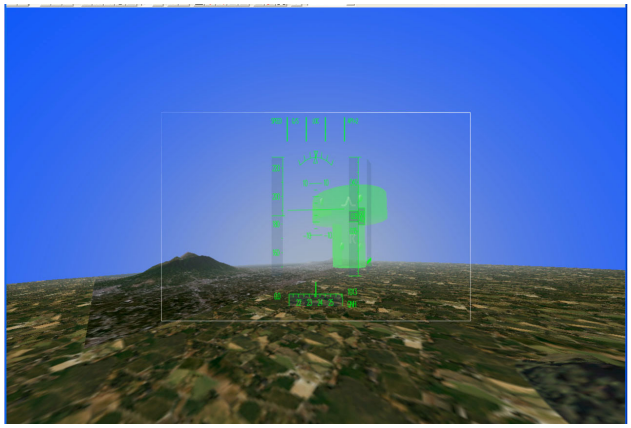
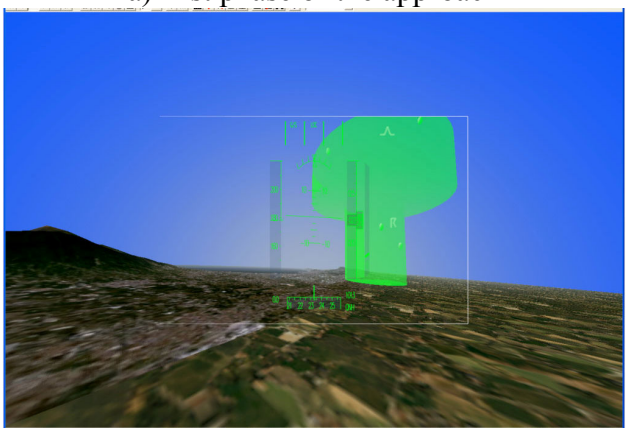


Fig. 5 Distribution of weather 3D symbols

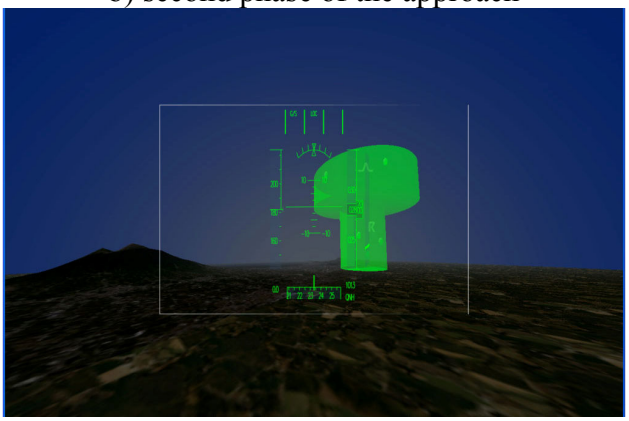
The images resulting from the overlapping of the HUD layers on the external view are depicted in Fig. 6 a), b) and c).
These are snapshots taken from a virtual movie generated with the visual application. The white frame inside the picture is the simulated HUD edge. In the HUD a set of non-conformal data is depicted together with the 3D volume of the CB and the weather symbology. While the terrain representation in the two layers is continuous, the HUD symbology is only visible if inside the viewport cone. As a matter of fact, in fig. 6 b) the CB volume is only partially visible, since it refers to a view that is closer to the CB. The last image (Fig.6 c) represents the same scenario in night visibility conditions.



a) first phase of the approach



b) second phase of the approach



c) approach in night visibility conditions

Fig. 6 Snapshots from the virtual video

5 Evaluation

There are two evaluation objectives that we intend measuring to validate the concept of displaying 3D conformal symbols onto the HUD in our simulator. They mainly concern the capability of the system of providing useful information on:

- The enhancement of the cockpit overall usability.
- The improvement of pilot's SA (Situation Awareness).

Nevertheless, being in the early stage of the development, the evaluation has been organized in three phases. The first one is the communication of the concept and aims at giving the first information regarding the system and its architecture. During this phase the members of the focus group are asked to familiarize with it and provide suggestions for the refinement of the prototype. The refinement of the system follows this phase. Finally, the measure of the overall usability and the SA should be performed via questionnaires on the refined system. Currently, only the first phase has been completed and is presented in this paper.

During the first phase of the evaluation, the images from the previous section have been presented to the ALICIA EEAG (External Expert Advisory Group). The EEAG is a group of experts coming from airliners, helicopter and aircraft operators and other representative stakeholders of the project results. End user representatives are involved from the early stages and during the life of the project to deliver operational and practical feedback to partners' proposals [1].

The presentation mainly focused on the following three points:

- Description of concept objectives;
- General description of prototype concept;
- Presentation of snapshots and movies in a virtually reconstructed scenario.

Several comments rose from the members of the EEAG. Such comments mainly concern the information presented in the HUD reconstructed images both in the type of information and on the way of presentation/visualization. In particular, it has been highlighted that it would be important to specify the weather context (including the time of observation measurement) of the images and the weather data source in order to allow the focus group to have a clear idea of the scenario in which the prototype is adapted. Moreover, it has been asked to specify the correlation between the navigation display and the virtual

HUD proposed. In order to increase the credibility of the simulated scenario, it has been suggested to include a flight path vector and to adjust the altitude indicator. It has been remarked that the capability to read the instruments is a very important issue to be considered. A further remark has been made on the use of shaded shapes. Specifically, during the night condition it may result too bright and a wireframe approach could be preferable.

Finally, it has been proposed to add the hail, icing and wind shear 3D symbols to the library and to think about the discrimination between real data and forecast.

If more information layers are present in the horizon, one layer should suppress the others. Thus, priority areas should be defined. Moreover, it has been pointed out that any proposal covering the horizon runs into certification problems.

6 Conclusions

In this paper we have presented the first prototyping and evaluation phases for the generation of a novel cockpit HMI. This is based on the use of conformal data on the HUD and exploits a virtual reality system for the concept development. Information displayed onto the HUD should:

- •Supplement the conventional head down representation of weather with complimentary head up data.
- Provide 3D support for the "detect and avoid" strategies.

Very encouraging comments have been collected during the first evaluation phase. Further developments are already planned. They concern the refinement of the VR based HUD simulator and the evaluation of the workload and Situation Awareness measures related to the concept.

7 References

[1] http://www.alicia-project.eu/

[2] Sheridan, T.B. "Humans and Automation", John Wiley & Son, Inc., Publications, Santa Monica,2002.

[3] Wickens, C., D., Gordon, S., E., Yili Liu, "An Introduction to Human Factors Engineering – Second Edition", Prentice Hall, New Jersey, 2004.

[4] Theunissen, E., Koeners, G.J.M., Roefs, F.D., Rademaker, R.M., Jinkins, R.D., Etherington, T.J. "Guidance, Situation Awareness and Integrity Monitoring with an SVS+EVS", AIAA Guidance, Navigation, and Control Conference and Exhibit, San Francisco, 2005.

[5] Rouwhorst, W.F.J.A., Groeneweg, J.,Winckelmans, G., Luckner, R. "Piloted Simulation Results of an Onboard Wake Vortex Detection, Warning and Avoidance System and Cockpit Display Aspects", ICAS 2008, Anchorage, AK, 2008.

[6] Berth, C., Hüttig, G., Lehmann, O., "Research on Integrated Collimated Cockpit Visual and Flight Information System", ICAS 2008, Anchorage, AK, 2008.

[7] Warwick, G., "Simulator image-generation technology from CAE could boost helicopter avionics", Aviation Week and Space Technology, April 27, 2009, p.44, McGraw-Hill.

[8] http://www.eu-flysafe.org

[9] Johnson,N., Wiegmann, D., and Wickens, C., "Effects of Advanced Cockpit Displays on General Aviation Pilot's Decisions to Continue Visual Flight into Instrument Metereological Conditions", Proceedings of the Human Factors and Ergonomic Society, 50th Annual Meeting, San Francisco, 2006.

CEAS 2011

Advancements in Fluid-Dynamics 2

Fast Diagnostics for high enthalpy hypersonic flows

Antonio Esposito and Enzo Caso
Department of Space Engineering (DIAS)– University of Naples Federico II
Via Claudio 21 – 80125 Naples (Italy)

Keywords: *high enthalpy flows- hypersonic-diagnostics*

Abstract

Simulation of the aerothermal environment experienced by space vehicles during atmospheric reentry can be performed by arc-jet facilities, which generate high enthalpy hypersonic flows for long time. Such is the case of Small Planetary Entry Simulator (SPES) at the Department of Aerospace Engineering of University of Naples (DIAS), described in the paper.
To support arc-jet tests within SPES it is important to dispose of a simple tool which evaluates all aerothermodynamic flow properties by using the few experimental measurements available and a chemical model of the high enthalpy flow. In the paper we present the model implemented and used for arc-jet tests performed for AVIO company, to evaluated performances of its thermal protection material for the European space vehicle named IXV.

1 Introduction

Arc-jet facilities are used to produce the heating environment of high-speed planetary entry flights. One important parameter that such tests must simulate is the flow enthalpy; in the test section of an arc-jet facility, the enthalpy of the core region is referred to commonly as the "centerline" enthalpy; it is this centerline enthalpy that must simulate the flight value. In contrast, the mass average of the enthalpy in the stream, including the portion in the cold outer region, is called the "mass-averaged" enthalpy. The ratio of the centerline enthalpy to the mass-averaged enthalpy, which is by definition greater than unity, is an important parameter in the use of the arc-jet facilities. Because heat transfer rate is proportional to enthalpy, any uncertainty in this value translates directly to the uncertainty in the performance of the tested heatshield material. Then, considering the importance of the accuracy of the enthalpy values for arc-jet testing, it is essntial to dispose of a tool which evaluates the centerline enthalpy and other flow properties by using the few experimental measurements available, i.e the mass-averaged enthalpies, pressures and heat fluxes.

In the present paper we describe the numerical tool used for arc-jet tests performed at the Small Planetary Entry Simulator (SPES) of the Department of Aerospace Engineering of University of Naples (DIAS). Tests are executed for the aerospace company AVIO, to evaluate the performances of its material to be used as heatshield for some parts of the European space vehicle named IXV.

2 Facility Description

2.1 Configuration

The Small Planetary Entry Simulator (SPES) functional scheme is shown in fig.1. An overall view is shown in fig.2.

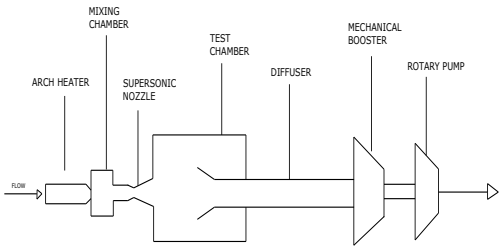


Fig. 1 SPES layout

Main SPES subsystems are:

1) Electric arc-heater (or plasma torch), 80 [kW] maximum power, operating with pure inert gases (nitrogen). The arc-heater is powered by an AC/DC rectifier, while start is obtained by a high-voltage, radiofrequency discharge;
2) a mixer where the nitrogen plasma gas can be mixed with cold oxygen, to simulate hearth planetary atmosphere;
3) four conical nozzles (area ratio 4, 20, 56 and 100) for operation in supersonic and hypersonic flow regimes (Mach numbers 3-5-6-7 respectively);
4) a cylindrical vacuum test chamber equipped with glass windows, with free jet plasma flow exhaust. Ultimate pressure in the test chamber is 50 [Pa];
5) two cooling systems. One operating with distilled water, cools the arc-heater, the nozzle and the gas heat-exchanger. The second one, operating with tap water, cools the test chamber, the vacuum pumps and the first cooling system water itself;
6) a panel, equipped with controls and measurement instruments for the manual control of the facility;
7) an electronic control and data acquisition system is used to manage the acquisition of the signals from SPES.



Fig. 2 SPES overall view

2.2 Instrumentation

The gas mass flow rates are measured by two thermal mass flow-meters; coolant flow rates are measured by variable-area flow-meters. Electrical arc parameters (i.e. voltage and current) supplied to the arc-heater are measured by digital instruments. A dedicated electronic data acquisition system allows to measure and store the temperatures from thermocouples located on the arc-heater, the nozzle and other components. Pressure measurements are made by absolute transducers which have been calibrated by a capacitive vacuum gauge.

3 Experimental Measurements

Total Enthalpy Measurements

The primary method to evaluate the total enthalpy of the stream is the energy balance method [1].
The mass-averaged enthalpy of the gas leaving the arc heater and entering the mixer is evaluated by subtracting from the input electric power the losses due to the cooling water:

$$H_{hm} = \frac{V \cdot I - (\dot{m} \cdot c_p \cdot \Delta T)_{hc}}{\dot{m}_{gh}} \qquad (1)$$

The mass-averaged enthalpy of the total flow (nitrogen plus oxygen) leaving the mixer and entering the nozzle is :

$$H_{mn} = \frac{H_{hm} \cdot \dot{m}_{gt}}{\dot{m}_{gh}} \qquad (2)$$

Finally the mass-averaged enthalpy of the flow at the nozzle exit is :

$$H_{ne} = H_{mn} - \frac{(c_p \cdot \dot{m} \cdot \Delta T)_{nc}}{\dot{m}_{gt}} \qquad (3)$$

where ***V*** and ***I*** are the arc voltage and current, $\boldsymbol{c_p m \Delta T_{hc}}$ and $\boldsymbol{c_p m \Delta T_{nc}}$ are the power losses for the arc-heater and the nozzle, $\dot{m}_g h$ is the nitrogen mass flow rate, $\dot{m}_{gt}$ is the total mass flow rate.
The relative uncertainty on the measured enthalpy was estimated to range from 10% to 20%, depending on the arc power levels. At our typical test conditions (low power and low gas flow rates) we extimate more probable the upper uncertainty level (20%).

Pressure Measurements

During each test the following pressures are measured:
1) at the downstream end of the mixing chamber
2) at the nozzle outlet
3) in the test chamber
4) the impact pressure on a probe located in the test section

The relative uncertainty on the measured pressures was estimated to range from 2% to 5% of measured values.

Heat flux Measurements

The slug-type heat-flux sensor consists of a small piece of metal (slug), inserted into the face of the model but isolated from it, whereby it measure the heat flux. The presence of an insulator is required to minimize the heat transfer to and from the model in order to approximate to the one-dimensional problem. A thermocouple measures the temperature on the rear surface the slug.

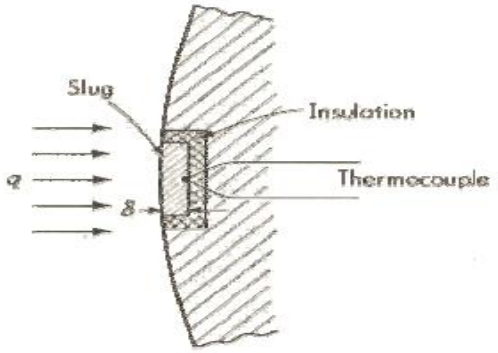


Fig 3 Slug-type heat-flux sensor

The heat flux is obtained by means of an energy balance:

$$\dot{q} = \rho \cdot c \cdot l \cdot \frac{dT}{dt} \qquad (4)$$

where $\dot{q}$ is the heat flux, ρ and c are respectively the density and the specific heat of the slug, T is the temperature of the slug, t is the time and l is the slug length.
The main characteristics of this sensor are the simple construction, no need of water cooling, low cost compared to other types of sensor, easy to install on the parts to be monitored.

4. Flow Diagnostics

A simple numerical model describing the flow evolution in SPES, starting from available experimental data, has been specifically developed. The numerical code follows the gas evolution across the arc-heater, the mixing-reservoir chamber and the nozzle; the flow parameters at the nozzle exit are considered to be the "free stream" ones in the test section.

To analyze the flow process properly and evaluate the total stream enthalpy, it was necessary to determine both the thermodynamic and chemical state of the gas at the exit of arc heater, mixing chamber and nozzle.

a)facility configuration

As first step we give the geometrical configuration for the plasma torch (anode and cathode), the mixing chamber (swirl or radial) and the nozzle.

The second step is physical input for :

- plasma torch (power, nitrogen mass flow rate, cooling water flow rate, temperature rise)
- mixer (oxygen flow rate plus cooling water flow rate and temperature rise if water cooled)
- nozzle(total mass flow rate, cooling water flow rate, temperature rise)

As output from this section, the numerical code calculate the averaged total enthalpies of the plasma torch, the mixing chamber and the nozzle.

b)plasma torch

Available experimental data are :
-nitrogen mass flow
-average total enthalpy at the torch exit
-plasma torch nozzle geometry

It is well extablished that for industrial plasma torch (such is our case) the pressure in the arc column is normally greather than 0.3 (atm), so we can assume thermodynamic and chemical equilibrium at the plasma torch exit. In this way, starting from the total enthalpy and nitrogen mass flow, it is possible to evaluate chemical composition of the nitrogen plasma and all thermofluidynamic parameters.

c)mixing-reservoir chamber

Available experimental data are :
-nitrogen and oxygen mass flow
-average total enthalpy at the reservoir exit
-pressure at the end of reservoir chamber
-reservoir geometry

In the mixing chamber, due to the low total pressure attained by the flow - in all test this pressure is lower than 0.5 atm) – the chemical composition of the gas is assumed to freeze upon entering the chamber (in particular downstream the mixing nitrogen-oxygen) and to remain frozen during the travel through the chamber.
A confirmation of this assumption can be obtained computing the atom recombination time and the residence time of the gas species in the chamber; tipically recombination times are O[10^{-2}] (s) while the residence times are O[10^{-4}] (s).
In this way, starting from the measured total enthalpy and total mass flow we can compute chemical composition (we consider the Lightill model for air).
Following Pope [2] we use its modified version of the frozen sonic flow method to calculate the centerline total enthalpy. First, the mass fraction of dissociated species α was used to calculate the value of the frozen isentropic exponent :

$$\gamma = \frac{4+3Z}{4+Z} \qquad (5)$$

where Z=1+ α and α is the total mass fraction of dissociated species.

Using the total pressure and the mass flow we can evaluate the total gas temperature at the beginning of expansion (T_0) by:

$$T_o = \frac{1}{Z}\cdot\sqrt{\frac{(C\cdot F(\gamma)\cdot p_m\cdot A^*)}{m_g}} \qquad (6)$$

where $C=(m_o/R)^{1/2}$, m_o molecular weight of the undissociated gas, R universal gas constant, A^* the cross-sectional area of nozzle throat and:

$$F(\gamma)=\sqrt{\gamma\cdot\left(\frac{2}{\gamma+1}\right)^{\left(\frac{\gamma+1}{\gamma-1}\right)}} \qquad (7)$$

With the total temperature and the chemical composition now calculated, it is a simple matter to compute the centerline total enthalpy of the flow.

d) Nozzle

Available experimental data are :
-total mass flow
-averaged total enthalpy at the nozzle exit
-pressure at the nozzle exit
-nozzle geometry

On the basis of the sudden freezing effects discussed above, we assume the flow to be frozen and during expansion in the nozzle.

Also we note from experimental measurements that the enthalpy decrement in the nozzle is generally small (from 10 to 20% max); this fact allows to consider the expansion of the flow in the nozzle near isentropic.
In this way, starting from the measured total enthalpy and total pressure we can compute all thermofluidynamic parameters at nozzle exit.

5 Results

5.1 General Flow Diagnostics

A certain number of runs with SPES has been made to validate the flow model, see the following table.

HN	Hah	HM
MJ/kg	MJ/kg	MJ/kg
5,88	8,38	6,71
6,86	10,08	8,07
7,73	12,08	9,66
8,35	13,52	10,81
9,26	15,75	12,60
8,33	13,31	10,65
9,61	15,64	12,51
10,81	18,13	14,51
12,43	21,07	16,86

As one can see from figure 4 the total pressure in all runs is under 50.000 [Pa], validating the ipothesys of frozen flow.

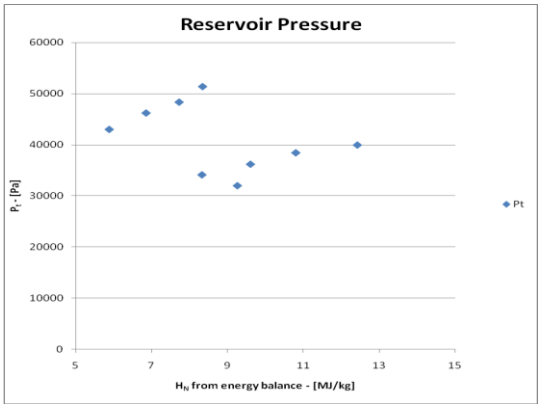


Fig 4 Total pressure measurements

Also in figure 5 we can see the centerline enthalpies as calculated following Pope and Park suggestions; the agreement is good, considering the uncertainty of the energy balance (up to 20%).

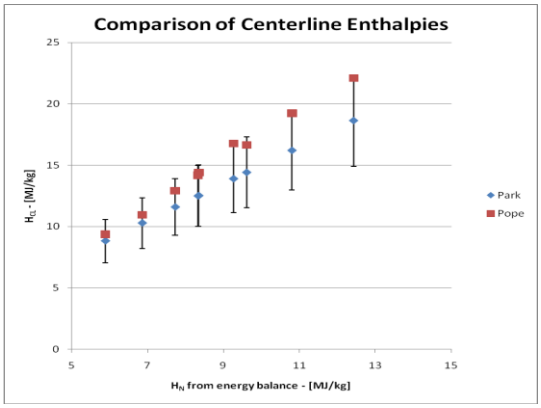


Fig 5 Centerline Enthalpy Measurements

5.2 Flow Diagnostics for Ablation Tests

The main aim of the arc-jet test campaign was to evaluate the thermo-ablative behavior of an AVIO material, when exposed to heat load conditions comparable to the ones specified for IXV ablative Thermal Protection System (TPS).

The dimensioning conditions of IXV Ablative TPS are hereafter reported, in terms of:
o Heat flux time history (Figure 6a)
o Pressure time history (Figure 6b)

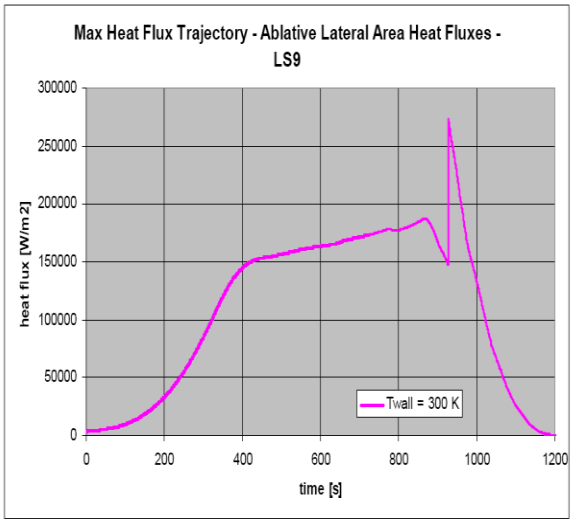


Figure 1b: Dimensioning heat flux time history

Fig. 6a IXV heat flux history

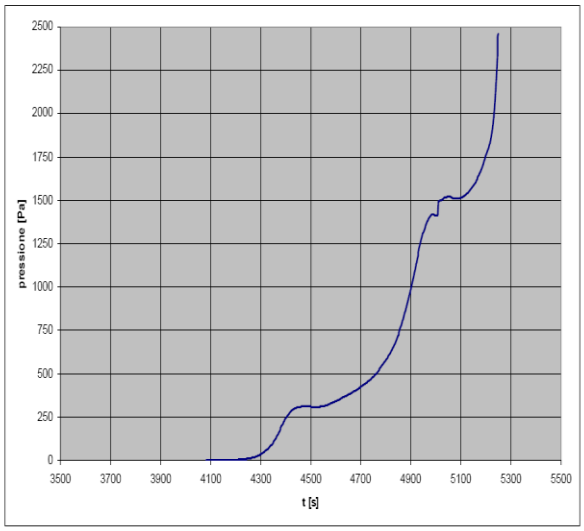


Figure 2: Dimensioning pressure time history

Fig. 6b IXV impact pressure history

In order to select in the SPES operating envelope the points that allow for achieving the aforementioned conditions, a dedicated calibration activity has been performed, by means of a slug calorimeter and a water-cooled Pitot probe. The result is that the maximum heat load of 140000 kJ/m2 acting on the ablative TPS can be reproduced in the arcjet facility, applying a 300 kW/m2 heat flux for about 9 minutes, with a resulting impact pressure on the sample equal to 2900 Pa. About the design heat flux value (300 kW/m2) it is the minimum value that ensures stable and safe operations of the plasma facility. The setting of the heat flux value automatically fix the resulting impact pressure value (2900 Pa), because of the fixed value of electrical power and mass flow in the plasma jet.

For the specified Test Requirements :

qs = 320 [kW/m2] stagnation-point heat flux
ps = 2900 [Pa] impact pressure

the more suitable Facility Test Parameters are :

Nitrogen mass flow : 0.4 [g/s]
Oxygen mass flow : 0.1 [g/s]
Electrical Arc Power : 9000 [W]

that, applied to the facility, lead to the following measured Thermofluidynamic (TFD) Test Parameters :

Hah = 13.1 [MJ/kg] (plasma torch)
Hm = 10.5 [MJ/kg] (mixer)
Hne = 7.5 [MJ/kg] (nozzle)
pm = 29000 [Pa]
pne = 160 [Pa]
pts = 50 [Pa] (test section)

using these together the numerical code, we obtain the calculated TFD parameters :

Total temperature : 5876 [K]
Total pressure : 29.000 [Pa]
Gamma : 1.482
Mach number : 4.01
Centerline Enthalpy : 13.95

Molecular Nitrogen mass fraction: 0.727
Atomic Nitrogen mass fraction : .073
Molecular Oxygen mass fraction : .004
Atomic Oxygen mass fraction : .196

A fist check for the numerical code is made comparing the calculated centerline enthalpy with the one obtained by the well-assessed formula :

Centerline Enthalpy = 1.5Averaged Enthalpy

as recently verified by Park on the basis of spectrometically determinated centerline enthalpy values [3].

For this case the check gives :

H_{CL} (Pope) = 13.95 [MJ/kg]
H_{CL} (Park) = 11.25 [MJ/kg]

A second check to verify the goodness of calculations is on the heat flux value. We follow the simple method proposed by Zoby [4], which suggests an empirical relation in the form :

$$\dot{q}_{fc} = K \cdot H_{ne} \cdot \sqrt{p_s / R_{eff}} \qquad (8)$$

where $\mathbf{R_{eff}}$ is the effective nose body radius [cm], $\mathbf{p_s}$ is the impact pressure [atm], **K** is a dimensional constant and **Hne** is the total enthalpy [MJ/Kg]. The constant **K** for Air is 117. $\mathbf{R_{eff}}$ can be assumed as 2.9 times the body radius as stated by Pope [2].
With the our-case values : Hne = 7.5 [MJ/kg] - ps = 2900 [Pa] - Reff = 0.058 [m] we obtain : $q_{s,fc}$ = 614 [kW/m2]. This heat flux value is for a fully catalytic wall condition; such is not the case of the copper calorimeter, then we must correct the heat flux value using the well-know Goulard theory [5]; with the suitable correction, the calculated heat flux becomes q_s = 363 [kW/m2], that is close to the experimental one.

A third check is on chemical composition; in the simple model implemented only nitrogen and oxygen mass fractions can be extimated, but for the experimental conditions considered the percentage of other species is expected to be very small. We use a commercial CFD code to evaluate the chemical composition for the test conditions, and the results are the following :

Molecular Nitrogen mass fraction:	.71
Atomic Nitrogen mass fraction :	.08
Molecular Oxygen mass fraction :	0.05
Atomic Oxygen mass fraction :	0.15

6 Ablation Tests

In the stagnation point configuration the test samples are placed normally to the flow. In the test section, located 63 (mm) downstream the nozzle exit, a model support system (MSS) supporting the ablative model and the calorimeter is mounted, figure 7. Calorimeter and test model have the same geometry.

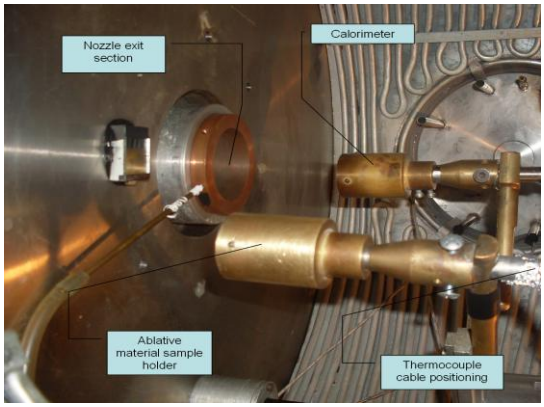


Figure 7 . Stagnation-Point Configuration

Test models have a maximum diameter of 40 mm; this value is given as the superposition of the ablative sample, an insulating material surrounding the sample and the metallic case to fix the model to the MSS. The test begins placing the calorimeter in the test section, in order to measure the incident heat flux. After this calibration phase, the calorimeter is removed and replaced by the ablative material model.

The material sample has been instrumented with six thermocouples, see fig.8:

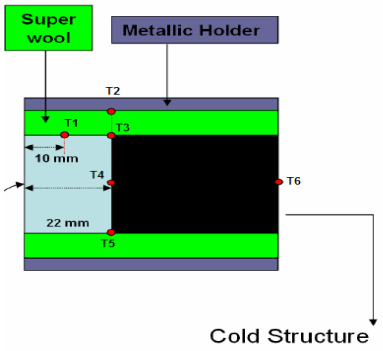


Fig. 8 Thermocouples positions

The same test conditions have been repeated four times in SPES. From Fig. 9 to Fig. 14 the acquisition of thermocouples is showed. Each color refers to a test.

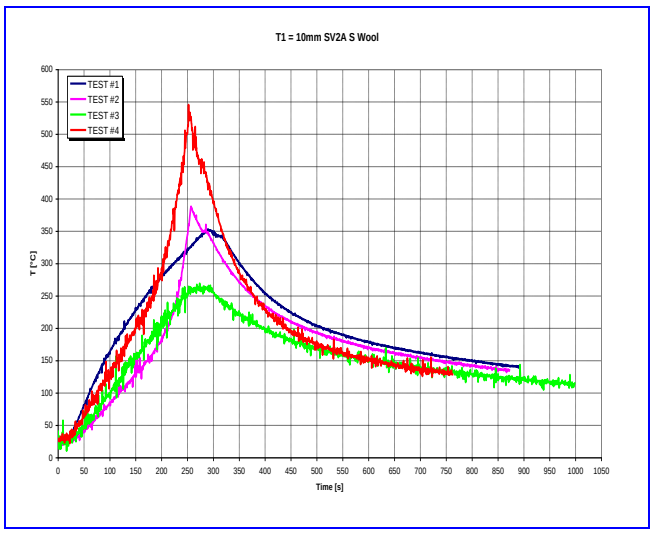


Fig. 9 Acquired temperature at TC 1

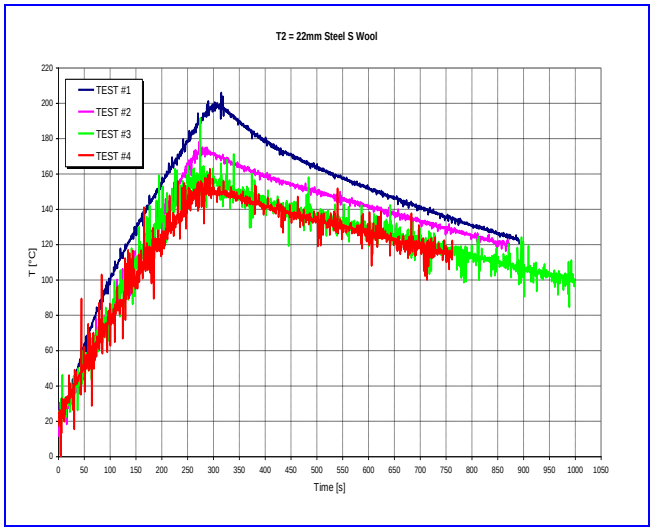


Fig. 10 Acquired temperature at TC 2

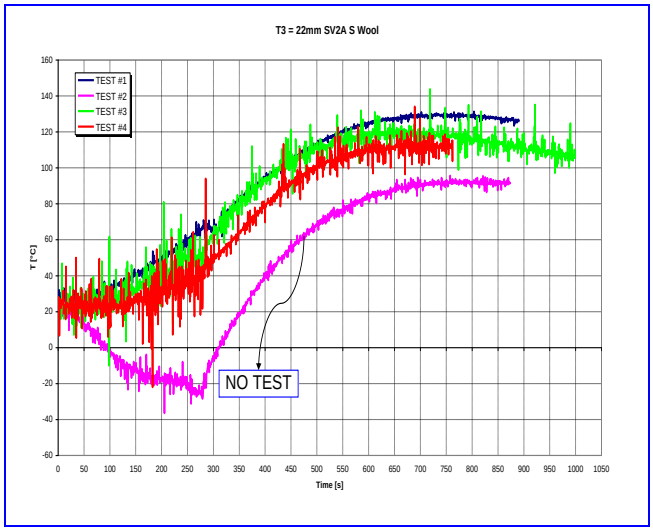


Fig. 11 Acquired temperature at TC 3

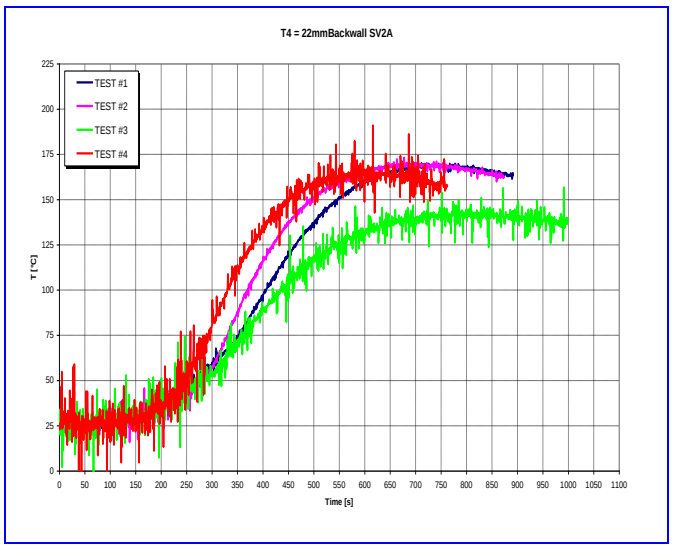


Fig. 12 Acquired temperature at TC 4

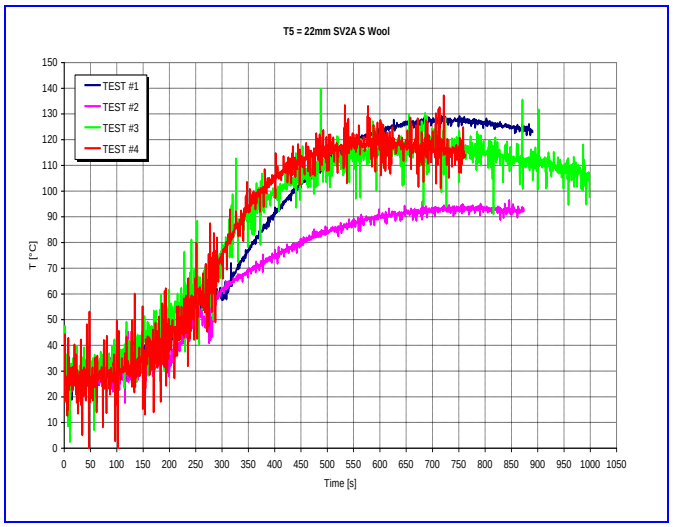


Fig. 13 Acquired temperature at TC 5

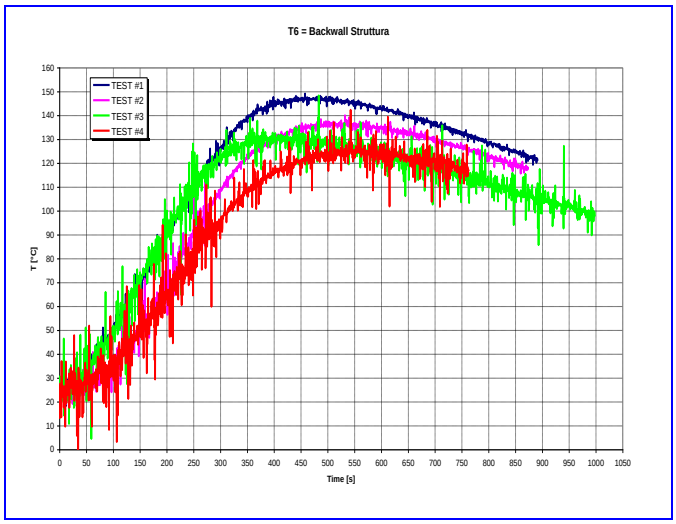


Fig. 14 Acquired temperature at TC 6

In figures 15 – 20 the test sequence is depicted.

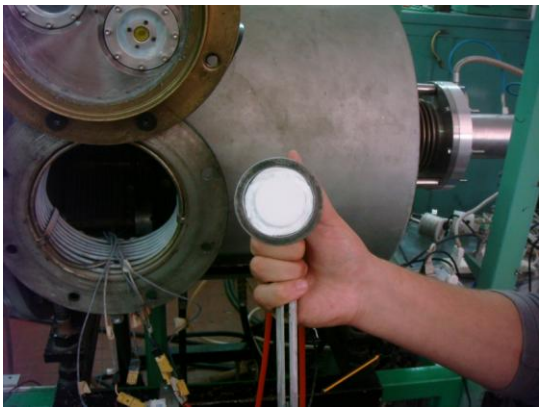


Fig. 15 Model ready to be mounted



Fig. 18 Model after test



Fig. 16 Model mounted



Fig. 19 Sample extracted from the model

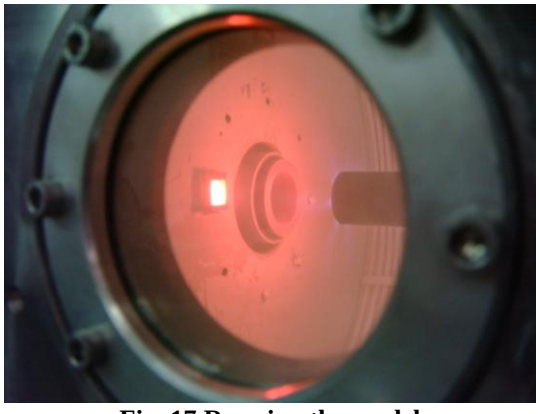


Fig. 17 Running the model



Fig. 21 Sample (detail)

As can be observed, a swelling phenomenon occurs during the test, due to the fact that the material expands with the carbonization.

6 Conclusions

A good characterization of a high enthalpy hypersonic flow is very important for TPS testing with arc-jet facilities. The process engineer need of a simple and fast diagnostic tool to evaluates all aerothermodynamic flow properties by using the few experimental measurements normally available during a test. In the paper we described the tool implemented and used for ablation tests, performed at the SPES facility at the Department of Aerospace Engineering of University of Naples, for AVIO company, to evaluated performances of its thermal protection material for the European space vehicle named IXV.

References

[1] ASTM Standard E341-81 "Standard practice for measuring arc gas enthalpy by energy balance"

[2] R. B. Pope “Measurements of enthalpy in low-density, arc-heated flows”, AIAA Journal vol.6, N.1, Jan. 1968, pp.103-110

[3] C. Park et alii, “Comparison of Enthalpy Determination Methods for Arc-Jet Facility” AIAA Journal of thermophysics and heat transfer, Vol. 20, No. 4, October–December 2006

[4] Zoby, Ernest “Empirical stagnation point heat transfer relation in several gas mixtures at high enthalpy levels” NASA TN D – 4799, October 1968

[5] R. Goulard, “On catalytic recombination rates in hypersonic stagnation heat transfer”,Jet Propulsion, Vol.28,No. 11, Nov. 1958, pp. 733-745

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Assessment of an improved Myers model for airfoil icing

E. Borghi, A.Guardone and G. Quaranta
Politecnico di Milano, Dipartimento di Ingegneria Aerospaziale, Milano, Italy

Keywords: Ice accretion, Rime and glaze ice, Messinger model, Myers model

Abstract

The problem of the ice accretion over an airfoil is studied numerically in both wet and dry regimes. The code developed starts with flow field computation using a potential method, which is coupled to an integral method for solving the boundary layer equations. The second step is the determination of droplets trajectories using a Lagrangian approach and then the collection efficiency. The convective heat transfer coefficient is determined in the laminar and turbulent case from an integral boundary layer solution that accounts for the roughness due to the ice formation. The last step is the numerical solution of the local thermodynamic balance and the determination of the ice accretion rate. In this work are described three fundamental steps that allows the icing accretion determination: the particle tracking and the successive determination of the collection efficiency, the solution of mass and energy equations and the geometry updating. Results are presented for ice accretion on a NACA 0012 airfoil and compared with results from other codes. The numerical solution of mass and thermal balances are performed using a non-uniform grid that allows to have better discretizazion near the stagnation point of the airfoil. Numerical results compare fairly well with the reference cases in the open literature.

1 Introduction

Ice growth on aircraft is a fundamental problem in the aerospace industry: various in-flight incidents have been shown to be directly attributable to the ice accretion. For this reason in last years this topic become the centre of numerous researches and various methods have been developed to study this problem. Icing prediction algorithms are generally divided into two major component. First, the air flow and the particle trajectories are computed, to become the inputs for the second step in which the ice accretion and the water flow are determined. The simulation of the ice accretion is an iterative procedure that consist of the successive computation of the air flow, water droplets trajectories, collection efficiency, convective heat transfer coefficient and finally the determination of the ice formation.

In this work the aerodynamic model is based on the approximation of the potential flow: the flow field around the airfoil is computed using the Hess-Smith panel method with an integral boundary layer correction, as described in [6]. Water droplets trajectories are determined using a Lagrangian approach. Droplets trajectories are then used to determine the collection efficiency and the incoming quantity of water that hits the airfoil surface. The cornerstone of most icing codes is the simple one-dimesional equilibrium energy balance first described by Messinger [1]. The thermodynamic model relies on the

heat and mass balance equations and it is enforced on each panel to determine the quantity of ice formed. During ice accretion roughness varies over the ice and airfoil surface: the roughness of the surface influences the boundary layer growth which in turns affects the convective heat transfer coefficient and hence the ice shape. This parameter is taken into account in the numerical simulations presented in this work by resorting to the empirical Shin and Bond correlation [4].

In the process of the ice accretion two different regimes of growth can be singled out: dry growth and wet growth. The first one occurs at lower temperatures and liquid water content: in these conditions each impacting droplet freeze nearly immediately on impact. This regime is called rime ice, it is characterized by the absence of the water, it has low density and the appearance of the ice is white and opaque. On the other hand during the wet growth the droplets are unable to release all latent heat of fusion at the impact: they remains over the surface in liquid form and flow along the airfoil. This is the case of the glaze ice in which water appears together with ice. The resulting ice has an higher density and it appears transparent and glossy. Glaze ice is favorite by warmer temperatures and higher liquid water content. The typical situation is the initial growth of the rime ice which turns into glaze ice when the ice thickness is so large that the temperature at the ice-water boundary reach that of the freezing temperature. The simulation code has to be able to treat also the intermediate situation, when some panels of the discretization are under rime ice condition while the water appears over others. In the case of the rime ice accretion the ice thickness can be easily determined through a mass balance while when the ice thickness is covered by a thin film of liquid water is necessary to consider a model that represents water flow and accretion.

In this latter situation, the first possibility is to follow the model developed by Myers [2]: in this case conduction through water is neglected, water is considered to have the freezing temperature and its accretion is not performed. This model take into account the water in the thermal balance by considering the evaporating heat flux and the heat flux entering the control volume and leaving it. The thermal flux that enters into the control volume is due to the mass of water that flows from the previous volume. The energy equation and mass balance are combined to provide a single first order non linear differential equation for the ice thickness, which is solved numerically. Another possibility is to account for the dynamics of the water and the flow of the water over the airfoil towards the leading edge. In this case the problem is governed by coupled mass and energy balances. The numerical solution gives not only the ice height but also the water layer thickness. The model used for solving the glaze condition is valid for thin water layer.

In this paper the determination of the collection efficiency and the numerical solution of the mass and energy equations using the improved Messinger model over realistic aeronautical geometries is presented. The results presented are obtained neglecting the water dynamics: this is the model followed by other existing icing codes used to validate the results [4, 8, 9].

2 Droplets trajectories computation

In the Lagrangian approach used in the present code, the trajectory of each droplet is computed based on a force balance: the water droplet is considered as a rigid sphere with a diameter ≤ 500 μm subjected to drag, gravitational an inertial forces. For two dimensional case the balance reads:

$$\begin{cases} m_d \frac{d^2 x_d}{dt^2} = -D\cos(\gamma) + m_d g \sin\alpha \\ m_d \frac{d^2 y_d}{dt^2} = -D\sin(\gamma) - m_d g \cos\alpha \end{cases} \quad (1)$$

where

- m_d droplet mass
- x_d, y_d position of the droplet
- t time
- D drag force

- g gravity acceleration
- α angle od attack
- γ angle between the local droplet velocity and the air velocity, namely

$$\gamma = \arctan\left(\frac{V_d - V_a}{U_d - U_a}\right)$$

where U_d, V_d, U_a, V_a represents respectively the components of the droplets and the air velocity in the x and the y direction, as in figure 1.

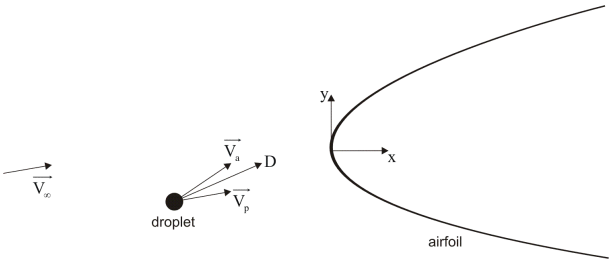


Figure 1: Velocities definition

The drag force acting on the droplet is expressed as a function of the drag coefficient C_d for a spherical drop. The expression used in this work is the Shiller-Neumann coefficient, which can be found in [5] and it is based on the droplets Reynolds number Re as follows:

$$\begin{cases} C_d = 0.4 & Re_d > 1000 \\ C_d = \frac{24}{Re_d}\left(1 + 0.15\, Re_d^{0.687}\right) & Re_d < 1000 \end{cases}$$

The drag force is written for a droplet of area A_d as:

$$D = C_d\, \rho_a\, \frac{V_\infty^2}{2} A_d$$

The system of equations (1) is then solved numerically using a BDF scheme and needs initial conditions: droplets are released at a distance from the airfoil so that the flow is not affected by the perturbation due to the presence of the body. The initial condition for droplets velocity is then equal to the free stream velocity components.

Trajectories are computed to define the impingement limits on the airfoil surface. Each trajectory is computed until it intersects the body or passes over it. The impingement limits are found following an iterative bisection process. two trajectories are chosen: one hitting the surface and the other that does not intersect the body. The new trajectory computed stay halfway between these two and if it impacts on the airfoil it became the new hitting trajectory while if it misses the body it is considered as the new trajectory that does not hit the surface. This procedure is repeated until the gap between the initial position of two droplets is smaller than a specified tolerance.

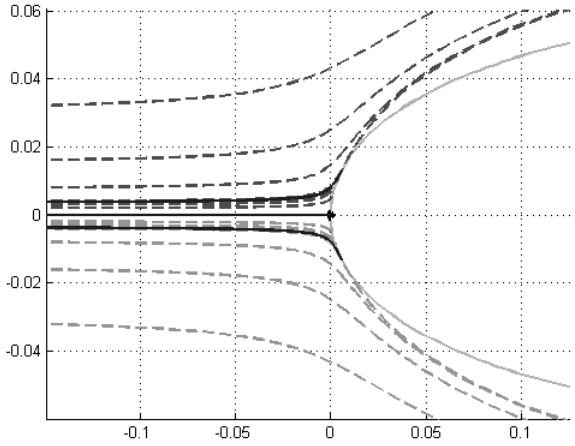


Figure 2: Search for impingement limits

Once the impingement limits are computed it is possible to perform an arbitrary number of trajectories that hit the airfoil: for each of this trajectory are known respectively the release and the impingement position. This allow to determine the collection efficiency, defined as (cf. figure 3):

$$\beta = \frac{\Delta y_0}{\Delta s} \tag{2}$$

In order to perform the collection efficiency β is necessary to know the velocity field around the airfoil: the air velocity is computed for all the droplet positions along its trajectory starting from the panel method used to solve the flow field. In figure 4, the predicted collection efficiency for the NACA 0012 is compared to that

obtained by available models in the literature. The agreement is fairly good and confirms the correctness of the proposed scheme.

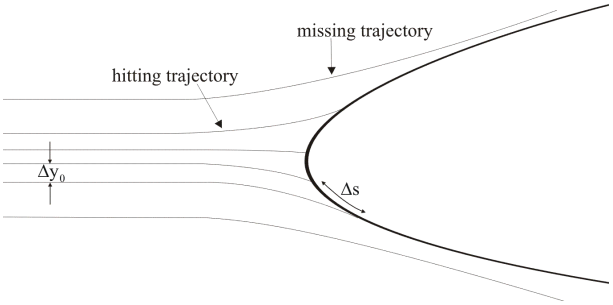


Figure 3: Collection efficiency definition

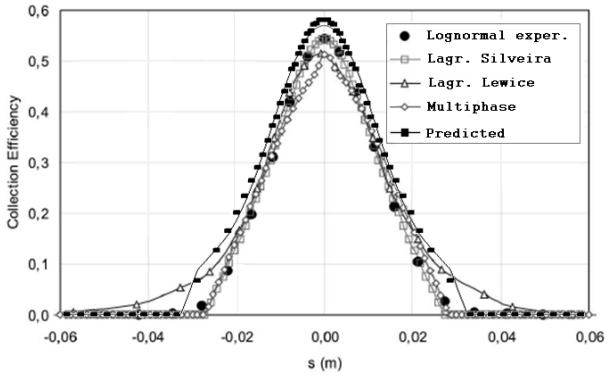


Figure 4: Collection efficiency for NACA0012

3 Mathematical problem formulation

The computation of the ice accretion is made here following the Myers model [2]. In this model, ice and water layers are considered separately. Conduction through the ice layer and phase change are also modeled. At the moving water ice interface the energy balance lead to a phase change or Stefan condition [7], which take the form of a first-order ordinary differential equation.

The typical situation investigated in this work is represented in the figure 5: an ice layer of thickness B rest on the top of the solid surface. In the case of the glaze ice the ice thickness is covered by a water thickness. The model presented is that for the glaze ice: in the case of rime ice the equations used are still true setting the water thickness h to zero.

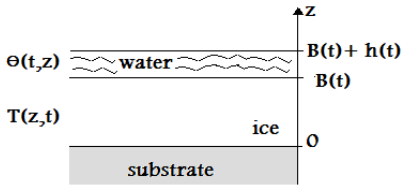


Figure 5: Scheme of the ice-water condition

The problem is governed by a system of four equations: mass balance, heat equation at the water and the ice interface and a phase change at the ice-water interface.

$$\begin{cases} \dfrac{\partial T}{\partial t} = \dfrac{k_i}{\rho_i\, c_i}\, \dfrac{\partial^2 T}{\partial t^2} \\ \dfrac{\partial \theta}{\partial t} = \dfrac{k_w}{\rho_w\, c_w}\, \dfrac{\partial^2 \theta}{\partial t^2} \\ \rho_i\, \dfrac{\partial B}{\partial t} + \rho_w\, \dfrac{\partial h}{\partial t} = \beta\; LWC\; V_\infty \\ \rho_i\, L_F\, \dfrac{\partial B}{\partial t} = k_i\, \dfrac{\partial T}{\partial z} - k_w\, \dfrac{\partial \theta}{\partial z} \end{cases} \tag{3}$$

where

- T, θ: ice and water temperature
- k_i, k_w: ice and water thermal conductivity
- c_i, c_w: ice and water specific heat
- ρ_i, ρ_w: ice and water density
- LWC liquid water content of the air
- V_∞ free stream velocity
- z direction normal to the surface
- L_F latent heat of fusion

The system of equation above requires one to specify suitable boundary and initial conditions as follows:

- the substrate is initially clean

$$B = h = 0 \qquad \text{at} \qquad t = 0 \tag{4}$$

- at the substrate the ice is in perfect thermal contact with the surface and so the temperature has the same value of the substrate temperature T_s

$$T\,(0,t)) = T_s \tag{5}$$

- the temperature is continuous at the ice-water interface and reach the freezing value $T_f = 273.15\ K$

$$T\,(B,t) = \theta\,(B,t) = T_f \tag{6}$$

- the last condition is the imposition of the Fourier's law at the ice-water interface. This is equivalent to impose the balance of thermal fluxes Q. In the case of the rime ice the same equation is written for the ice temperature at the ice-air interface.

$$\begin{aligned} -k_w \tfrac{\partial \theta}{\partial z} = Q_e + Q_{inc} + Q_c + Q_{out} \\ -\,(Q_{visc} + Q_k + Q_{ent}) \end{aligned} \tag{7}$$

Equation 7 is an example of the basic Messinger energy balance, where

  - convective heat transfer at the water surface

$$Q_c = H_w\ (\theta(B+h,t) - T_\infty) \tag{8}$$

    where H_w is the convective heat transfer coefficient

  - cooling of the incoming droplets

$$Q_{inc} = LWC\,\beta\,V_\infty c_w\,(\theta(B+h,t) - T_\infty) \tag{9}$$

  - heat lost due to the evaporation that in this work is expressed as a function of the saturation pressure P_{sat}

$$Q_e = 0.7 H_w L_E \left[\tfrac{\Phi\ P_{sat}(T_\infty) - P_{sat}(\theta(B+h))}{P_\infty c_a}\right] \tag{10}$$

    - L_E latent heat of evaporation
    - ϕ relative humidity

  - heat flowing out of the control volume with the liquid mass water

$$Q_{out} = \dot{m_{out}}\,\theta(B+h)\,c_w \tag{11}$$

  - heat due to the kinetic energy of the incoming droplets

$$Q_k = LWC\,\beta\,V_\infty \frac{V_\infty^2}{2} \tag{12}$$

  - aerodynamic heating

$$Q_{visc} = r\,H_a\,\frac{V_\infty^2}{2\,c_a} \tag{13}$$

    - r recovery factor
    - H_a air convective heat transfer coefficient
    - c_a specific heat of the air

  - heat flowing into the control volume from the preceding

$$Q_{ent} = \dot{m_{ent}}\,\theta(B+h)\,c_w \tag{14}$$

The system of equation (3) is replaced with a simpler one under the consideration that the ice growth rate is considerably slower than the heat conduction rate: the temperature can be assumed to be in equilibrium as the ice slowly accumulates. The problem of the ice accretion can be solved under the pseudo-steady approximation: with the thickness of the ice layer slowly growing with time. With this approximation the temperature profile is linear in the z direction. Another approximation made in this model is that to neglect the conduction through the water layer: the water temperature is constant and has the value of the freezing temperature $\theta\,(B+h) = T_f$. With the introduction of these simplification, applying initial and boundary condition described and following the procedure describe by Myers [2] is possible to gain the final equation in which the only unknown is the growth rate of the ice. This equation have to be coupled with a mass balance giving the final system of equations:

$$\begin{cases} \rho_i\,L_F\,\frac{\partial B}{\partial t} = \frac{k_i\,(T_f - T_s)}{B} \\ \qquad\qquad + (Q_e + Q_{inc} + Q_c + Q_{out}) \\ \qquad\qquad - (Q_{visc} + Q_k + Q_{ent}) \\ \dot{m_{ent}} + \dot{m_{inc}} = F\,(\dot{m_{ent}} + \dot{m_{inc}}) + \dot{m_e} + \dot{m_{out}} \end{cases} \tag{15}$$

where

- $\dot{m_e}$ mass rate of evaporation
- $\dot{m_{ent}}$ water mass flowing into the control volume
- $\dot{m_{inc}}$ mass of incoming water droplets
- $\dot{m_{out}}$ water mass flowing out the control volume

The freezing fraction F is defined as the ratio between the mass of water that freeze in the control volume and the total mass of water entering. In the case of the rime ice it is set at a constant value equal to one while in the glaze condition is $0 \leq F \leq 1$.

The numerical solution of the system (15) allows to compute the ice thickness. In this work the time step used for the numerical integration of the equations is 10^{-3} seconds.

4 Geometry updating

The updating of the airfoil geometry to account for the ice thickness is a very important issue in an icing code. The ice accretion analysis follows the progressive accumulation of the ice over the surface for a finite time: it is an iterative procedure that requires a regular updating of the geometry. The presence of the ice near the stagnation point modifies the shape of the surface and this has effects on the flow field around the airfoil such as on the impingements points of the droplets that hits the surface and so the change in the shapes of the body influences the collection efficiency. Before starting the simulation the time step for the updating is chosen: the entire temporal interval is divided into smaller parts and at each time-step the complete updating is performed starting with the determination of the flow field for the new shape, passing through the trajectories calculation and the definition of the collection efficiency and arriving at the new computation of the ice thickness. The choice of the updating time step is important because this parameter has a great influences on the final result: smaller time steps allow to have accurate conditions for the ice calculation and this produces more realistic results with an high level of precision. On the other hand the continuous update of the flow field, trajectories and the ice determination leads to a great computational cost for the icing code: the choice of the time step for the geometry update is a key point in the icing code because from this parameter depends the accuracy of the result and the computation time, see figure 6.

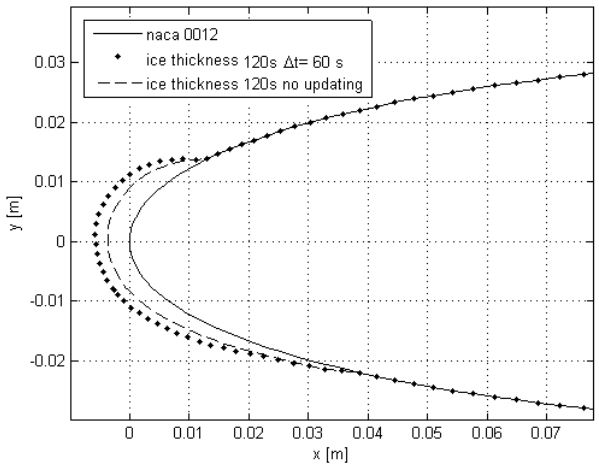


Figure 6: Ice accretion on NACA0012 after 120 s: comparison between ice thickness obtained without the geometry update and the accretion obtained updating the geometry after 60 s

In this work the computation of the flow field is made with a panel method and this requires a smooth geometry: at each time step the new geometry for the aerodynamic model is obtained connecting the ice thickness with the airfoil so that the edges between the ice and the body are deleted as shown in fig. 7.

The results presented in the next section are obtained using a time step for the upgrading of 60 second.

5 Results

The results obtained with the code illustrated in the paper are validated by comparison against experimental and those obtained from other existing icing codes such as LEWICE [4], [10], DRA code [10], CANICE [11] and CIRA code [8]. Various atmospheric conditions such as different flight conditions are considered. In this

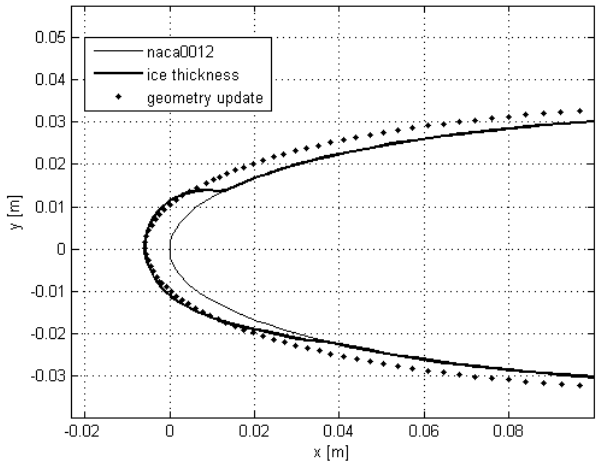


Figure 7: Geometry updating: connection between the ice thickness and the airfoil

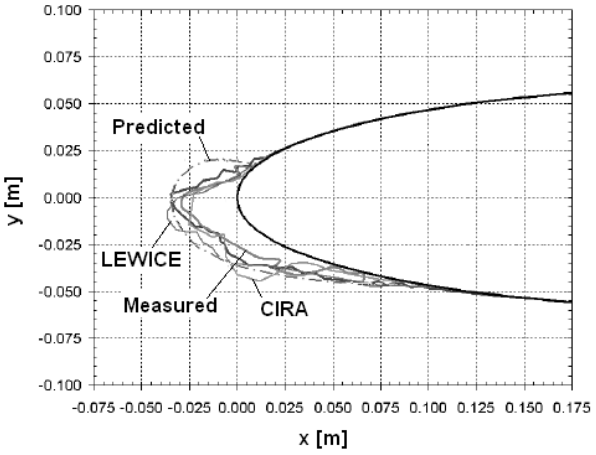


Figure 8: Comparison among the code presented, experimental results, LEWICE and CIRA code. Analysis refer to the conditions illustrated in the table 1

paper only two test cases are analyzed: the main difference between them, whose conditions are illustrated in the table 1 and 2, is the angle of attack which is null in the second case and equal to 4°in the first.

Accretion time	360 [s]
Angle of attack	4°
Chord	0.5334 [m]
Airspeed	67.05 [m/s]
Atmospheric temperature	261.9 [K]
Atmospheric pressure	101300 [Pa]
Liquid water content	1 [g/m^3]
Droplets diameter	20 [μm]

Table 1: Definition of test case 1

Accretion time	120 [s]
Angle of attack	0°
Chord	0.152 [m]
Airspeed	130.5[m/s]
Atmospheric temperature	260.7 [K]
Atmospheric pressure	90500 [Pa]
Liquid water content	0.5 [g/m^3]
Droplets diameter	17.5 [μm]

Table 2: Definition of test case 2

Figure 8 shows a good agreement between the ice accretion predicted and those computed by other icing codes. The most relevant difference is the profile of the ice shape: the geometry of the predicted ice is smoother than the others. This is attributable to the ice roughness models used in the existing codes, which are more accurate and produces more realistic ice shapes.

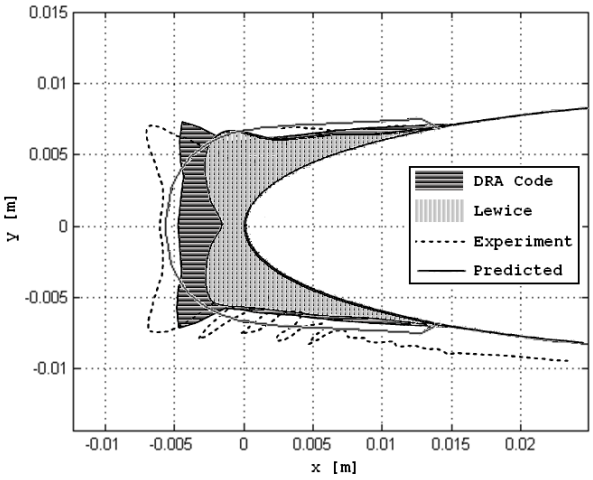


Figure 9: Comparison among results obtained from the code presented, experimental results and those presented by the LEWICE and the DRA code under conditions illustrated in the table 2

In figure 9 the ice thickness predicted has the same order of that obtained using other codes but the shape of the accretion is quite different for the lack of the horns that characterized data

gained through experiments. This is believed to be related to the method used to upgrade the geometry at each time step that, as explained in the previous section make the geometry smooth, creating a new airfoil that contains the ice thickness. This is a limit of the code and the first aspect that have to be improved. Another important improvement for the icing code presented is the capability to consideResultr the dynamics of the runback water. The computation not only of the ice accretion but also of the water flowing and accretion could give a more precise understanding of the icing accretion problem and results that are closer to the reality.

6 Conclusions

The modified Meyers model, that was derived for a flat plate geometry, was applied to the prediction of ice accretion over a simmetric NACA 0012 airfoil at zero and at four degree incidence. The model was included into a very simple ice accretion code, which was developed with the purpose of evaluating different ice accretion models. Numerical results agree fairly well with those obtained experimentally and by other simulation codes, with the exception of the well-known ice horns that are the hallmark of ice accretion over airfoil. This is believed to be related to a lack of a sophisticated roughness model in our code and it will be the object of future studies.

References

[1] B.L. Messinger *Equilibrium temperature of an unheated icing surface as a function of air speed* Journal of the Aeronautical Sciences, pp. 29–42, January 1953.

[2] T.G. Myers *Extension to the Messinger model for aircraft icing*, AIAA Journal, Vol. 39, No. 2, February 2001.

[3] T.G. Myers, J.P.F. Charpin, C.P. Thomphson *Slowly accreting ice due to supercooled water impacting on a cold surface*, Physics of fluids, Vol. 14, No. 1, January 2001.

[4] J. Shin, T. Bond *Experimental and Computational Ice Shapes And Resulting Drag Increase for a NACA 0012 Airfoil*, NASA Technical Manual 105743, 1992.

[5] Y. Borgault, W.G. Habashi, J. Dompierre, G.S. Baruzzi *A Finite Element Method Study of Eulerian Droplets Impingement Models*, International Journal for Numerical Methods in Fluids, Vol. 29, pp. 429-449, 1999.

[6] J.L. Hess, A.M.O Smith *Calculation of Potential Flow About Arbitrary Bodies*, Progress in Aeronautical Sciences, Elmsford, New York, Pergmon Press, 1967.

[7] G.M. Hill *One - Dimensional Stefan Problems: An Introduction*, Longman Science Technical, Harlow, England, UK, 1987.

[8] G. Mingione, V. Brandi *Ice Accretion Prediction on Multielements Airfoils*, Journal of Aircraft, Vol.35,No. 2 , 1998.

[9] G. Fortin, A. Illinca, J.L. Laforte, V. Brandi *New Roughness Computation Method and Geometric Accretion Model for Airfoil Icing*, Journal of Aircraft, Vol.41 ,No. 1 , 2004.

[10] W.B. Wrigth, R.W. Gent, D. Guffond *DRA/NASA/ONERA Collaboration on Icing Research, Part II-Prediction of Airfoil Ice Accretion*, NASA Report 202349, Lewis Research Center, 1997.

[11] I. Paraschivoiu, F. Saeed, *Ice Accretion Simulation Code CANICE*, International Aerospace Symposium, Bucharest, Romania, Oct. 2001

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Experimental and Numerical Investigation on the Electro Hydro Dynamic Interaction generated by a Dielectric Barrier Discharge

A. Cristofolini, G. Neretti, F. Roveda and C. A. Borghi
Department of Electrical Engineering, University of Bologna, 40136 Bologna, Italy

Keywords: *weakly ionized plasma, dielectric barrier discharge, EHD interaction, schlieren diagnostic.*

Abstract

A surface dielectric barrier discharge (DBD) actuator has been experimentally investigated. An electrode pair separated by a dielectric sheet constitutes the actuator. Two types of dielectric materials and several a.c. supply conditions have been utilized to generate a surface dielectric barrier discharge with different characteristics. An Electro Hydro Dynamics (EHD) interaction was induced into still air and several fluid-dynamic regimes were obtained. Using electric and fluid dynamics measurements, the energy transfer mechanisms caused by the EHD interaction have been investigated. The visualization of the plasma boundary layer during the discharge ignition phase, characterized by hot vortexes, and during the steady regime has been obtained by means of Schlieren diagnostics technique. Vortex morphology and propagation velocities for both actuator types at all supply conditions have been evaluated. Pitot velocity profiles have been taken in the steady regime operation at several distances along a line perpendicular to the actuator surface. Along this line and as a function of the position the integral of the pixel intensities of the Schlieren image has been calculated. The function obtained matches with a good agreement the Pitot velocity profile for all distances and in all the supply conditions investigated. Numerical simulations were performed to validate theoretically this result. The calculations confirm the relationship between flow velocity distribution in the boundary layer and the gas density distribution.

1 Introduction

The interest in active flow control, i.e. the ability to directly manipulate a flow in order to create a desired change, has showed a quick growth in the last decade [1]. The application of a Dielectric Barrier Discharge (DBD) fluid-dynamic actuator in aerodynamic flow control, is in a continuous development. Many experimental works have dealt with the application of Electro Hydro Dynamic (EHD) actuators to be used in the aeronautical [2]-[11] and aerospace [12], [13] fields, and for the turbine blades[14], [15]. The development of the discharge and the establishment of the EHD interaction have been investigated in several numerical works [16]-[20] too. In particular, the ability of the actuator to prevent or to induce flow separation and the possibility to reduce the drag and to enhance the lift of airfoils have been demonstrated.

The plasma jet generated by a DBD fluid-dynamic actuator is thus able to alter the velocity inside the boundary layer and can modify the interaction between fluid and body surface (a wing or a blade) without moving parts. Hence the EHD phenomenon allows the direct conversion from electric into kinetic energy with response times that can be considered negligible compared to the characteristic times of the fluid dynamics. Furthermore the DBD actuator could lead to

reduce the size of the moving parts and their load, and thus to reduce the input power and to enhance the operational flexibility.

In this work we present an experimental investigation finalized to achieve a better knowledge of the momentum transfer process involved in the EHD phenomenon. The visualization of the wall-jet generated by the actuator during the ignition phase and the steady regime using a Schlieren technique has been carried out. The optical diagnostic is based on the variation of the refractive index of the medium in which the light beam propagates. The complete non-intrusiveness of this technique is the advantage of its use if compared to other diagnostics [21]. Pitot tube diagnostics allows the determination of velocities and velocity profiles. However it is an intrusive diagnostic and the cost of using it is linked to the flow perturbation introduced by the Pitot tube. Particle Image Velocimetry (PIV) needs seeding of the fluid. For a proper operation of PIV the right type and concentration of the seed can involve a long work. Moreover the seeding particles may get electrically charged and may affect the plasma dynamics and the gas flow. The Schlieren technique, applied to a DBD fluid dynamic actuator, is "self seeded" by the density gradients created by the hot wall jet [22], [23].

In a previous work [24] we pointed out that the velocity induced by a DBD actuator in the boundary layer, measured flush to the wall, is mainly related to the active power delivered by the power supply to the electrodes. A desired velocity can be reached with a defined amount of average power obtained with different supply voltage and frequency values. These results have been confirmed by measurements performed on actuators with different geometries, dielectric materials and a.c. supply conditions.

The transfer energy mechanism has been investigated also by means of the evaluation of the efficiency of the EHD interaction through electrical and fluid dynamics measurements.

2 Experimental setup

All the performed experiments have been realized with single surface Dielectric Barrier Discharge actuators immersed in still atmospheric pressure air. The actuator configuration is shown in Fig. 1. The electrodes are made of copper strips with a thickness of 0.08 mm. The upper electrode is 5 mm wide. Two different materials have been used as dielectric: a 2 mm thick teflon tape and a kapton tape electrode realized by 4 kapton layers, 60 µm thick each. With the teflon dielectric, the buried electrode is 7 mm wide. When the kapton tape is used the buried electrode is 10 mm wide. In both cases no gap is present. The spanwise length of the electrodes (In the z direction in Fig. 1) is 50 mm.

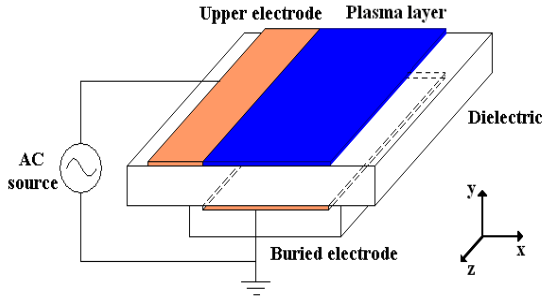


Fig. 1. DBD actuator

The actuators have been fed by the power supply the schematic of which is shown in Fig. 2. A HP-Agilent 33120A signal generator produces an a.c. signal delivered to an Elgar a.c. amplifier able to provide voltages up to 260 V with apparent powers up to 3 kVA. For increasing the voltage to the desired value a ferrite high voltage ferrite transformer is utilized.

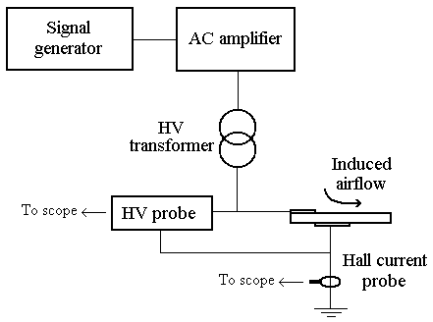


Fig. 2. Power supply electrical scheme

3 Diagnostics

The time behaviors of the electrode voltage drop voltage *v(t)* and of the electric discharge current *i(t)* have been measured. A Tektronix P6015A capacitive compensated high voltage probe, with a bandwidth up to 75 MHz, is used for *v(t)*. A Tektronix TCP312 Hall current probe, with a bandwidth up to 100 MHz, is utilized to measure *i(t)* (Fig. 2). Therefore the value of the instantaneous power $p(t)=v(t)i(t)$ to feed the discharge is derived. The average power absorbed is obtained as the average power in a waveform period.

Schlieren images have been taken by a Z-type configuration setup [21] (Fig. 3). The light source is a tungsten halogen low voltage lamp equipped with a rear reflector. The condenser is a Schneider-Kreuznach Xenon 40 mm double-Gauss lens with a *f/1.9* focal ratio. The light beam is reflected by two off-axis parabolic mirrors of a diameter of 138 mm and a *f/3.5* focal ratio. The light beam is in the *z*-direction (Fig. 1).

The knife edge was parallel to the actuator surface (x-direction in Fig. 1) so that the Schlieren diagnostic can record density gradients perpendicular to it. The images are detected by a PCO CCD camera equipped by a super video graphics array (SVGA) resolution and a pixel size of 6 x 6 μm^2. The peak quantum efficiency is 55 % at a wavelength of 380 nm. Velocity profiles have been determined by means of a glass Pitot tube moved by a step motor with linear resolution of 0.03 mm. The Pitot tube has been connected to a DCAL401 Sursense ultra low pressure sensor with a sensitivity of 32 mV/Pa.

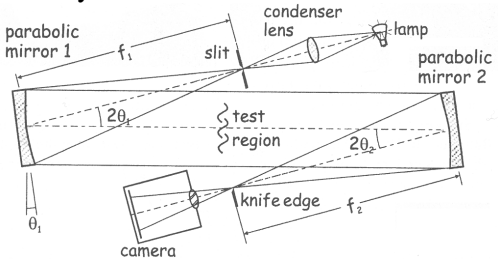


Fig. 3. Z-type configuration Schlieren setup

4 Experimental results

The electrical supply conditions have been chosen so that to deliver the maximal amount of the average power transferred to the discharge as high as possible, avoiding plasma inhomogeneities (Table 1). For each actuator type, different voltage-frequency combinations, which correspond to the same average power, have been selected. The average power values are referred to the beginning of the actuator life. As a matter of fact, when an actuator is operated for several tens of minutes, its electrical characteristics change. In particular, for a fixed voltage, the average power increases with the run time. This behavior could be related to the electrical surface conductivity increase due to chemical reactions activated by the plasma.

Table I. Electrical supply conditions

Average Power [W/m]	Dielectric material	Voltage peak to peak [kV]	Frequency [kHz]
96	teflon	35	5
		27	8
140	Kapton tape	14.5	5
		11.4	10
		10.4	15

In the kapton tape actuator the observed plasma covers the first 6 mm on the dielectric surface after the upper electrode. In the teflon actuator the plasma completely overlay the dielectric surface over the bottom electrode that is of 7 mm

Velocity profiles in the y-direction and the related standard deviations obtained in the teflon actuator, acquired with the Pitot probe at the end of the plasma region in the different supply conditions, are reported in Fig. 4. The "y" direction is perpendicular to the actuator surface as depicted in Fig. 1. The measurements start at a distance in the y-direction equal to 0.5 mm, which is equal to the distance between the Pitot tube axis and the dielectric surface when the probe is placed on it. Thus this distance corresponds to the "zero height" of the Pitot tube.

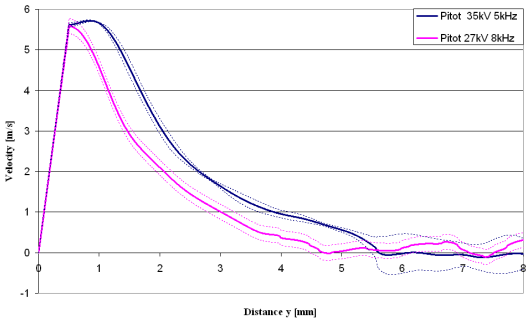


Fig. 4. Pitot velocity profiles and standard deviations for teflon actuator at the end of the plasma region.

Figure 4 shows as at "zero height" the velocities measured in the two supply conditions at the same average power are the same. However for the highest voltage condition the region in which the actuator expresses its action, is wider.

A similar behavior is obtained for the kapton actuator too. In Figure 5 the velocity profiles and the related standard deviations of the kapton tape actuator, measured by means of the Pitot probe placed at the end of the plasma region in different supply conditions, are shown.

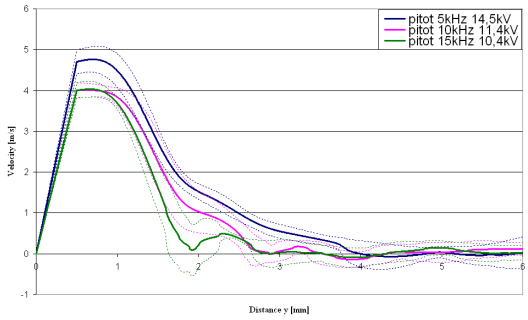


Fig 5. Pitot velocity profiles and standard deviations for kapton tape actuator at the end of the plasma region.

For the two supply conditions at low voltage, the speed at "zero height" is the same. The main difference is the thickness of the region in which the EHD interaction is observed. This region is larger for higher voltages. The supply condition at the largest voltage (14.5 kV) utilized, is characterized by the highest velocities measured and the largest region affected by the EHD interaction.

In order to consider the actuator performance with reference to the supply conditions related to the momentum transfer mechanism, the electro-mechanical efficiency η is defined. In still air, this quantity can be calculated as the ratio between the mechanical power P_{mec} and the electrical power P_{elec}:

$$\eta = \frac{P_{mec}}{P_{elec}} \qquad (1)$$

For both teflon and kapton actuators and for all the supply conditions used, the electrical power vales have been reported in Table 1. The mechanical power has been determined by using the Pitot velocity profiles $v_P(y)$ under the assumption of stationary flow. The mechanical power per unit length corresponds to the kinetic energy density flow rate [26] and can be expressed as

$$P_{mec} = \frac{1}{2}\rho\int_0^\infty v_p^3(y)\,dy \qquad (2)$$

where ρ is the air mass density.

In Table II the actuator efficiencies have been reported.

Table II. Electrical supply conditions and related actuator efficiencies.

Dielectric material	Voltage peak to peak [kV]	Frequency [kHz]	Efficiency %
teflon	35	5	$1.5\cdot10^{-1}$
	27	8	$0.86\cdot10^{-1}$
Kapton tape	14.5	5	$0.42\cdot10^{-1}$
	11.4	10	$0.25\cdot10^{-1}$
	10.4	15	$0.22\cdot10^{-1}$

These values are largely lower than 1 per cent. A large amount of average power is used for the ionization and the vibrational and the rotational excitation modes. The wall jet nature of the DBD fluidynamic actuator significantly reduces the surface velocity due to the skin friction. A third mechanism of the power absorption is the dielectric surface resistance as it was described at the beginning of this paragraph. Moreover an amount of average power is transferred to the plasma due to Joule heating.

As discussed above, for the selected dielectric material the efficiency is higher for higher applied voltages. The teflon actuator shows higher efficiencies than the kapton one.

This can be due to the different discharge regimes in the two dielectric cases. The inhomogeneity of the plasma is larger when the kapton tape is used. It is well known that, when a streamer behaviour is observed, the efficiency of the actuator decreases because a larger amount of active power is converted in heat and not in net thrust [24].

Visualization of the vortexes generated by the actuator during the ignition phase and the wall jet regime of the steady state, had been carried out by means of Schlieren imaging. In order to detect the vortex propagation behavior, the CCD camera has been triggered with increasing delay times. In this way the vortex behavior is obtained by several discharge ignitions. The validity of this measurement methodology has been checked. Several images, acquired at the same delay time and corresponding to different vortexes, coincide. The exposure time of the camera has been set to be 1 ms. A stable light source and the lack of external light radiation are essential conditions in order to achieve a uniform image background.

Schlieren images of the vortex evolution for the kapton actuator fed by 14.5 kV - 5 kHz is shown in Fig 6.

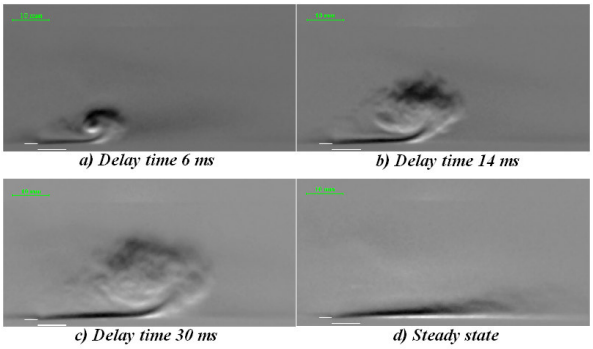


Fig. 6. Schlieren images of teflon actuator jet at different delay times.

The white lines at the bottom of each image show the electrode positions and the green line at the top of it represents a length of 10 mm.

Images a), b) and c) are referred to delay times of 6, 14 and 30 ms. For the teflon and the kapton actuator in all the supply conditions considered the vortex behaviour has been detected by the Schlieren techniques. Vortex propagation velocities in the range of 0.5÷1.5 m/s have been measured.

Vortexes appear during the transient following the ignition. After about 100 ms steady state regime is reached. During this regime vortexes are not observed. The image d) of Fig. 6 shows the region affected by the actuator in the steady state regime. A dark region over the upper electrode is observed. Here the gas density is lower due to electrode heating. This region is shown in the image a), b) and c) too. The EHD effect is present till the edge of the image, thus roughly up to 70 mm away from the upper electrode. A bright region between the actuator surface and the dark region within the jet core can be observed. As it will be shown later in the following, this area corresponds to the boundary layer region in which the velocity increase from zero to its maximum value.

The EHD effect caused by the actuator in the steady state regime had been investigated also. Pitot measurements on a line in the boundary layer perpendicularly to the actuator surface (y-direction of Fig. 1) at different distances in the x-direction from the upper electrode, have been performed. The velocity profiles, obtained by the Pitot technique, have been compared with the profiles obtained by the pixel intensities of the Schlieren image along the same line. These comparisons have been done for a teflon actuator supplied with a frequency of 5 kHz and four different supply voltages (Table III). The boundary layer profiles had been determined at 10, 20 and 30 mm from the upper electrode (Fig. 7).

Table III. Electrical supply conditions used for Schlieren-Pitot comparisons.

Condition	Voltage peak to peak [kV]	Average power [W/m]
A	20	20
B	25	66
C	30	94
D	35	140

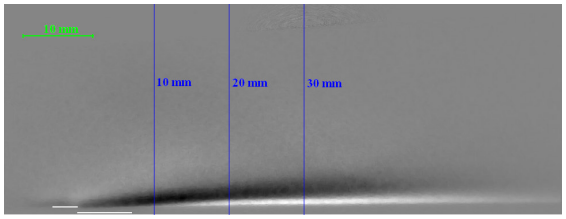


Fig. 7. Schlieren image at condition D. Outlined in blue are the three lines used for the comparisons.

The Schlieren images of the four supply conditions of Table III are shown in Figure 8.

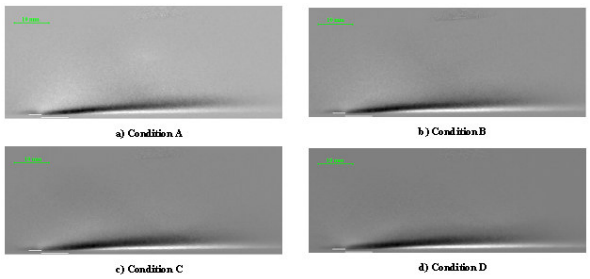


Fig. 8. Schlieren images of the wall jet produced by the actuator in the steady regime in the four supply conditions of Table III.

A comparison between the profile of the pixel intensities *I(y)* and the Pitot velocity profile along the line at 30 mm for the supply condition D of Table III is shown in Fig. 9.

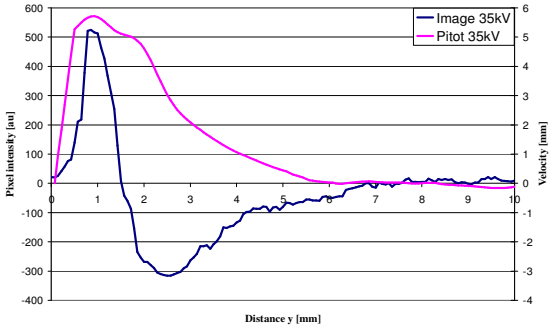


Figure 9. Pixel intensities and Pitot velocity profile along the line at 30mm for the supply condition D.

The distance "0" corresponds to the actuator surface. The Pitot measurements start at a distance of 0.5 mm which is the distance between the Pitot tube axis and the surface when the probe is placed directly against it. Both the optical (Schlieren imaging) and the fluid dynamic (Pitot technique) diagnostics clearly show that the region affected by the EHD interaction is about 7 mm wide.

The Intensity Integral function is defined as:

$$II(y) = -\int_{y}^{0} I(y')dy' \qquad (3)$$

where *I(y')* is the intensity of the at the distance *y'* from the actuator surface.

In Fig. 10 the comparison between the Pitot velocity profile and the profile obtained by the intensity integral *II(y)* at 30 mm from the upper electrode and at the supply condition D, is shown.

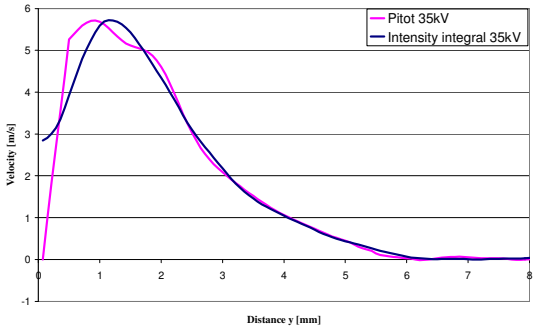


Figure 10: Comparison between the Pitot velocity profile and the intensity integral obtained for the supply condition D along the line at 30 mm

In order to compare the profiles, the intensity integral values have been normalized. The agreement between the two profiles is very good.

In Fig. 11 the comparison between the Pitot velocity profiles and the image intensity profiles along the line at 30 mm from the electrode and obtained for the supply conditions A, B and C of Table III is reported. The image integral values are all divided by the same constant value used for the comparison in Fig. 10. In these cases also a very good agreement is obtained.

The comparisons between Pitot velocity profiles and the intensity integrals at 20 and 10 mm from the electrode are shown in Figures 12 and 13 respectively. The intensity integral values reported here are normalized by means of the same constant used for *II(y)* of Fig. 10.

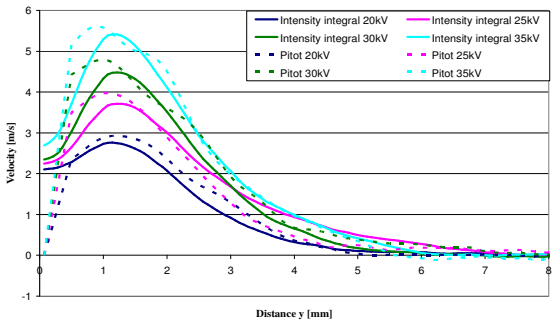


Figure 11: Comparison between the Pitot velocity profiles and the intensity integrals along the line at 30mm obtained for all the supply conditions

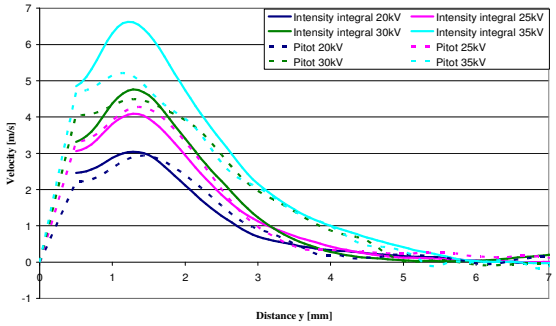


Figure 12: Comparison between the Pitot velocity profiles and the intensity integrals along the line at 20mm obtained for all the supply conditions

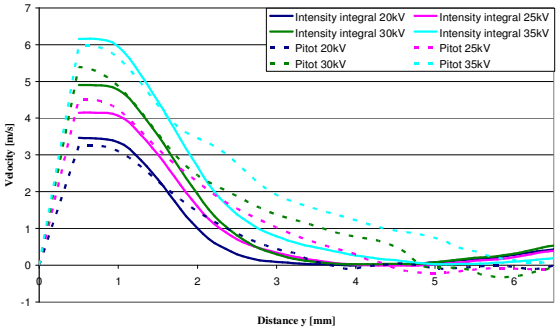


Figure 13: Comparison between the Pitot velocity profiles and the intensity integrals along the line at 10mm obtained for all the supply conditions

At 20 mm from the electrode (Fig. 12) the agreement is very good for conditions A and B. For condition C the two profiles are similar only within the firsts 2 mm. For condition D (where the highest voltage is utilized) close to the actuator surface the profile obtained by the intensity integral shows a higher velocity than that of the Pitot measurements.

At 10 mm (Fig 13) the profiles obtained by Pitot measurements and by the intensity integrals present very similar behaviors in the region within the firsts 2 mm from the surface. For y > 2 mm the Pitot velocity profiles have larger values. As a matter of fact here the Pitot measurements are less reliable because of the higher intrusiveness of the Pitot tube in comparison with the thin boundary layer thickness of the wall jet (compare Fig. 7).

As shown in Figures 11, 12 and 13 the intensity integral obtained from the Schlieren pixel intensity along a line perpendicular to the actuator surface is a good estimation of the velocity profile of the wall jet induced by the DBD actuator. The deviation of a light beam passing through the gas, and hence the refraction index, is due to mass density variations within it. Thus the light measured by the Schlieren diagnostic is related to the first derivative of the refraction index [21]-[25]:

$$I(x,y) \propto \int \left(\frac{\partial \rho(x,y,z)}{\partial x} + \frac{\partial \rho(x,y,z)}{\partial y} \right) dz, \qquad (4)$$

where $I(x,y)$ is the intensity of the pixel and the coordinates (x,y,z) are given in Fig. 1.

As the knife edge is positioned parallel to the x axis the Schlieren image intensity mainly depends on the density gradients along the y coordinate. Thus the derivative of the density along the x coordinate can be neglected. Moreover the velocity field produced by the DBD actuator can be studied as an almost 2D-field (i.e. no density variations are present in the z direction). Equation (4) becomes

$$I(x,y) \propto \Delta z \frac{\partial \rho(x,y,z)}{\partial y}, \qquad (5)$$

where Δz represents the width of the region in which the light beam is deflected. By integrating now both the terms of eq. (5) with respect to y' in the y-direction, it follows:

$$\int_0^y \boldsymbol{I(x,y')dy'} \propto \Delta z \int_0^y \frac{\partial \rho(\boldsymbol{x,y',z})}{\partial \boldsymbol{y'}} \boldsymbol{dy'} \quad . \qquad (6)$$

The left hand side of eq. (6) represents the intensity integral $II(y)$ at the distance x from the upper electrode. Equation (6) states that these

profiles are proportional to the mass density profiles of the air.

Due to the Schlieren facility construction the density variation in the y-direction only is recorded. Therefore the integral of the Schlieren image pixel intensities depicts the density distribution along the y-direction and can be directly related to the velocity distribution.

5 Numerical Simulations

Numerical simulations were carried out in order to verify the existence of the abovementioned relationship between gas velocity and intensity integrals. The opensource software OpenFOAM was used to perform this comparison. A turbulent transient CFD solver was modified to account for the body force generated by the EHD interaction and the ohmic heating due to conductive current in the discharge region. In eq. 7 ρ_{max} is the maximum value of the gas density perpendicularly to the actuator surface.

The quantity

$$\rho_{rel} = \rho_{\max} - \rho, \qquad (7)$$

is the relative gas density value. Its normalized value ρ_{sc} is defined by the coefficient c_ρ:

$$\rho_{sc} = c_\rho \rho_{rel}. \qquad (8)$$

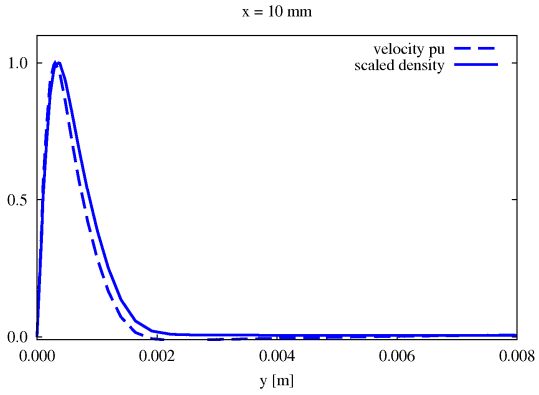


Figure 14: Numerical comparison between normalized flow velocity and relative scaled gas density (10mm).

Figure 14 shows the comparison between ρ_{sc} and the flow velocity normalized to its maximum value at a distance of 10 mm from the electrode along the x-direction obtained by the calculations. The numerical results prove that the buoyancy effects are negligible with respect to convection.

Figures 15 and 16 show the same comparison (when normalizing to maximum velocity value at 10 mm) at 20 mm and 40 mm from the electrode edge. The same value of c_ρ has been adopted in Figures 14, 15 and 16 and this also is in good agreement with measurements (compare Figures 12, 13 and 14). This demonstrates that the relationship between the two quantities (mass density and flow velocity profiles) does not depend on the location.

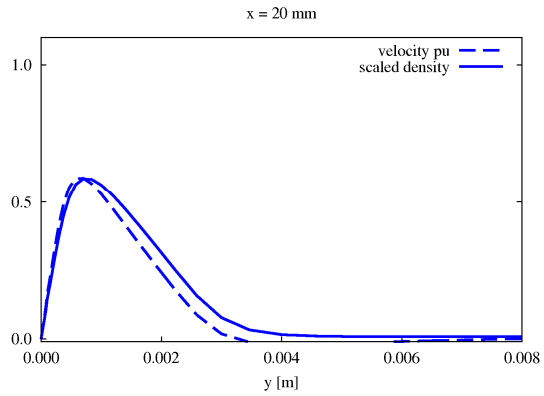


Figure 15: Numerical comparison between normalized flow velocity and relative scaled gas density (20mm).

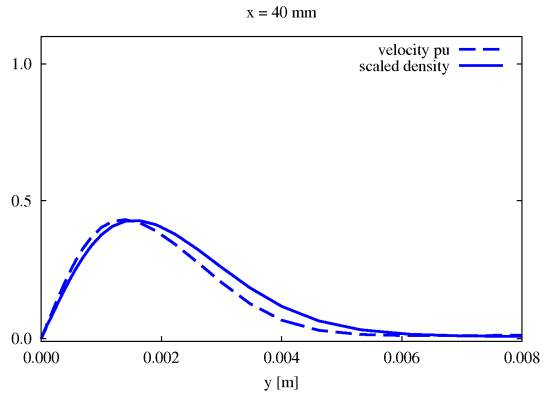


Figure 16: Numerical comparison between normalized flow velocity and relative scaled gas density (40mm).

6 Conclusion

The electro-fluid dynamic (EHD) interaction effects of a surface dielectric barrier discharge (DBD) plasma actuator, supplied with different a.c. voltage and frequencies, have been evaluated using fluid dynamic and optical

diagnostics.

Pitot velocity profiles have been performed on different actuators and different supply conditions. For low supply voltages the maximal speed is roughly the same at all the conditions. For increasing voltages the region in which the EHD interaction affects the boundary layer, rise. At high voltage a further growth of it is followed by an increase of the maximal value of the velocity. The velocity profiles and the electrical average power delivered to the electrodes have been used to determine the efficiency of the momentum transfer process involved in the EHD interaction phenomenon. The efficiency is always lower than 1 per cent. Higher values are reached for the teflon sheet actuator.

A Schlieren diagnostic, able to detect the morphology of the vortexes generated by the ignition phase of the discharge has been used. The successive images of the vortexes allowed to determine the vortex propagation velocity.

Comparisons between the Pitot velocity profiles and the integral of the pixel intensities of the Schlieren images of the hot wall jet at different positions away from the upper electrode have been done in steady state operations. For the three positions and the four supply conditions considered a good agreement has been found. Numerical simulations were performed using a modified fluid dynamic opensource code. The calculations confirm the relationship suggested by the experiments between the flow velocity and gas density.

A new use of the Schlieren technique is then defined. Quantitative information about the velocity flow field can be reached by a non-intrusive diagnostic that can be utilized in large scale facilities.

References

[1] Gad-El-Hak M 2000. "Flow Control". Cambridge University Press, 2000.

[2] Cristofolini, A., Neretti, G., and Borghi, C.A., "EHD Plasma Actuators Utilizing a Dielectric Barrier Discharge", *39th AIAA Plasmadynamics and Lasers Conference*, Seattle, WA USA, June 2008, Paper AIAA 2008-4286.

[3] Borghi, C.A., Carraro, M.R., Cristofolini, A., and Neretti, G., "Experimental Investigation on the EHD Interaction Induced by a Dielectric Barrier Discharge," Journal of Applied Physics, Vol. 103, pp. 063304-1 - 063304-11, ISSN: 0021-8979, 2008.

[4] Cristofolini, A., Borghi, C.A., Carraro, M.R., and Neretti, G., "A Study of The Electrical Supply System of a Barrier Discharge for EHD Flow Acceleration,", 38th AIAA Plasmadynamics And Lasers Conference, Miami, FL USA, June 2007, Paper AIAA 2007-4136.

[5] Borghi, C.A. Carraro, M.R., Cristofolini, A., and Neretti, G., "A Surface Barrier Discharge System for EHD Flow Acceleration," Proc. of the 7th Workshop on Magneto-Plasma-Aerodynamics in Aerospace Applications, Moscow, April 2007.

[6] J. Pons, E. Moreau, and G. Touchard. "Electrical and aerodynamic characteristics of atmospheric pressure barrier discharges in ambient air". Proc. ISNTPT, pages 370-410, 2004.

[7] T. Corke, D. Cavalieri, and E. Matlis. "Boundary layer instability on a sharp cone at mach 3.5 with controlled input". AIAA-2002-0350, 2001.

[8] D. Opaits, G. Neretti, S. Zaidi, M. Shneider, R. Miles, A. Likhanskii, and S. Macheret, "DBD Plasma Actuator Driven by a Combination of Low Freqeuncy Bias Voltage and Nanosecond Pulses", 46th AIAA Aerospace Sciences Meeting and Exhibit, Reno Nevada, Jan 7-10, 2008, AIAA-2008-1372.

[9] M.P. Patel, Z.H. Solwe, T.C. Corke, and C. He. Autonomous sensing and control of wing stall using a smart plasma slat. 44th AIAA Aerospace Sciences Meeting and Exhibit, AIAA 2006-1207, 2006.

[10] M.L. Post and T.C. Corke. Flow control with single dielectric barrier plasma actuators. 35th AIAA Fluid Dynamics Conference and Exhibit, AIAA 2005-4630, 2005.

[11] Borghi, C.A., Cristofolini, A., Neretti, G., Gurioli A. and Pertile R., "Experimental results in DBD plasma actuators for airflow control", 41st Plasmadynamics and Lasers Conference, Chicago, Illinois, June 28-1, 2010. AIAA-2010-4763.

[12] B. Arad, Y. Gazit, and A. Ludmirsky. "A sliding discharge device for producing cylindrical shock waves". J. of Appl. Physics, 20: 360-367, 1987.

[13] S.B. Leonov and D.A. Yarantsev. "Mechanisms of flow control by nearsurface electrical discharge generation". AIAA-2005-0780, 2005.

[14] J. Huang. "Documentation and control of flow separation on a linear cascade of Pak-B blades using plasma actuators". PhD thesis, University of Notre Dame, 2005.

[15] D.K. Van Ness II, T.C. Corke, and S.C. Morris. "Turbine tip clearance flow control using plasmas actuators". AIAA meeting, AIAA-2006-0021, 2006.

[16] 1 C.S. Kalra, M.N. Shneider, and R.B. Miles. "Numerical study of shockwave induced boundary layer separation control using plasma actuators". 46th AIAA Aerospace Science Meeting and Exhibit, AIAA-2008- 1095, 2008.

[17] J.P. Boeuf and L.C. Pithford. "Electrohydrodynamic force and aerodynamic flow acceleration in surface dielectric barrier discharge". Journal of Applied Physics, 97-103307, 2005.

[18] T.C. Corke, B. Mertz, and M.P. Patel. Plasma flow control optimized airfoil. 44th AIAA Aerospace Sciences Meeting and Exhibit, AIAA 2006-1208, 2006.

[19] D.V Gaitonde, M.R. Visbal, and S. Roy. A coupled approach for plasma-based ow control simulations of wing sections. 44th AIAA Aerospace Sciences

Meeting and Exhibit, AIAA 2006-1205, 2006.

[20] A. Likhanskii, m. Shneider, D. Opaits, R.B. Miles and S. Macheret, "Numerical modeling of DBD plasma actuators and the induced air flow", 38th AIAA Plasmadynamic and Laser Conference, Miami, Florida, June 25-28 2007.AIAA-2007-4533.

[21] G.S. Settles, "Schlieren and Shadograph Techniques", Springer 2001.

[22] R. Sosa, E. Arnaud, G. Artana, "Schlieren Image Velocimetry applied to EHD flows", International Simposium on Electrohydrodonamics (ISEHD), Buenos Aires, Argentina, 2006.

[23] D. Opaits, A. Likhanskii, G. Neretti, S. Zaidi, M. Shneider, R. Miles and S. Macheret, "Experimental Investigation on Dielectric Barrier Discharge Plasma Actuators Driven by Repetitive High-Voltage Nanosecond Pulses with DC or Low Frequency Sinusoidal bias", Journal of Applied Physics 104, 043304 (2008).

[24] Borghi, C.A., Cristofolini, A., and Neretti, G., "Experimental investigation on the energy transfer in a DBD plasma actuator for airflow control", 10th International Magneto-Plasma Aerodynamics, 22-24 March, Moscow.

[25] Goldstein, R.J., Kuehn, T.H., "Fluid Mechanic Measurements", Taylor and Francis, 1996.

[26] Moreau E., Labergue A., and G. Touchard, "About the kinetic power induced by AC and DC discharges", Proc. IEEE-CEIDP (Nashville, USA, October 2005.

CEAS 2011

Space Systems Design and Analysis 2

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Rosetta mission: on-comet operation planning for the sampler, drill and distribution subsystem

P. DiLizia, P. Francesconi, F. Malnati, F. Bernelli-Zazzera, and A. Ercoli-Finzi
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Milano, Italy

Keywords: Rosetta mission; Cometary soil; Drilling; Sampling; Operation planning

Abstract

This work describes the operation planning activities related to the sampler, drill and distribution (SD2) subsystem onboard the lander Philae of the Rosetta mission. Rosetta will reach the comet 67P/Churymov-Gerasimenko in 2014 and will release the lander Philae on the comet surface for in-situ investigation. SD2 is devoted to soil drilling, samples collection, and their distribution to three scientific instruments. The journey of Rosetta will last about 10 years. During its cruise to the comet, the status of SD2 has been checked during thirteen payload checkouts. After landing in 2014, Philae will accomplish a first scientific sequence lasting about 5 days, with a limited possibility of uploading telecommands from ground. Thus, SD2 operations must be accurately planned to be performed automatically. A dedicated software is being developed at Politecnico di Milano to support planning.

1 Introduction

Comets are believed to be the primitive leftover of the solar system formation process. They carry records of the solar system in a very early phase and are a key to our understanding of its origin and development. Moreover, they contain information on the compositional mixture from which the planets formed about 4.6 billion years ago and are also known to harbour complex organic molecules, to the great interest of those who study the origins of life on Earth [9].

Unfortunately, the real nature of comets remains unknown and, over the last three decades, several space missions have pursued the task of approaching a comet for close investigation. The first probe to visit a comet was the ISEE-3/ICE probe, which reached the comet Giacobini-Zinner in 1985, proving the "dirt-snowball" theory with its observations. Then, five spacecraft from the USSR (Vega 1 and 2), Japan (Sagigake and Suisei), and Europe (Giotto) were launched to make a rendezvous with Halley's comet in 1986. A more ambitious challenge was undertaken by space scientists in 1999 with the comet sample return mission known as Stardust. Its spacecraft had a close encounter with comet Wild 2, collecting samples of comet dust in a return capsule. More recently, NASA's Deep Impact probe was launched to reach comet Temple 1 in 2005 and send a projectile to it, creating an impact crater to carry out spectroscopy of the internal structures.

A more spectacular approach to comet exploration is near to completion, thanks to the European Space Agency (ESA)'s probe Rosetta. Rosetta will enhance our understanding of the formation and evolution of the solar system as well as the origin of life by investigating the 4 km-wide comet 67P/ Churyumov-Gerasimenko. Unlike past missions, Rosetta will be the first probe to carry out a close study of the nucleus,

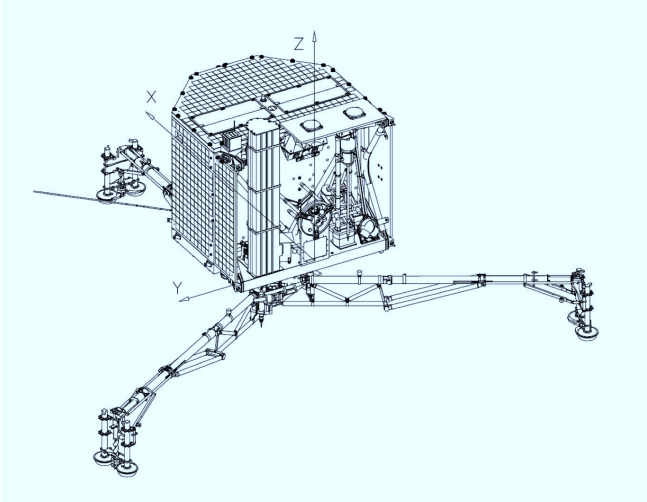


Figure 1: The lander Philae onboard Rosetta mission.

orbiting a spacecraft around the target comet. Even more challenging is its further goal of attempting the first ever landing on a comet surface, enabling in-situ analysis of cometary soil [9].

Rosetta was launched from the European spaceport of Kourou, French Guyana, in 2004, starting its ten-year journey to comet 67P/Churyumov- Gerasimenko. Along its complex trajectory to the target comet, Rosetta has already performed four gravity assist maneuvers about Earth and Mars, and two flybys of asteroid Šteins (September 2008) and Lutetia (July 2010). The spacecraft is currently in a special hibernation mode, which will be followed by the comet approach phase, in which Rosetta will enter its orbit around the comet and stay with it as it journeys in towards the Sun.

Right after arrival at the target comet in May 2014, Rosetta will start investigating the comet nucleus and the gas and dust ejected from it, by combining remote and direct sensing techniques. In November 2014, the lander Philae will be commanded to self-eject from Rosetta for a soft landing on the comet nucleus (see Fig. 1).

Philae will then act as an independent spacecraft, as it is provided with all subsystems needed to survive and work alone on the comet. The orbiter will be used only as a communication relay with the Earth. The ten scientific instruments on- board Philae will start performing their surface measurements after touch down. Instruments include camera systems (ÇIVA, ROLIS); evolved gas analyzers (Ptolemy, COSAC); an alpha-x-ray fluorescence spectrometer (APX), a combined magnetometer and simple plasma monitor (ROMAP); instrumentation to investigate the physical properties of the surface material (SESAME, MUPUS); and a radiowave experiment (CONSERT).

A crucial subsystem to ÇIVA, Ptolemy, and COSAC activities is the sampling, drilling, and distribution device (SD2). ÇIVA, Ptolemy, and COSAC are aimed at analyzing small comet samples in special conditions [1]. SD2 is devoted to support their activity by drilling into the cometary soil, collecting samples, and distributing them to the scientific instruments. The development of SD2 was committed to Politecnico di Milano and Galileo Avionica by the Italian Space Agency. These institutions currently manage the instrument for all in-flight and future on-comet operations.

This paper describes the development and planning of SD2 operations. The journey of Rosetta to the comet will last about 10 years. During its cruise to the comet, the probe has already been activated for thirteen payload checkouts, during which the status of SD2 was checked and its mechanisms exercised for standalone and combined tests with the other scientific instruments. This involved developing and testing dedicated operation procedures. Moreover, after landing in November 2014, a first scientific sequence (FSS) lasting about 5 days will be accomplished by Philae. Due to the scarce duration and the communication delay, the FSS activities must be carried out automatically, with a limited possibility of uploading telecommands from ground. To this aim, an accurate operation planning of SD2 activities is needed, for which a dedicated software is being developed at Politecnico di Milano.

The paper is organized as follows. The description of SD2 and the facility available at Politecnico di Milano for the development of the mission plans is reported in Sect. 2. The activities carried out during the payload checkouts of

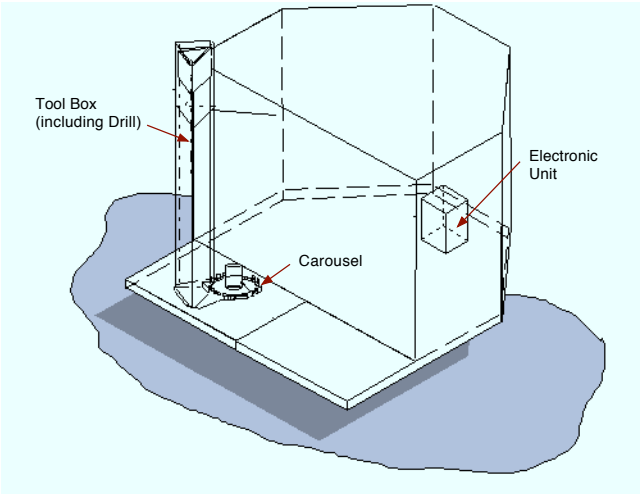


Figure 2: SD2 onboard the lander Philae.

the cruise phase and the issues related to operation planning for the on-comet phase are presented in Sect. 3. The software tool to support operation planning is presented in Sect. 4. Section 5 concludes the paper.

2 SD2 Description

The sampler, drill and distribution subsystem [5, 6] is a device that provides in-situ operations to collect and distribute samples to the scientific instruments ÇIVA, Ptolemy, and COSAC onboard the lander Philae of Rosetta mission. SD2 is mounted on the lander base plate (see Fig. 2), and it is equipped with:

- a drill able to collect several samples from 10 to 40 mm^3 at different depths. The maximum depth is 230 mm and samples retrieval is assured by the extraction of a sampling tube from the drill bit;
- a rotating platform (carousel) carrying 26 ovens. The carousel is suitably rotated to release the sample into an assigned oven and to place the filled oven under the proper scientific port;
- an electronic unit, which incorporates all electronics to control the drill and the carousel.

SD2 was designed to work in very low gravity and wide thermal excursion environment, and to optimize cutting performances with a very low power consumption (maximum mean power request of about 20 W during drilling at maximum torque).

Table 1: Main environmental parameters for SD2 design.

Parameter	# Range
Comet strength	50 Pa ÷ 50 MPa
Temperature	-140 °C ÷ +50 °C
Pressure	10^{-5} mbar ÷ 1 bar

The main environmental parameters for the design of SD2 system were comet soil strength, temperature and surface pressure. Wide parameters ranges were taken into account to assure SD2 functionality even in a highly uncertain environment (see Table 1).

SD2 has a total mass of 5100 g. Its mechanical unit [8] consists of a tool box, the drill, and the carousel. The tool box, built in carbon fibre, avoids drill damages due to vibrations and shocks during launch and landing phases. The drill, with a maximum extent capability of 581.6 mm from the lander balcony and a radius of 12 mm, is made of aluminium alloy. Sampling is achieved by a dedicated tube, which is commanded to extract from the drill bit. The collection of the samples is then performed by a pressure contact, which guarantees simplicity and flexibility of the solution adopted. Once the material is picked up, the carousel rotates to discharge the sample in one of the ovens. Then, a further carousel rotation takes the oven under the scientific port of ÇIVA, Ptolemy, or COSAC for subsequent analyses. Two kinds of oven are available onboard: 10 Medium Temperature Ovens (MTOs), which are able to heat up the sample for medium temperature experiments (+180 °C), and 16 High Temperature Ovens (HTOs) suited for high temperature experiments (+800 °C).

The electronic unit is installed into the warm compartment of the lander and incorporates all electronics to control the mechanical unit. Its

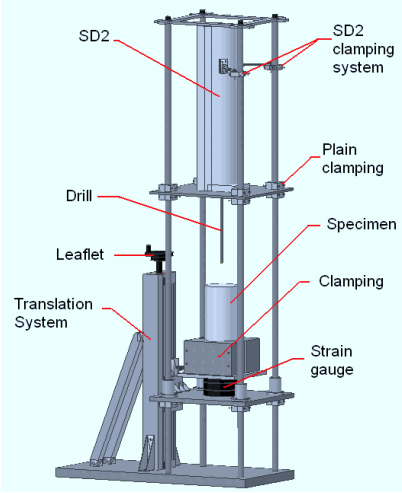


Figure 3: SD2 facility at Politecnico di Milano.

hardware and the software provide the interface between the mechanical unit and the Command Data and Management System (CDMS), which controls the lander [2]. SD2 is supplied with a 28 V line from the lander primary bus and some auxiliary power lines (±5 V, ±12 V) from the lander secondary converters.

2.1 SD2 Facility at Politecnico di Milano

The journey of Rosetta to the comet will last about 10 years. During this transfer time, the status of SD2 must be regularly checked during the passive and active payload checkouts. This involves developing dedicated procedures, which must be tested on ground before being uploaded and executed onboard. In addition, SD2 behavior needs to be tested in the many different scenarios that can be encountered on the comet.

A dedicated facility has been designed and realized in the laboratories of Politecnico di Milano to address the previous tasks. The facility is equipped with (see Figure 3):

- the SD2 flight spare (FS) model;
- a support structure that replicates the clamping system on the SD2 flight model (FM). The structure is realized with four tubular beams (2000 mm high), three aluminum base-plates, and four additional beams clamped at the top of the structure to reduce vibrations during the drilling phase;
- a sensor system, to measure the drill mechanical behavior: a biaxial strain gauge that measures the normal force and the torque during perforation, and a current sensor to measure SD2 FS power consumption during all drilling phases;
- a translation system and different soil specimen, which are meant to simulate the actual on-comet scenario. Three kinds of specimens are currently studied: gasbeton specimens, which are produced with different inclination of the perforated surface to simulate different shapes of the cometary soil; graphitic foam specimens, which are filled with water and then frozen to assess SD2 performances on icy soils; multilayer specimens, where a first granular gravel layer is followed by a gasbeton layer to study SD2 behavior in soils with varying granularity levels;
- an acquisition system to measure and elaborate the drill kinematics behavior in real time.

The SD2 facility available at Politecnico di Milano is instrumental to the management of SD2 FM during all mission phases, as it allows the designer to perform the following operations:

- mechanical verification of the perforation system;
- simulation of many different conditions of cometary soil;
- mechanical verification of the sampling and distribution mechanism;
- verification of SD2 behavior during mission plan development and execution;
- implementation of contingency operations;
- assessment of SD2 power consumption during different drilling and sampling procedures.

3 SD2 Operations

The Philae operations centers during Rosetta's journey to the comet are the Lander Control Centre (LCC) at DLR, Cologne, and the Science Operations and Navigation Centre (SONC) at CNES, Toulouse. These centers include all the necessary equipment to support all data processing and distribution tasks, and are responsible for all lander operations, including:

- lander operations planning and verification;
- data monitoring and control of the subsystems and instruments;
- distribution and archiving of all received lander data.

An additional crucial responsibility of LCC and SONC is the coordination of the lander instruments, including SD2, during all operations. Nominal lander operations include regular health checks of the system, subsystems and instruments during cruise, through the so-called "payload checkouts", as well as the science operations during the on-comet phase. Politecnico di Milano is responsible for the development, verification, and delivery of the mission plans necessary to accomplish SD2 tasks during Philae operations.

3.1 Payload Checkouts

A total of 13 payload checkouts (PCs) have been performed by Rosetta during its cruise period between commissioning and deep space hibernation. During each PC, the instruments have the opportunity to activate their units, exercise mechanisms, refresh EEPROM memories, and perform calibrations and software updates. Interactive operations are only possible during the active PCs, whereas passive PCs allow for standard procedures for general monitoring of the instruments and lander status.

A list of all SD2 activities during all payload checkouts is reported in Table 2. Only SD2 translation and rotation resolvers are activated during passive payload checkouts. This allows the SD2 team to check the instrument status and to measure the drill and carousel positions that are compared with expected values. Active payload checkouts were devoted to exercise the drill and the carousel. More specifically:

- downward (DW) and upward (UW) drill translations were executed to verify the drill translation motor functioning;
- clockwise (CW) and counterclockwise (CCW) drill rotations were executed to check the drill rotation motor status;
- carousel movements were performed for stand-alone and combined tests with ÇIVA, Ptolemy, and COSAC;
- the EEPROM memory was refreshed before hibernation.

All the above activities involve developing dedicated operation procedures. For each operation, a specific list of commands is generated to optimize data return. The resulting mission plan is checked on ground through the SD2 facility available at Politecnico di Milano and sent to SONC and LCC. The requests by all instruments are then merged to form the lander operation plan to be tested on the lander Ground Reference Model available at LCC before being executed onboard.

Following onboard execution, the generated telemetry is distributed to all instruments for immediate analysis. A software tool for a fast and reliable analysis of SD2 telemetry has been developed at Politecnico di Milano [7]. The software allows the designer to:

- verify SD2 status;
- check the drill and carousel positions during the execution of all operations;
- confirm the achievement of the operation objectives and the fulfillment of the constraints;
- analyze errors in case of non-nominal behaviors.

The downloaded telemetry shows that SD2 behaved nominally during all PCs. The amount of collected data was consistent with the performed activities, and the telemetry for both

Table 2: SD2 activities during payload checkouts.

PC#	Type	Activities
PC0÷3, PC5, PC9	passive	Standard checkout
PC7, PC11	-	Cancelled
PC4	active	7 carousel rotations for combined tests with COSAC and ÇIVA
PC6	active	• 7 carousel rotations for combined tests with COSAC and ÇIVA • 1 CW and 1 CCW drill rotation (21 s)
PC8	active	• 6 carousel rotations for combined tests with ÇIVA • 1 CW and 1 CCW drill rotation (21 s) • 1 DW (to 0.3 mm) and 1 UW (to 0 mm) drill translation • EEPROM refresh
PC10	active	• 8 carousel rotations for combined tests with COSAC and ÇIVA • 1 CW and 1 CCW drill rotation (90 s) • 1 DW (to 2.5 mm) and 1 UW (to 0 mm) drill translation
PC12	active	• 2 carousel rotations for combined tests with COSAC • EEPROM refresh
PC13	active	Standard checkout

the carousel and the drill movements matched expectations within the admissible tolerances. Based on these results, SD2 is considered fully commissioned and ready for the next mission phases.

3.2 First Science Sequence

Rosetta is currently in a Deep Space Hibernation (DSH) mode. No flight activities are performed during this phase, which will end in 2014, when the lander Philae (like all other Rosetta components) will be re-commissioned and data from the orbiter instruments will be evaluated for the planning of the detailed landing scenario.

After landing in November 2014, a first scientific sequence (FSS) lasting about 5 days is planned. The lander will be powered to a large extent by its primary battery (capacity: 1 kWh) and several instruments and subsystems will be operated simultaneously. SD2 will act as a lander subsystem, sampling the cometary soil and distributing the samples to the proper scientific port. This is a non-trivial task, which involves

- collecting the sampling requests by ÇIVA, Ptolemy, and COSAC (i.e., number of samples, depths, ovens and scientific ports to be used);
- developing an optimal operation plan to fulfill the requests, while minimizing SD2 power consumption and taking into account operations priorities.

Besides the scarce duration of the FSS, significant communication delays might appear during on-comet operations due to the relative visibility between the lander and the orbiter, and the visibility of the orbiter from ground. This limits the possibility of uploading telecommands during FSS activities, which must be executed automatically, minimizing ground support. Thus, a strong interaction between the SD2 team and the other scientific instruments onboard Philae is ongoing, under the coordination of SONC and LCC, to perform an accurate operation plan-

ning. The aim is to define the exact sequence of mission plans to be loaded onboard prior to landing and executed during the entire FSS for all sampling procedures. Energy consumption is a key parameter for planning, as the mission scientific outcome must be maximized while satisfying the constraint of the energy stored in the primary battery. Moreover, the operation sequence must be robust to the uncertainties of the comet environment, in order to minimize the risk of failures and the need of uploading the associated recovery procedures.

3.3 Long Term Science

The FSS will be followed by the long term science (LTS). During LTS, the lander will rely on the power from the solar generator and a secondary battery. Instruments operations will mainly be performed in sequence and the data evaluation will be carried out primarily offline. The lander instruments operations are planned to last a few months on the comet surface.

During this phase, SD2 will continue to support ÇIVA, Ptolemy, and COSAC. However, recent studies have proven the existence of a correlation between the drill behavior during perforation and the mechanical characteristics of the cometary soil. Thus SD2 was recently proposed to be used also as a scientific instrument [6]. A dedicated strategy is being developed, which is based on performing multiple perforations with varying speed levels, in order to identify the drill working zones in the space of the drill rotation and translation speeds. Information about the mechanical characteristics and inhomogeneity of the cometary soil will be obtained by comparing the working zone measured on the comet during the LTS phase and the working zones of a database of specimens available on ground.

Being based on multiple perforation, this strategy is time and power consuming, and requires specific mission plans to be developed for each drilling procedure. Several drilling strategies need to be analyzed before landing in order to define the operation plan. More specifically, the overall duration of the sampling operations and SD2 power consumption need to be assessed to compare the strategies. Moreover, for each perforation, the associated list of telecommands must be developed and tested onground before being uploaded onboard.

Table 3: Steps of a typical sampling procedure.

Step	Description
1	Rearming oven under drill
2	DW translation for drill bit rearming
3	UW translation to home position
4	Carousel rotation to home position
5	Drilling to desired depth
6	Sampling tube release
7	Drill rotation for 1 min
8	UW translation to home position
9	Desired oven under drill
10	DW translation for sample delivery
11	UW translation to home position
12	Desired oven under scientific port

4 SD2 Activity Planner

A software tool to support SD2 operation planning in the mission operation scenario presented in Sect. 3 is under development at Politecnico di Milano. The software is based on the subdivision of a typical sampling procedures in elementary mission plans (EMPs). Given ÇIVA, Ptolemy, and COSAC requests, SD2 operation engineer defines the associated sequence of EMPs. Then, the software tool estimates the duration of the overall operation and SD2 energy consumption, and generates the associated mission plans.

4.1 Elementary Mission Plans

Table 3 reports the steps of a typical SD2 sampling procedure [4]. It is worth observing that not all the steps differ for each sampling: some of them can be grouped to build elementary blocks that can be repeated identically in any sampling procedure. The resulting common elementary blocks are listed hereafter:

- *Drill bit rearming*: Steps 1, 2, 3, and 4. Description: the drill bit is rearmed and SD2 is prepared for drilling. It is worth observing that this will be the first block to be executed during FSS, as SD2 is currently flying with the drill bit in unarmed position.

- *Sample retrieval*: Steps 6, 7, and 8. Description: once the desired depth has been reached by the drill during Step 5, this block extracts the sample from the cometary soil and moves the drill back to its home position.

- *Oven feeding*: Steps 10 and 11. Description: once the oven to be used is under the drill (see Step 9), this block delivers the sample into the oven.

For all the above elementary blocks of activities, the same sequence of commands is executed during any sampling procedure, and the same settings can be used for SD2 at any occurrence. Consequently, it is convenient to develop one mission plan for each common elementary block of activities, which will be loaded on Philae and used at each occurrence during all sampling procedures. This reduces onboard memory utilization and maximizes the number of commands that can be stored in the dedicated buffer, so limiting the need of uploading telecommands from ground.

Specific mission plans must be prepared and uploaded for Steps 5, 9, and 12:

- Step 5 refers to the roto-translation necessary to drill into the cometary soil until the desired depth. The commands to be executed depend on the depth to be reached, which is different for each sampling procedure;

- Step 9 rotates the carousel to take the oven under the drill for sample delivery. The commands depend on the specific oven to be used, which is different for each sampling procedure;

- Step 12 rotates the carousel to take the oven with its sample under the proper scientific port. Similarly to Step 9, the associated commands depend on the specific oven and the scientific port to be used, which might differ for each sampling procedure.

As the commands associated to Steps 5, 9, and 12 are specific to each sampling procedure, specific mission plans must be designed for each sampling procedure.

Based on the above analysis, a total number of 5 EMPs can be identified. Drill bit rearming, sample retrieval, and oven feeding are common EMPs, which are repeated identically during any sampling procedure. Specific EMPs must be developed for the drill roto-translations and the carousel rotations in Steps 5, 9, and 12.

4.2 Software Description

Besides optimizing onboard memory consumption, the definition of the EMPs eases the task of defining SD2 operations. More specifically, any sampling procedure can be seen as a sequence of EMPs. Thus, the definition of any procedure is now reduced to the identification of the sequence of EMPs that satisfies ÇIVA, Ptolemy, and COSAC requests. Once the sequence is available, the overall duration of the activity and SD2 power consumption can be estimated, and the associated mission plans can be generated for onboard execution. The software tool "SD2 Activity Planner" (SD2-AP) is being developed at Politecnico di Milano to serve this purpose.

The software guides the operation engineer through the definition and entry of the EMP sequence, thanks to a proper Graphical User Interface (GUI), which is reported in Fig. 4. The left side of the GUI is devoted to input definition. The user can select the EMP to be added to the sequence through a dedicated drop-down menu (see detail in Fig. 5). Depending on the EMP selected, the necessary parameters are supplied in dedicated sections. More specifically, no parameters are necessary for the common EMPs, which do not depend on the specific procedure. The translation and rotation speed levels, as well as the desired depth are asked for the drill roto-translation movements (Step 5 in Table 3). The oven to be used and the scientific port to be

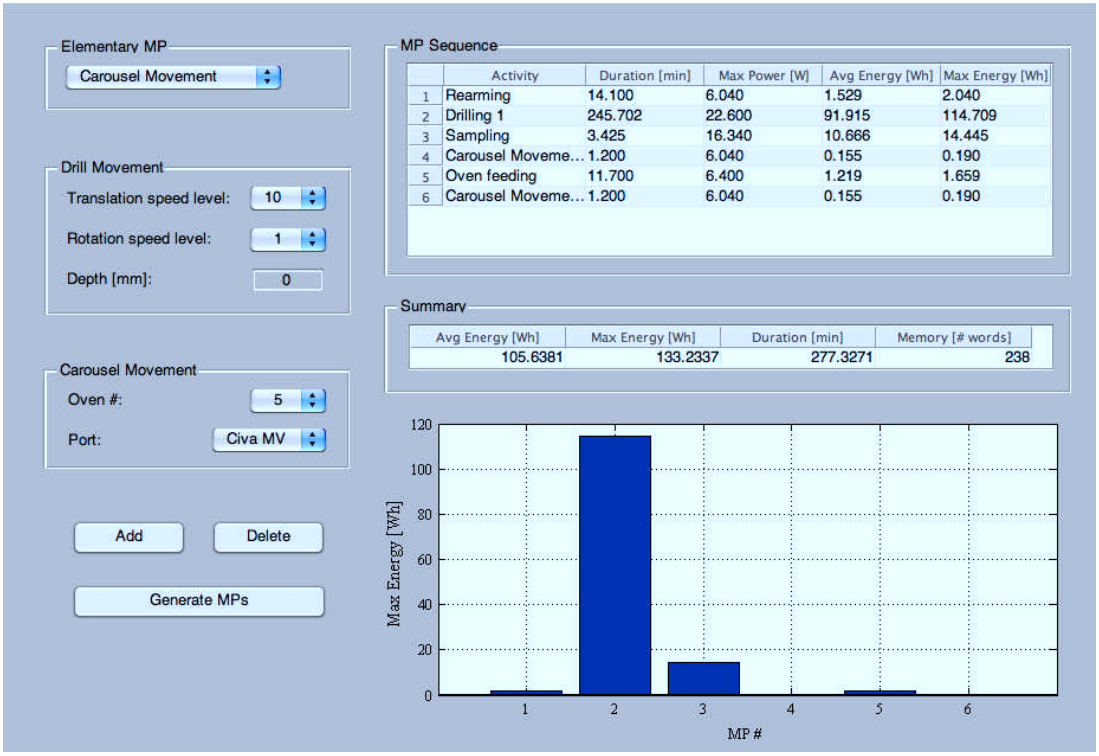


Figure 4: SD2 Activity Planner: main GUI.

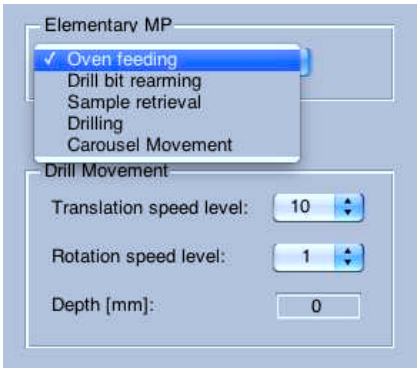


Figure 5: SD2 Activity Planner: EMP selection.

served are the inputs to the specific carousel movements (Steps 9 and 12 in Table 3).

For each EMP, the duration of the associated activity, and the average and maximum power and energy consumption are estimated based on laboratory and in-flight tests [3]. The estimates are reported in a dedicated table at the top-right corner of the GUI. The overall results for the entire sequence are resumed in a summary table, illustrating also onboard memory consumption in terms of number of words to be stored in the telecommand buffer. For the sake of an immediate comparison among the activities, the energy consumption of each EMP is also plotted on a bar chart.

Once the sequence is completely defined, the button "Generate MPs" can be pushed to generate the mission plans associated to all EMPs in the sequence. One mission plan for the common EMPs is produced for all their occurrences. Different mission plans are instead generated for each occurrence of a specific EMP. Figure 6 illustrates the typical list of commands for the common EMP "Sample retrieval". The first row reports the position on the telecommand buffer where the mission plan will be stored (131 in figure). The SD2 commands are reported in the subsequent rows. The generated mission plans are ready to be tested onground and uploaded on Philae for execution.

5 Conclusion

The operation planning for the SD2 subsystem onboard the lander Philae of the Rosetta mission has been described in this work. The cur-

```
[MISSION_PLAN] 131
STARTOP 1
MHIT 1 87
SARE
DRGO 18 7 0 60
DELAY 320
DRTC 19 7 0 0 21750
ONOF 0 1 1 0 0 0 0 0
DELAY 90
ONOF 0 0 0 0 0 0 0 0
MHIT 0 0
STOPOP 1 1
```

Figure 6: Mission plan for the elementary block "Sample retrieval".

rent status of the instrument, as well as the open work for the on-comet phase has been presented. SD2 operation planning will be supported by the use of the SD2-AP software, which is being developed at Politecnico di Milano. SD2-AP is an easy and reliable tool to compare operation procedures in terms of overall duration and energy consumption, and will support SD2 operation engineers towards the definition the optimal operation strategy. Future developments of the tool will address data budget estimation and the inclusion of checks for constraint satisfaction between consecutive elementary mission plans.

Acknowledgements

The authors would like to acknowledge the support of the Italian Space Agency to this work, which was funded through the contract I/062/08/0.

References

[1] Biele, J., and Ulamec, S., "Capabilities of Philae, the Rosetta Lander", *Space Science Reviews*, Vol. 138, 2008, pp. 275-289.

[2] Crudo, E., and Bologna, P., "CDMS-SD2: Data Interface Control Document, TECNOSPAZIO S.p.A., 2002.

[3] DiLizia, P., Malnati, F., and Ercoli-Finzi, A., "SD2 In-Flight Power Consumption", Politecnico di Milano, 2009.

[4] DiLizia, P., Malnati, F., and Ercoli-Finzi, A., "SD2 Elementary Mission Plans", Politecnico di Milano, 2010.

[5] Ercoli-Finzi, A., Magnani, P.G., Re, E., Espinasse, S., and Olivieri, A., "SD2 - How to Sample a Comet, *Space Science Reviews*, Vol. 128, 2007, pp. 281-299.

[6] Francesconi, P., Malnati, F., and Ercoli-Finzi, A., "SD2, a Technological/Scientific Instrument to Understand a Comet", *Journal of Aerospace Science, Technology and Systems*, Vol. 88, 2010, pp. 129-138.

[7] Malnati, F., and Ercoli-Finzi, A., "SD2 Telemetry Analyzer User Manual", Politecnico di Milano, 2004.

[8] Malychev, M., and Magnani, P.G., "Rosetta Lander SD2: Subsystem Specification", TECNOSPAZIO S.p.A., 2001.

[9] Ulamec, S., Balazs, A., Biele, J., Espinasse, S., Fantinati, C., et al., "Rosetta Lander - System Status after Five Years in Space", *Journal of Aerospace Science, Technology and Systems*, Vol. 88, 2010, pp. 121-128.

3rd CEAS Air&Space Conference
21st AIDAA Congress

Sampler, Drill and Distribution system for in-situ analysis of a cometary nucleus

Pietro Francesconi, Pierluigi DiLizia, Fabio Malnati, Franco Bernelli-Zazzera, and Amalia Ercoli-Finzi
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Italy
sd2@aero.polimi.it

Keywords: *small bodies exploration, cometary analysis*

Abstract

This paper describes the activities related to the Sampler, Drill and Distribution subsystem onboard the lander Philae of the Rosetta mission. The flight spare model of the subsystem has been integrated in a dedicated facility at Politecnico di Milano to develop the mission plans to be executed onboard and assess the subsystem performances on a range of realistic scenarios. The possibility of a scientific use of the subsystem is also presented. More specifically, a test campaign is ongoing to identify a correlation between telemetry data and mechanical characteristics of the cometary soil. The results on the gasbeton specimens and on the new icy specimens are illustrated.

1 Introduction

Rosetta is the third cornerstone mission of the European Space Agency scientific program "Horizon 2000". Launched in March 2004, it will reach the comet 67P/Churymov-Gerasimenko in 2014 and will be the first spacecraft to orbit around a comet nucleus. Its probe is made up by two systems (see Fig. 1):

- the orbiter, *Rosetta*: a pseudo cube of 2.8 x 2.1 x 2.0 m, on which all subsystems and payload equipment are mounted;
- a Lander, *Philae*: a pseudo hexagonal prism that will self-eject from Rosetta for a soft landing on the comet nucleus [1].

Rosetta is the first space mission to travel beyond the main asteroid belt relying solely on solar cells for power generation, rather than the traditional radio-isotope thermal generators. The new solar-cell technology used on the orbiter two solar panels (32 m tip-to-tip in deployed configuration) allows it to operate over 800 million kilometres from the Sun, where solar radiation is only 4% of that on Earth. Hundreds of thousands of specially developed nonreflective silicon cells generate up to 8700 W in the inner Solar System and around 400 W for the deep-space comet encounter, what is in any case a very small amount to perform scientific experiments as foreseen.

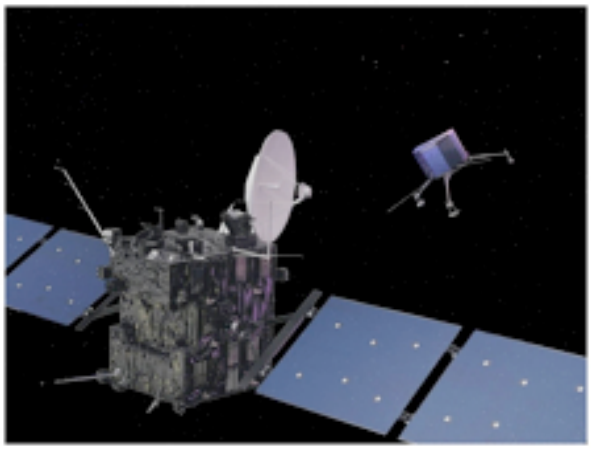


Fig. 1 The orbiter Rosetta and the lander Philae.

Rosetta will reach 67P/Churyumov-Gerasimenko at about 450 millions kilometers from the Earth, after three Earth gravity assists (GA), one Mars GA and two asteroid fly-bys. During its rendezvous with 67P/Churyumov-Gerasimenko, Rosetta will enter its orbit around the comet and fly with the comet toward the inner solar system. This will enable a close study of the nucleus, as well as the investigation of its evolution throughout its approach to the Sun, by remote sensing techniques. The measurements will provide data to understand the nucleus internal structure, the nucleus nature and the mineralogical, chemical and isotopical composition.

After its arrival on the comet, the lander Philae (see Fig. 2) will perform the first soft landing on a comet nucleus to study in-situ the comet nucleus composition. The lander is provided with all subsystems needed to survive and work alone on the comet. Ten scientific instruments will perform their surface measurements after touch down. The mission goals include the determination of the elementary and mineralogical composition, the identification of traces elements, and isotopic composition of cometary material from the surface and subsurface.

Comet's surface strength, density, texture, porosity, ice phases and thermal properties will also be investigated together with soil structure. To this aim, individual grains will be analyzed by three scientific instruments:

- COSAC: one of the two evolved gas analyzers onboard Philae, designed to detect and identify complex organic molecules from their elemental and molecular composition;
- PTOLEMY: an evolved gas analyzer designed to perform accurate measurements of the isotopic ratios of light elements;
- ÇIVA: a visible light microscope coupled to an infrared spectrometer to provide data on the composition, texture and albedo of comet samples.

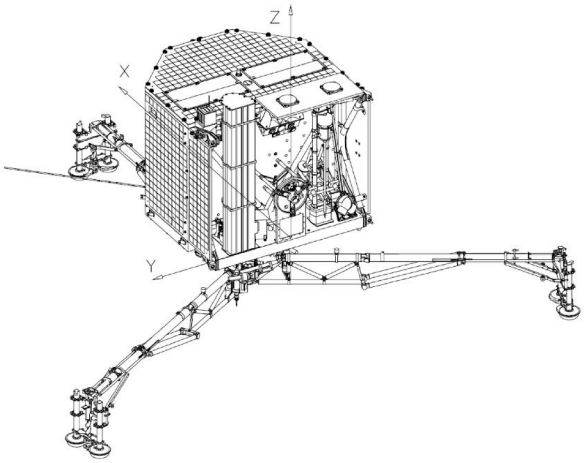


Fig. 2 The Rosetta lander Philae

A crucial subsystem to ÇIVA, PTOLEMY, and COSAC activities is the sampling, drilling, and distribution (SD2) device. SD2 is devoted to support their activity by drilling into the cometary soil, collecting samples, and distributing them to the scientific instruments. The design and management of SD2 was committed to Politecnico di Milano and Galileo Avionica by the Italian Space Agency.

2 SD2 Description

The Sampler, Drill and Distribution subsystem is the multifunction device that provides in-situ operations in order to collect and distribute samples to ÇIVA, PTOLEMY, and COSAC. SD2 is mounted on Philae's baseplate (see Fig. 3), and it is equipped with a drill able to collect several samples of 10-40 mm^3 at different depths (the maximum depth is 230 mm) from the same hole or different ones. It was designed to operate in critical thermo-vacuum environment and to meet the demanding mass/power resources limits of Rosetta mission. To this aim, very innovative technological solutions are adopted [4]: sampling tube for the drill/sampling operations, composite materials, dry lubrication, stepper motors, medium temperature ovens design.

Table 1 Environmental parameters for SD2 design.

Parameters	Range
Comet strength	50 Pa ÷ 50 MPa
Temperature	-140° C ÷ +50° C
Pressure	10^{-5} mbar ÷ 1 bar

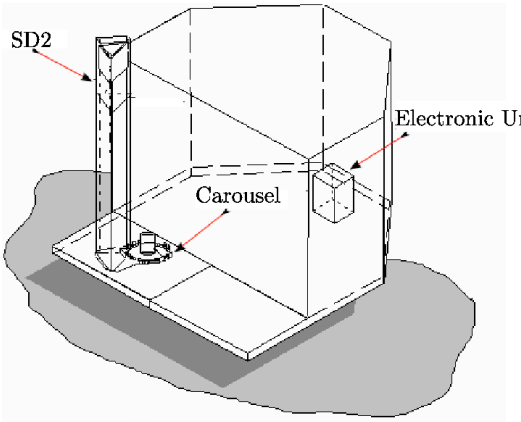


Fig. 3 SD2 components distribution on Philae.

The driving environmental parameters for the subsystem design were comet soil strength, temperature and surface pressure (see Table 1). Parameters ranges, taken into account during the design phase, are wide in order to assure functioning in presence of high uncertainty.

SD2 has a total mass of 5100 g and is composed by a *mechanical* unit (3700 g), an *electronic* unit embedding SD2 software (1000 g), and the *harness* for electrical connection between the mechanical and electronic units (400 g).

The SD2 mechanical unit consists of the Tool Box, the Drill, and the Carousel:

- The Tool Box is built in carbon fibre and avoids drill damages due to vibrations and shocks during launch and landing phases.
- The Drill is made of aluminium alloy, and has a radius of 12 mm and a maximum extension of 581.6 mm from the lander balcony; polycrystalline diamonds have been used to reinforce the drill bit for hard soil drilling; position, shape and geometry of the bits have been optimized by theoretical analysis, numerical simulations and experimental tests, in order to maximize the cutting capability with a low vertical trust (100 N) and a low power consumption. The power consumption during operations has a maximum average value of about 20 W. A sampling tube is embedded into the Drill bit and extracted to pick up the sample from the cometary soil. This solution was chosen for its simplicity and flexibility: the collection of the samples is performed by a pressure contact, as well as the release.
- The Carousel is a rotating platform on which some small containers, the ovens, are mounted. The material is picked up and discharged in the ovens. The Carousel task is to position the oven and its sample under the scientific ports of ÇIVA, PTOLEMY, and COSAC. The ovens provide the interface between the collected sample and the scientific instrument. According to the scientists requests, two kinds of ovens are on-board: 10 medium temperature ovens with an optical sapphire prism, suited for the analysis by visible I/R microscopes, before heating up for medium temperature experiment (+180°), and 16 high temperature ovens suited for high temperature experiments (+800°).

The electronic unit is installed into the warm compartment of the lander and incorporates all electronics to control the mechanical unit. The hardware and software installed provide the interface between the mechanical unit and the lander control system, the Command Data and Management System (CDMS) [2]. SD2 is supplied by Philae's power subsystem with a 28 V line from the lander primary bus, devoted to the mechanical unit, and some auxiliary power lines (±5 V, ±12 V) from the lander secondary converters.

3 SD2 Facility at Politecnico di Milano

A dedicated facility was designed and realized at Politecnico di Milano to test SD2 behavior and assess its performances in different realistic scenarios [5]. The operations

that can be carried out at the SD2 facility are listed hereafter:

- mechanical verification of the drilling system; i.e., analysis of the drill structural behavior and assessment of the force and torque transmitted by the drill to the specimen;
- simulation of different comet-like soils and drilling scenarios;
- mechanical and functional verification of the sampling and collecting system;
- assessment of SD2 behavior during mission plan execution;
- definition of the optimal perforation strategies;
- design of the recovery procedures in case of SD2 non-nominal behavior;
- assessment of SD2 power consumption during drilling/sampling activities.

The design of the SD2 facility was accomplished to meet the requirements issuing from the previous tasks. The resulting facility is equipped with (see Fig. 4):

- the SD2 Flight Spare (FS) model;
- a support structure, designed for easy inspection and replicating the clamping system on the Flight Model (FM);
- an acquisition system used to acquire and process data during tests.

The support structure is made up by (see Fig. 5):

- four tubular beams (2 m high and with a radius of 30 mm);



Fig. 4 SD2 facility at Politecnico di Milano

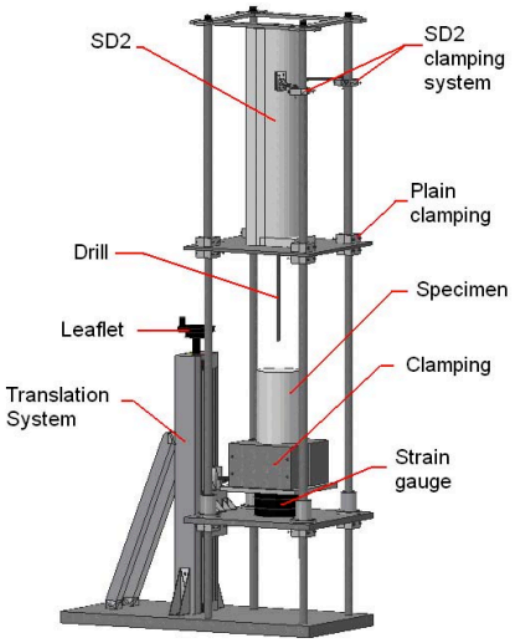


Fig. 5 Support structure for the SD2 FS model.

- three aluminum plates used as baseplate, SD2 FS and sensor system support;
- four beams clamped at the top of the structure to reduce vibrations or distortions during drilling phase.

The plates are clamped to the tubular beams to avoid SD2 FS movements. A translation system allows the sensor system and the specimen to translate to a known position. This decreases the problem of misalignment between the drill and the drilled surface.

The sensor system is composed by a biaxial strain gauge that measures the normal force and the torque applied to the specimen. The strain gauge has:

- torque full scale of 25 Nm with a resolution of 0.01 Nm;
- vertical force full scale of 1000 N with a resolution of 1 N.

The strain gauge must measure load contributions given by the drill during perforation (maximum load of 100 N), the specimen, and the clamping system (220 ÷ 500 N). The interface between the strain gauge and the structure is composed by eight screws, each of them tolerating a load range of 20 ÷ 65 N. A current sensor is added to measure SD2 FS power consumption during operation.

4 SD2 Scientific Contribution

Recent investigations suggest the possibility of correlating the drill behavior during perforation with the mechanical characteristics of the cometary soil. This outlines the opportunity of using SD2 not only as a tool to support ÇIVA, PTOLEMY, and COSAC, but also as a scientific instrument itself. To this purpose, a correlation between SD2 telemetry data and the cometary soil characteristics must be identified [5].

Unfortunately, due to the nature of the drill rotation and translation stepper motors, the variation of power consumption can not be used as an indicator of a variation of the soil characteristics. In fact, a stepper motor works with a fixed current value for each imposed speed value (see [3] for more details). Consequently, an alternative strategy has been developed, which is based on the identification of the drill *working zones*.

More specifically, the designed strategy relies on the observation that, during the perforation phase, small discontinuities in the specimen (e.g. grains of different dimension and strength) or a layer of high strength cause one of the following behaviors:

- no current variation: the resistance torque is lower than its maximum allowed value and the drill goes on drilling with the same speed;
- motor stops: the resistance torque is greater than its maximum allowed value and the drill cannot continue drilling with the given speed and torque levels. Thus the motor stops and a system failure occurs.

The failures the system can show can be gathered in two main categories:

- *force failure*: during the perforation phase the force suddenly increases (max value 100 N). This failure is characterized by a steep slope in the vertical force diagram and by large oscillations in the torque value;
- *speed failure*: during the perforation phase, the measured rotation speed is lower than expected. The system switches off power supply to the motor and a failure message is reported. Moreover, a fast and unexpected increase of the vertical force value is observed.

The described behaviors can be used to identify the maximum and minimum drill speed levels for each torque level on different comet-like materials and specimens' shapes. Given a fixed torque level, the strategy consists in commanding a sequence of perforations at different drill rotation and translation speed levels to identify their minimum and maximum values that allow a complete perforation to be performed without drilling failures. The resulting ranges define an area in the drill translation speed - drill rotation speed plane, which is the drill working zone for the imposed torque and drilling scenario (e.g., material and specimen's shape, see Fig. 6).

The strategy is applied for different torque levels, comet-like materials and specimen's shapes, to build a database of working zones. A mathematical model will then be developed to correlate the mechanical characteristics of the specimens to the identified working zones. The same procedure will be repeated on the comet by the SD2 flight model to obtain the working zone specific to the cometary soil. The mathematical model will then be used to determine the drilled material nature or possibly the category the material belongs to. Thus, SD2 can be exploited as a scientific instrument capable of identifying the material nature, a possible comet stratification and, in particular, significant unhomogeneity in the upper layers of the comet soil.

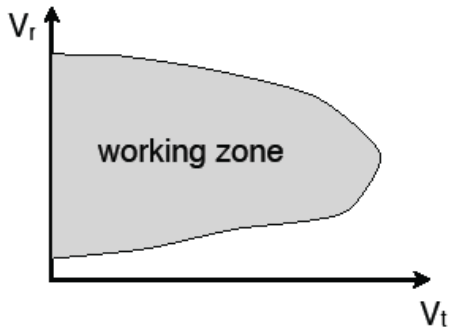


Fig. 6 Schematic representation of a working zone.

Table 2 Gasbeton mechanical characteristics.

Property	Range	Mean
Density [kg/m^3]	300÷900	500
E [GPa]	1.5÷1.9	1.75
σ_C [MPa]	1÷9	4
σ_T [MPa]	0.1÷2.25	1
τ_S [MPa]	0.25÷2.7	2

4.1 Preliminary results

This section describes the preliminary results of the test campaign aimed at building the database of working zones. The main material used for the test campaign is Gasbeton, which has mechanical characteristics comparable with the expected properties of the target comet 67P/Churyumov-Gerasimenko. Table 2 lists the mechanical characteristics of the Gasbeton material.

Specimens with different inclination of the perforated surface are used to simulate the irregular shape of the comet soil under the drill and the associated drilling scenario. In addition, SD2 behavior on multi-layer specimens is investigated to simulate the possible variation of soil granularity. The multi-layer specimens vary in terms of Gasbeton vertical thickness, grains dimension in the granular layers and layer alternance. The set of specimens available at the SD2 laboratory for the test campaign are illustrated in Fig. 7. A picture of the clamping system for the Gasbeton specimens is reported in Fig. 8.

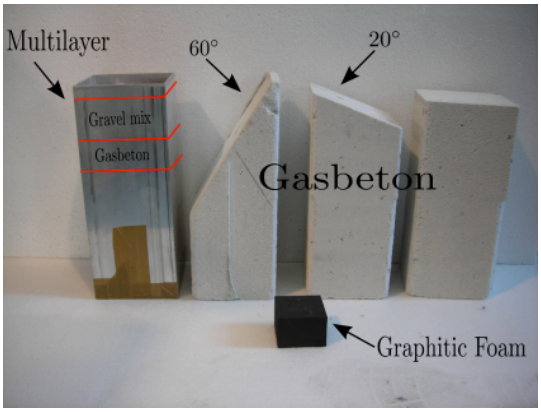


Fig. 7 Specimens available at SD2 laboratory.



Fig. 8 Gasbeton specimen clamped for a drilling test.

The methodology described in the previous section was used to build the working zones for the Gasbeton specimens with different surface inclinations. Some drilling test diagrams are illustrated in the followings to better understand how the minimum rotation speed is identified. Then, some working zones for the Gasbeton specimens are shown, which will contribute to the database for the correlation model between telemetry data and mechanical characteristics of the cometary soil.

First of all, the results of a successful 2 cm perforation are shown in Fig. 9. More specifically, the vertical force, torque, and current profiles during drilling are reported. The different drilling phases can be readily recognized from the vertical force diagram: variations of its value occurs at the beginning of each phase. In addition, the drill rotation and translation motors switching-on can be clearly identified on the current diagram, due to the sudden increase of current consumption.

The results of Fig. 9 can be compared with the diagrams in Fig. 10 and Fig. 11, which report the force, torque, and current profiles for a force failure and a rotation speed failure respectively. The vertical force and the torque diagrams differ significantly for the two failure cases. In the force failure case (Fig. 10), the torque oscillation increases with the vertical force. The drilling test is aborted when the force value reaches 80 N. The rotation speed failure (Fig. 11) is characterized by a high slope of the vertical force diagram. The system shuts down the stepper motors and returns an error message.

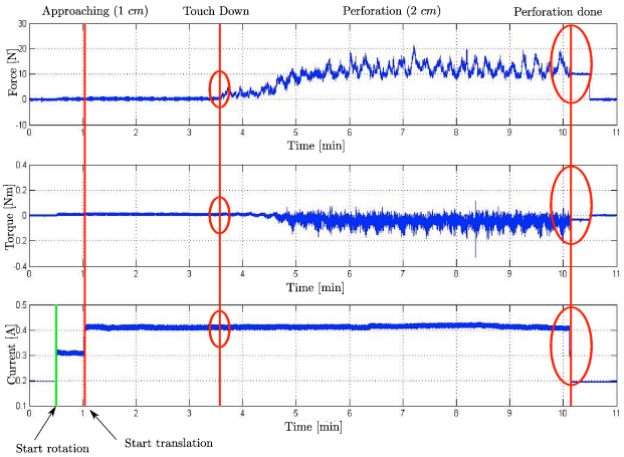


Fig. 9 Successful drilling test on a Gasbeton specimen with flat surface.

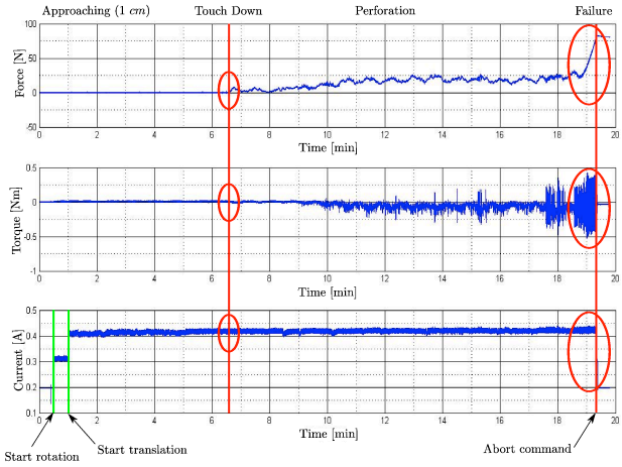


Fig. 10 Force failure test example.

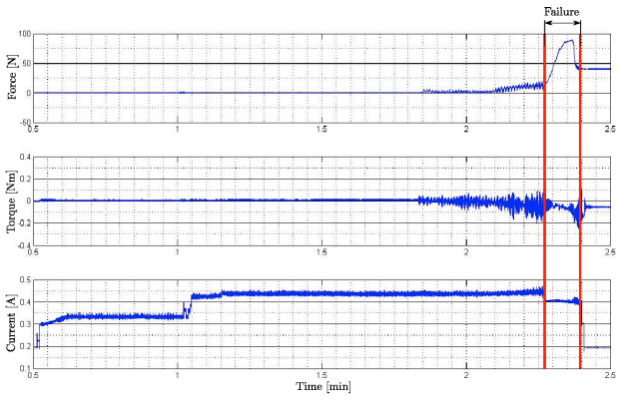


Fig. 11 Rotation speed failure test example.

The results on a Gasbeton specimen with 0° inclination, corresponding to commanded translation (C_t) and rotation (C_r) torque levels 4 and 2 respectively, is illustrated in Fig. 12. About 60 perforation tests were necessary to identify the working zone, which is represented by the area inside the red line in the translation speed (V_t) – rotation speed (V_r) plane in figure. The force failure and speed failure cases are plotted with red dots and yellow dots respectively.

The working zone on a Gasbeton specimen with 20° inclination, corresponding to the same C_t and C_r, is reported in Fig. 13 and compared with the results obtained for the 0° inclination specimen. Tests at different inclinations are necessary to study the possible effects of the soil shape on the working zones.

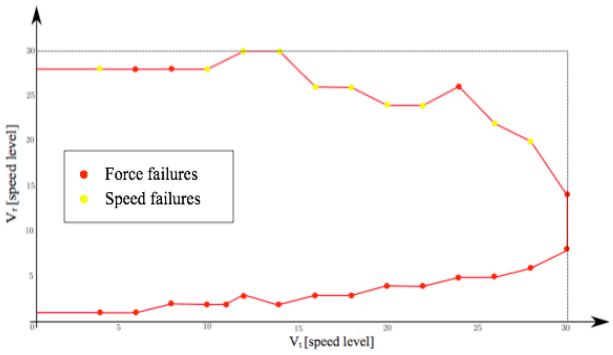


Fig. 12 Working zone for a Gasbeton specimen with 0° inclination ($C_t = 4$, $C_r = 2$).

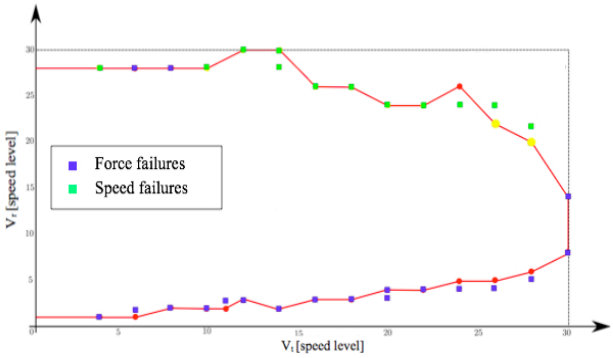


Fig. 13 Working zone for a Gasbeton specimen with 20° inclination ($C_t = 4$, $C_r = 2$).

5 Low temperature operations

The first results of the new test campaign on icy specimens are presented in this section. The purpose of test campaign is twofold:

- verify the effectiveness of the sampling mechanism on icy materials;
- identify the working zone of icy specimens.

A new specimen slot and clamping system have been developed to allow drilling/sampling test operations in low temperature. The low temperature system for the tests on icy specimens is made up of:

- a nitrogen tank;
- a cylindrical specimen slot, with a gap filled with nitrogen during tests (see Fig. 14);
- five insulating wood plugs for sensors system insulation.

A picture of the resulting low temperature system during tests is reported in Fig. 15.

A test campaign for the thermal characterization of the low temperature system has been carried out before the perforation tests. More specifically, the temperature distribution in the cylindrical specimen slot has been measured during tests in thermovacuum chamber (see Fig. 16). The aims of the thermal characterization were:

- verifying that the system keeps the temperature of the specimen below 0° C;

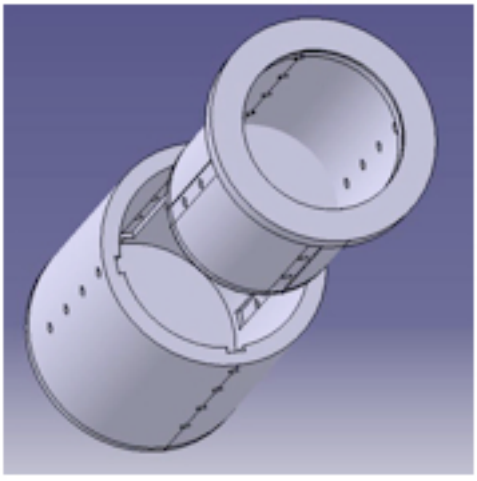


Fig. 14 Cad model of the slot for tests on icy specimens.



Fig. 15 Low temperature system: nitrogen tank, iced specimen slot, and insulating wood plugs.

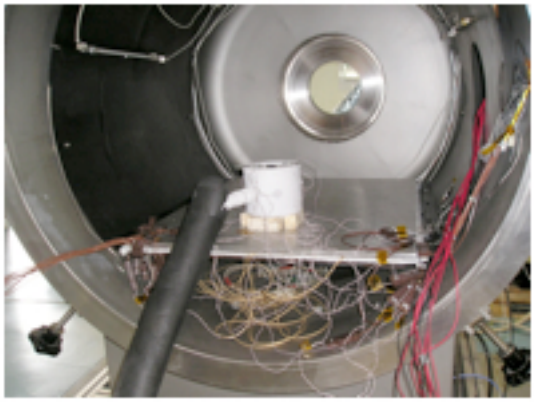


Fig. 16 Thermal characterization of the low temperature system in thermovacuum chamber.

- avoiding high temperature gradients in the specimen slot;
- assessing the performances of the insulating wood plugs at keeping the temperature of the sensor system within the admissible range.

First perforation tests have been performed on three different specimens:

- pure ice;
- ice and soil particles;
- graphitic foam with different water concentrations.

For the first specimen, a block of pure ice has been used. Unfortunately, the test campaign on pure ice specimens showed the inadequacy of the low temperature system to assess SD2 performances on icy specimens. More specifically, SD2 FS model was not able to drill pure ice. However, this is not attributable to an SD2 failure. Rather it is due to the significant

temperature difference between the specimen and SD2 FS model. During perforation, the water melts in contact with the drill cutting. Thus, instead of exiting the hole, the water tends to stay in the hole and to freeze again. Consequently, ice accumulates in the hole and blocks the drill (see Fig. 17).

Additional tests were performed on a new specimen: granular pieces of soil have been embedded into the ice (see Fig. 18). The drill was able to perforate few millimeters before experimenting a failure connected to the same phenomenon.

The same behavior was observed with the last icy specimen: a block of graphitic foam has been soaked with water at different concentration levels and then frozen (see Fig. 19). Similarly to the previous cases, the drill cannot perforate the entire block and SD2 fails perforation after few millimeters.

The above results prove that a more adequate low-temperature system is necessary for the test campaign on the icy specimens. More specifically, a isolated environment is necessary to guarantee an homogeneous temperature during tests and to reduce the temperature difference between the specimen and the drill cuttings.

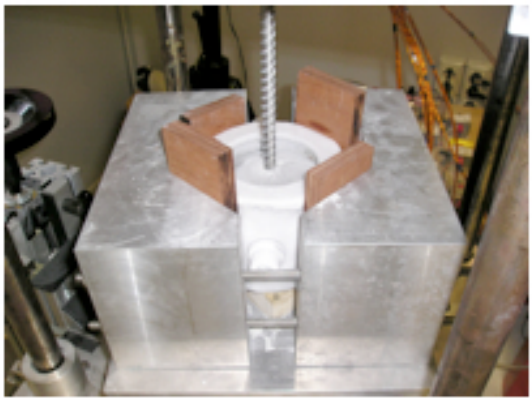


Fig. 17 Drill stuck into the pure ice specimen.

6 Conclusions

The scientific use of the SD2 subsystem onboard the lander Philae of Rosetta mission has been described. It is based on the identification of a correlation between drill behavior during perforation and mechanical characteristics of the cometary soil. The strategy

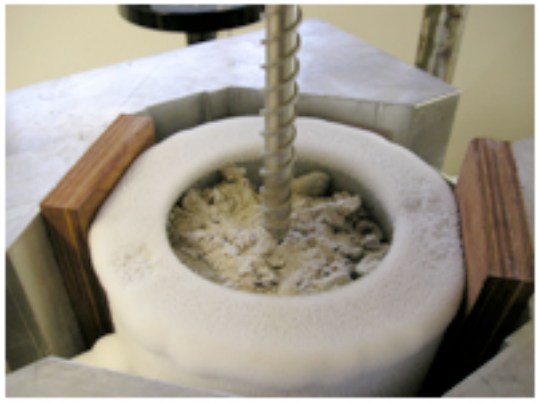


Fig. 18 Granular pieces of soil embedded into ice.



Fig. 19 Graphitic foam soaked with water.

relies on the identification of the drill working zones for different specimens and perforation scenarios. It is worth observing that the strategy expands SD2 capacities, but it is time and power consuming, due to the high number of perforations necessary to identify the drill working zone. Consequently, the scientific use of SD2 is considered as a back-up solution for the on-comet phase.

In addition, the low temperature system for the test campaign on icy specimens has been described. However, the first tests showed some criticalities of the system that need to be addressed. The design of a dedicated test-bed is ongoing to guarantee an homogeneous temperature during tests.

Acknowledgements

The authors would like to acknowledge the support of the Italian Space Agency to this

work, which was funded through the contract I/062/08/0.

References

[1] Berner, C., Bourillet, L., Ellwood, J., Kasper, M., Kletzine, P., Schultz, R., and Schwehm, G., "Rosetta: ESA's Comet Chaser", *ESA Bulletin*, Vol. 112, 2002, pp. 10-37.

[2] Crudo, E., and Bologna, P., "CDMS-SD2: Data Interface Control Document, TECNOSPAZIO S.p.A., 2002.

[3] Dainese, C., Ercoli-Finzi, A., and Malnati, F., "Test Facility for SD2 Comet Sampler Performance Improvement", 9th ESA Workshop on Advanced Space Technologies for Robotics and Automation (ASTRA 2006), ESTEC, Noordwijk (The Netherlands), 2006.

[4] Ercoli-Finzi, A., Magnani, P.G., Re, E., Espinasse, S., and Olivieri, A., "SD2 – How to Sample a Comet", *Space Science Reviews*, Vol. 128, 2007, pp. 281-299.

[5] Francesconi, P., Malnati, F., and Ercoli-Finzi, A., "SD2, a Technological/Scientific Instrument to Understand a Comet", *Journal of Aerospace Science, Technology and Systems*, Vol. 88, 2010, pp. 129-138.

Finite Element Analysis of a Solid Booster Flexible Joint for Thrust Vector Control

Luca Lampani, Francesca Angelini and Paolo Gaudenzi
Dipartimento Ingegneria Meccanica e Aerospaziale
Università di Roma "La Sapienza", Via Eudossiana 18, 00184 Roma, Italy

Renato Marocco and Mauro Fabrizi
Nozzle Design Group, Engineering Dept.
AVIO S.p.A., Via Ariana km 5.2, 00034 Colleferro Roma, Italy

Keywords: *flexible joint, thrust vector control, finite element*

Abstract

The aim of the work is to present the analysis of the behaviour in operative conditions of a component of the thrust vector control of a solid rocket motor. This component, a laminated flexible joint, is a fundamental part of the thrust vector control system of the motor. It is devoted to interface the nozzle with the booster case introducing the flexibility necessary to allow the nozzle to follow the movements imposed by actuators in order to obtain a deflection of the jet flow. The scope of this component, constituted as a sequence of reinforcement and rubber layers, is to work as much as possible as a spherical joint with centre (pivot point) along the axis of the motor. Most of the difficulties encountered during the study are related to the convergence of the solution, due to the strongly nonlinear behaviour of rubber under pressure and actuation. A precise characterization of the constitutive laws of the rubber material under high compression stress state become fundamental when it is modelled as a quasi-incompressible media.

1 Introduction

The thrust vector control system [1] is a device present in the guidance system of most jets, spacecraft, launchers and rockets to obtain a correct trajectory during flight [2]. This kind of steering system also operates to compensate the deviances due to wind shear, geometrical imperfection or asymmetries due to manufacturing process or misalignment of the masses. Two ways are possible to design such a system: with aerodynamic techniques, i.e. by introduction of fins and mobile surfaces, or with mechanical deviation of the thrust, i.e. with a modification of the thrust axis. The second type only can operate in absence of atmosphere, then more suitable for spacecraft and launchers. The present work is dedicated to this second class of devices. In particular the thrust can be oriented with an injection of liquid into the nozzle flowfield to unbalance the thrust or the asymmetry of the hot gas or pivoting the nozzle around the centre of a spherical joint with a kinematic or a flexible system. The design process of the flexible joint for solid boosters, through a numerical

approach with the finite element technique, is the focus of our work.

2 Configuration of a flexible joint

The flexible joint is a very delicate component of the chemical thrusters, complex to design and analyse because of its heavy operating conditions and the inherent nature of the object itself. It is a non-rigid connection between the motor case and the movable nozzle (Fig. 1), which then allows the deflection of the nozzle itself in any direction by a system of forces applied by two actuators placed at a distance of 90 degrees each other.

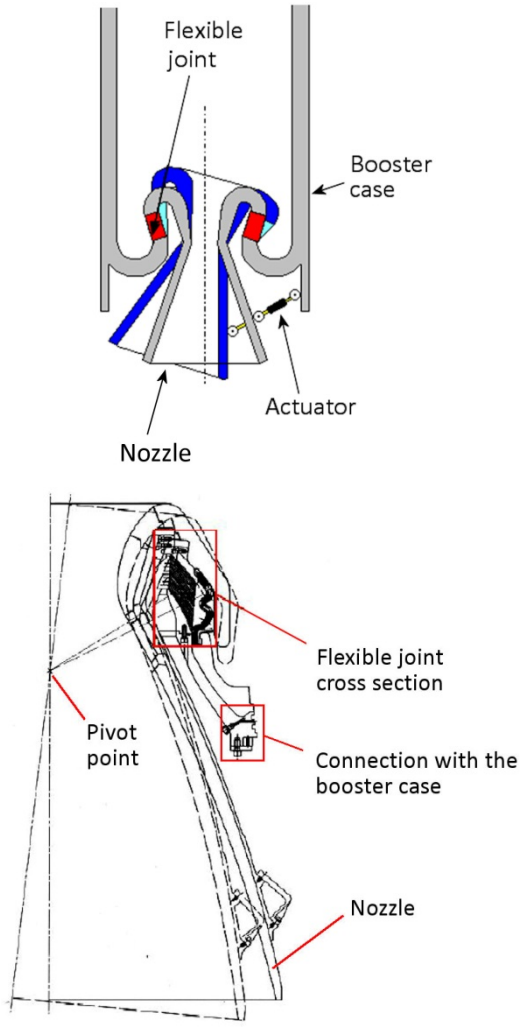


Fig. 1 Thrust vector control with flexible joint.

The deflection of the thrust allows to get a moment with respect to the center of gravity of the launcher, and this mechanical action can be used to develop a control subsystem for the trajectory of the launcher itself. The configuration of the flexible joint aims to reproduce a spherical kinematism by means of a flexible structure. In order to deflect the axis of the nozzle of some degree it is necessary to guarantee an high deformability without loose the spherical nature of its movement.

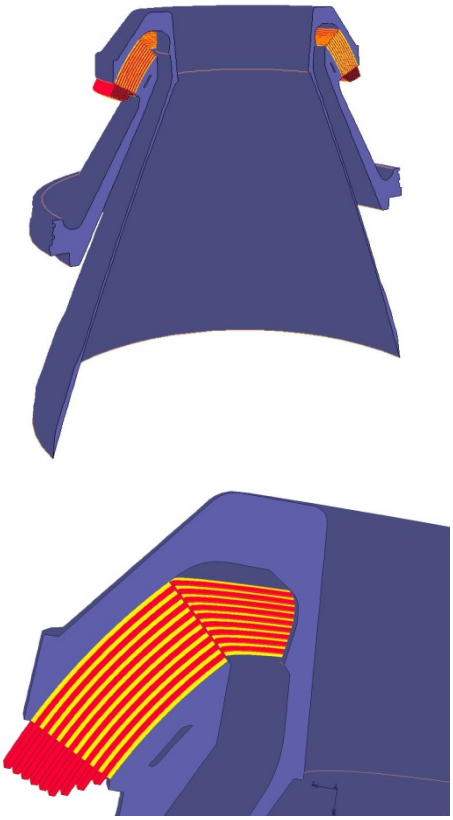


Fig. 2 Sketch of a nozzle with related flexible joint mechanism and close-up.

For this reason it is composed in a laminated way and consists of a series of layers of rubber and steel or composite with spherical section (Fig. 2), connected to the case and nozzle with aluminium supports. The joint can be closely exposed to the environment of the combustion chamber of the motor, it is then subjected essentially to two kind of actions, mechanical and thermal, at the same time. From a mechanical point of view the operative pressure of the motor is combined with the forces explicated by the hydraulic actuators. The heat fluxes coming

from the nozzle and the combustion chamber are mitigated by thermal protection and barriers but it needs to be taken into account in order to define the thermal requirements of the materials and the thermomechanical stresses. Flexible joints can be equipped with: metallic reinforcements, in order to withstand large stresses, but need a thermal protection to work in room temperature environment or composite reinforcements, in order to decrease weight and work without thermal protections, but withstanding lower stresses. The configuration with metallic shims needs of a flexible thermal barrier against the high heat fluxes. The configuration with composite shims can resist to higher temperature: in this case the shims protrudes outside the rubber layers and autoprotect the joint (close-up of Fig. 2). Such configuration is evaluated in our analyses.

The problems associated with the two load conditions are different and require different types of analysis. Under the high pressure load produced in the combustion chamber, the nozzle compresses the joint inducing translation of the geometrical pivot point, namely "off-set", from its nominal position. The joint deforms and consequently it modifies the position of the nozzle along the axis relatively to the booster case. The presence of the off-set influences the behaviour of the joint and produces a variation on the length of the actuators that affects the efficiency of the control system. Purpose of this work is to estimate such parameter too. Despite the control system is provided of a feedback loop, the prediction of the structural response of the joint to the applied pressure load allows to optimize the control system itself.

3 The finite element approach

Requirement of the flexible joint is to transfer the pressure and actuation loads with: a) the minimum displacement of the geometrical pivot point possible, b) the use of the minimum actuation force possible to deflect the thrust and c) without exceed the limits of resistance of the materials. It is then evident that the main design driver is to balance the stiffness required in the first point with the flexibility required in the second one. The stacking sequence with rubber layers and rigid shims allows to achieve this objective. The stiffness along the axis direction is guaranteed by the incompressibility of the rubber and the bending flexibility by big shear deformations. These parameters are usually evaluated through empirical relations, anyway extremely rare in literature [3] and with the counteract to be dependent by the size and configuration of the joint. In this way the extension of some results to develop different types of joint is difficult and the experimental test with full scale breadboards is expensive. All these considerations lead to develop a design methodology versatile and reliable, with the use of the finite element technique. Therefore the numerical problem to handle is complex. The structure of the joint is subjected to large displacements, non-linear behaviour of the rubber and quite total incompressibility. Most of the difficulties encountered during the study are related to the convergence of the solution, due to the strongly nonlinear behaviour of the rubber. A precise characterization of the constitutive behaviour of the rubber material under high compression stress state become fundamental when it is modelled as a quasi-incompressible media [4]. In order to achieve a reliable solution of the problem, the following items assume a relevant importance in the numerical procedure:

- performing an extensive finite element incremental analyses taking into account the large displacement of the model;
- characterizing the rubber material on the basis of energetic models derived from the nonlinear elasticity theory;
- to adopt a finite element mixed formulation with extended variational principle, where the primary variables are not only displacements but also ins and/or stresses, to take into account the quasi incompressibility of the rubber.

The non-axial symmetry of the loading conditions, summarized with a mechanical pressure and forces of actuation to rotate the nozzle around his pivot point, made necessary to elaborate a complete three-dimensional numerical model. Moreover the non-linearity of the analysis imposes the superposition of the loads and not of their effects.

4 The characterization of the rubber material

From an engineering point of view, rubbers are homogeneous, isotropic materials that exhibit very large elastic strains (in a tensile test, a specimen could reach ten times its original length). Moreover, one can observe from a tension test that the stress-strain behavior is strongly nonlinear. Rubber and rubber-like materials are usually modeled as incompressible hyperelastic materials. The nonlinear theory of elasticity is applied for the description of their mechanical behavior because structural components in the practice undergo large strain conditions. The finite-element method is suitable for the analysis of rubber and rubber-like materials provided that a nonlinear constitutive law and an iterative approach are used. In fact it is necessary to properly model the nonlinear stress-strain relations and take into account incompressible behavior of these materials, and the presence of large deformation.

Consider a body at a generic time t, a fundamental measure of the deformation of the body is given by the *deformation gradient*, defined as:

$$ {}^{t}_{0}\boldsymbol{X} = ({}_{0}\nabla\, {}^{t}\boldsymbol{x}^{T})^{T} \qquad (1) $$

where:

${}_{0}\nabla$: gradient operator

${}^{t}\boldsymbol{x}$: coordinates at time t

In this calculation it is used the *right Cauchy-Green deformation* tensor:

$$ {}^{t}_{0}\boldsymbol{C} = {}^{t}_{0}\boldsymbol{X}^{T}\, {}^{t}_{0}\boldsymbol{X} = 2\,{}^{t}_{0}\boldsymbol{\epsilon} + \boldsymbol{I} \qquad (2) $$

where ${}^{t}_{0}\boldsymbol{\epsilon}$ is the *Green-Lagrange strain* tensor; this can be written in terms of Cauchy-Green deformation tensor as:

$$ {}^{t}_{0}\boldsymbol{\epsilon} = \frac{1}{2}({}^{t}_{0}\boldsymbol{C} - \boldsymbol{I}) \qquad (3) $$

The energy conjugate stress measure to use with the Green-Lagrange strain tensor is the *second Piola-Kirchhoff stress* tensor ${}^{t}_{0}\boldsymbol{S}$ [5]. Hyperelastic material models assume that materials response is isotropic and isothermal. This assumption allows that the strain energy potentials are expressed in terms of strain invariants or principal stretch ratios. Except as otherwise indicated, the materials are also assumed to be nearly or purely incompressible. Material thermal expansion is also assumed to be isotropic. The constitutive behavior of hyperelastic materials are usually derived from the strain energy potentials. The isotropic hyperelastic effects are mathematically described by specifying the dependence of the strain energy density (per unit original volume) W on the Green-Lagrange strain tensor. The strain energy density W is written in terms of the invariants or stretches. In many cases, the strain energy density is conveniently written as the sum of the deviatoric strain energy density W_D and the volumetric strain energy density W_V. Depending on the specific material model used, several forms of strain energy potential, such as Neo-Hookean, Mooney-Rivlin, Polynomial Form, Ogden Potential, Arruda-Boyce, Gent, and Yeoh are available [5]. In our study, a good agreement with the experimental results, for all the cases where the compressibility effects are negligible, is obtained with the Ogden model [6]. The conventional Ogden model is

described by the deviatoric strain energy density per unit original volume:

$$W_D = \sum_{n=1}^{3} \frac{\mu_n}{\alpha_n}\left(\lambda_1^{\alpha_n} + \lambda_2^{\alpha_n} + \lambda_3^{\alpha_n} - 3\right) \qquad (4)$$

where μ_n and α_n are the Ogden material constants and the λ_i are the square roots of the principal stretches of the Cauchy-Green deformation tensor. The volumetric part is:

$$W_V = \frac{1}{2}\kappa(\lambda_1\lambda_2\lambda_3 - 1)^2 \qquad (5)$$

where κ is the bulk modulus. With the knowledge of how the strain energy density W depends on the Green-Lagrange strain tensor (through the invariants or stretches), the 2nd Piola-Kirchhoff stress tensor is evaluated using:

$$S_{ij} = \frac{1}{2}\left(\frac{\partial W}{\partial \epsilon_{ij}} + \frac{\partial W}{\partial \epsilon_{ji}}\right) \qquad (6)$$

and the incremental material elastic tensor is evaluated using the following expression:

$$C_{ijrs} = \frac{1}{2}\left(\frac{\partial S_{ij}}{\partial \epsilon_{rs}} + \frac{\partial S_{ij}}{\partial \epsilon_{sr}}\right) \qquad (7)$$

In order to characterize a particular rubber for this kind of description it is necessary to obtain the constants μ_n and α_n for n = 1, 2, 3 and the bulk modulus by experimental data. The constants follow from the least square method applied to data coming from several different experimental tests [7][8]. For incompressible elastomers, the basic strain states are simple tension, pure shear and simple compression. For experimental reasons the compression can be replaced by equal biaxial extension. For slightly compressible situations or situations where an elastomer is highly constrained, a volumetric compression test may be needed to determine the bulk behaviour. For this simple deformation states, the expressions in Eq.(6) assumes a particularly simple formulation.

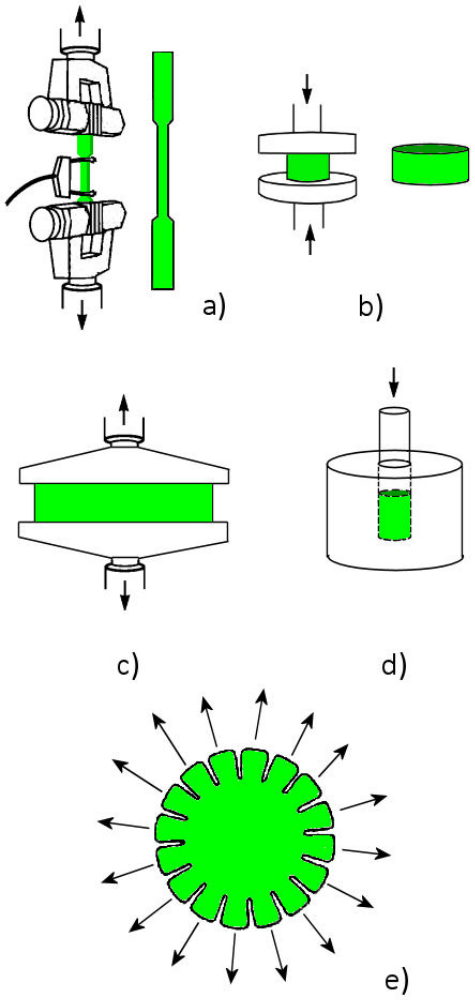


Fig. 3 Experimental tests for rubber characterization: a) simple tension, b) simple compression, c) pure shear, d) bulk modulus and b) equal biaxial extension.

The experimental tests for the characterization of the rubber are shown in Fig. 3. These tests are related to the standard ASTM D412 for the simple tension and ASTM D575 for the compression. A similar to D412 standard is adopted for shear, where the specimen must be at least 10 times wider than the length in the stretching direction. Because the material is nearly incompressible, a state of pure shear exists in the specimen at a 45 degree angle to the stretching direction. Volumetric compression is an experiment where the compressibility of the material is examined. In this experiment,

a cylindrical specimen is constrained in a fixture and compressed. The actual displacement during compression is very small and great care must be taken to measure only the specimen compliance and not the stiffness of the instrument itself. The initial slope of the resulting stress-strain function is the bulk modulus. This value is typically 2-3 orders of magnitude greater than the shear modulus for elastomers. All the results of the experimental tests are condensed in a data set to give the material constant for the material model used in the finite element formulation.

5 Finite element models

In order to study the thermal behaviour of the joint under prescribed temperatures and heat fluxes, a 2D finite element modal is developed.

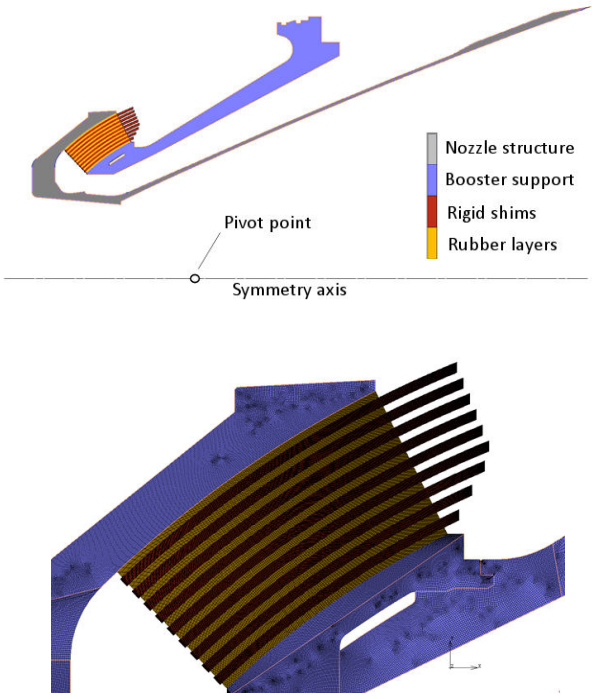
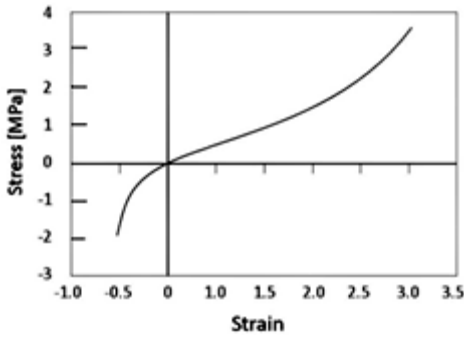


Fig. 4 2D axisymmetric finite element model of the joint with close-up.

The mesh of this model is also adopted for a preliminary study on the structure: the axial symmetry of the pressure load and constraints allows to obtain some preliminary results to be compared with the results of the more complex 3D model. On the contrary the out-of-axis actuation load can be applied on the 3D model only. The advantage of the 2D model is to have the possibility to use a particularly refined mesh.The 2D finite element model is composed up to 236140 axisymmetric elements quad-4, full & Hermann formulation, the 3D is composed of 197316 elements hex-8. The joint has an inner diameter of about 0.25m and is composed of rubber layers, glass-epoxy composite oriented 0/90°/±45° shims and aluminium parts (Fig. 4).

From a structural point of view the rubber is modelled using the Ogden's formulation with three parameters, the shims and the aluminium parts are modelled as linear elastic isotropic materials (Table 1).

Glass-Epoxy (isotropic)	
Young's Module (GPa)	12
Density (kg/m^3)	1850
Poisson's Ratio	0.31
Aluminium	
Young's Module (GPa)	72
Density (kg/m^3)	2700
Poisson's Ratio	0.33
Rubber	



κ (N/m^2)	2 10^9
Density (kg/m^3)	1000

Table 1. Structural material characteristics.

From a thermal point of view the rubber and the aluminium are modelled as isotropic materials, the shims orthotropic. The conductivity and the specific heat are

characteristics temperature-dependent (Fig. 5).

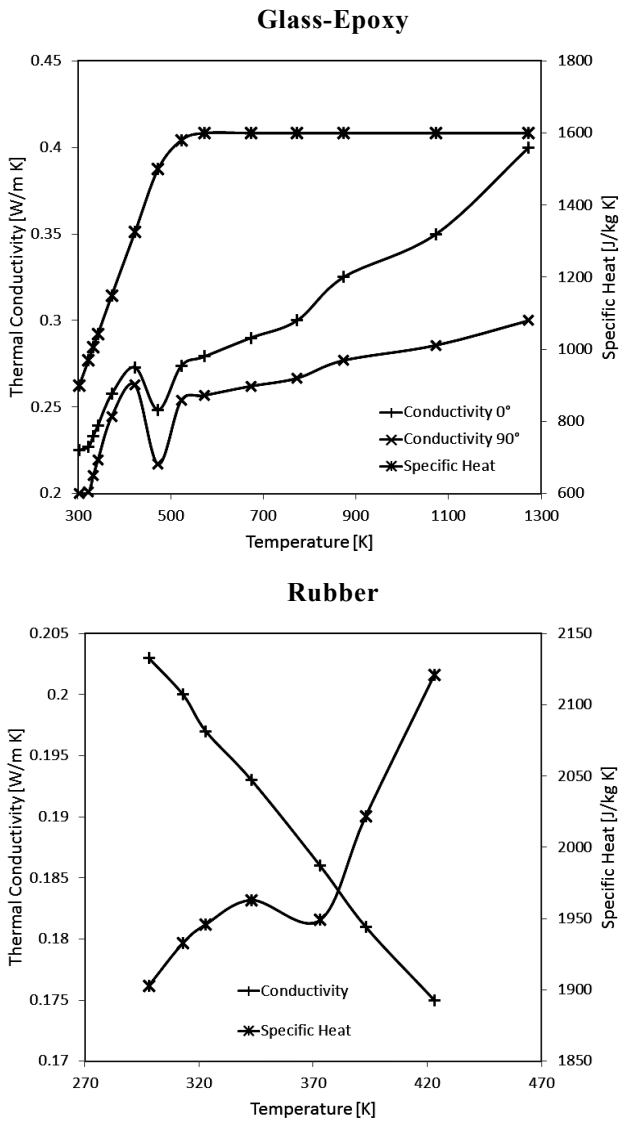


Fig. 5 Thermal material characteristics.

The thermal loading conditions are resumed in a prescribed temperature applied as in Fig. 6. The bounds of the joint subjected to this temperature are in contact with hot gas coming from the combustion chamber. A separate fluid-dynamics study allowed to define it on the surface of the joint with a magnitude of about 1800 K. The rest of the bounds of the model are subjected to adiabatic conditions. Such hypothesis can be done for the presence of a thermal protection (not modelled) on the nozzle structure and support.

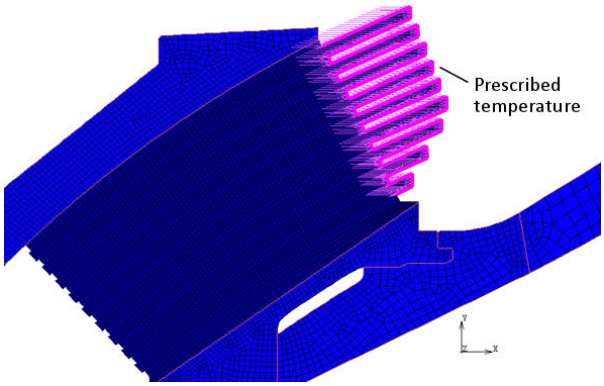


Fig. 6 Thermal loading conditions.

The mechanical conditions can be resumed in a clamping region where the support of the joint is connected with the booster case, a pressure load acting on the nozzle structure and, for the 3D model, a force explicated by the actuators to induce the rotation of the system around the pivot point (Fig. 7). The magnitude for the pressure is 1.1 the maximum expected operating pressure in the combustion chamber (about 10 MPa) corrected with a value taking into account the presence of the thermal protection not modelled.

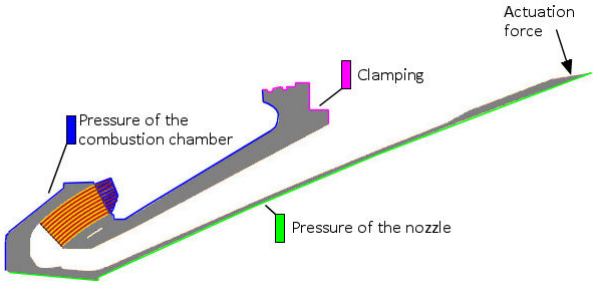


Fig. 7 Mechanical loading and boundary conditions.

6 Thermal and structural analysis

The thermal analysis is performed for a transitory period of 100 seconds with ambient temperature as initial condition. This kind of analysis allows to evaluate, in particular, the temperature distribution along the composite shims and rubber layers. As visible in Fig. 8, the overall protruding composite shims, after the transient period, reach the imposed temperature value, but, the

low values of conductivity of the materials and their high thermal capacities, diffuse the temperature inside the joint only for a little region about 8 mm thin.

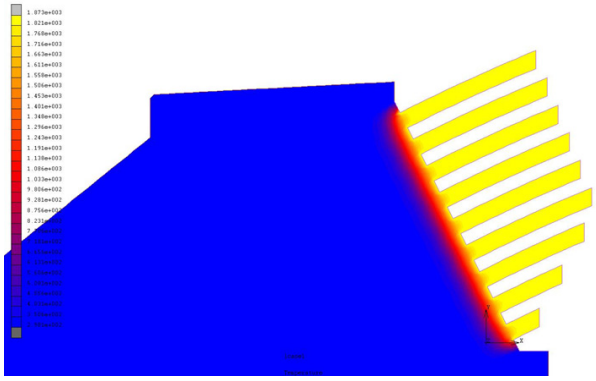


Fig. 8 Thermal analysis: temperature bending plot.

For the mechanical point of view we put in evidence the remarkable effects of the pressure load. The off-set of the pivot point is estimated about 6 mm.

The aluminium parts are subjected to a maximum effective stress of about 340 MPa (Fig. 9) but not in proximity of the high temperatures.

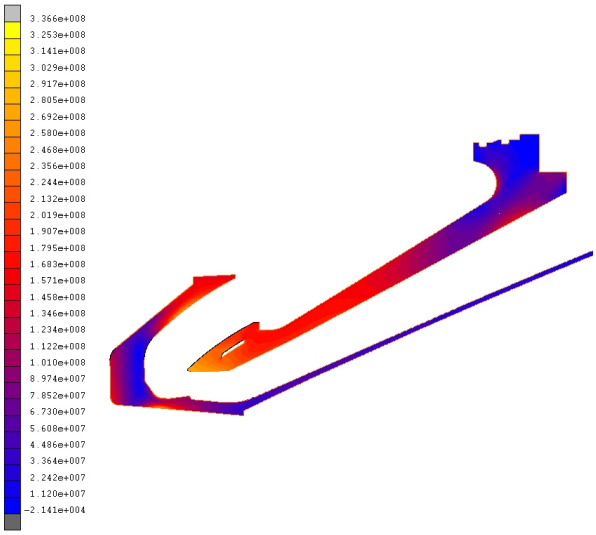


Fig. 9 Mechanical analysis: Von Mises stresses bending plot in aluminium parts.

The most solicited direction of the composite shims is circumferential (Fig. 10) with 130 MPa. The protruded part of the shims bends under pressure and reaches high values of stress. The function of this part is to protect thermally the rest of the joint and it is not structurally relevant. The composite is compressed in circumferential direction up to 80 MPa.

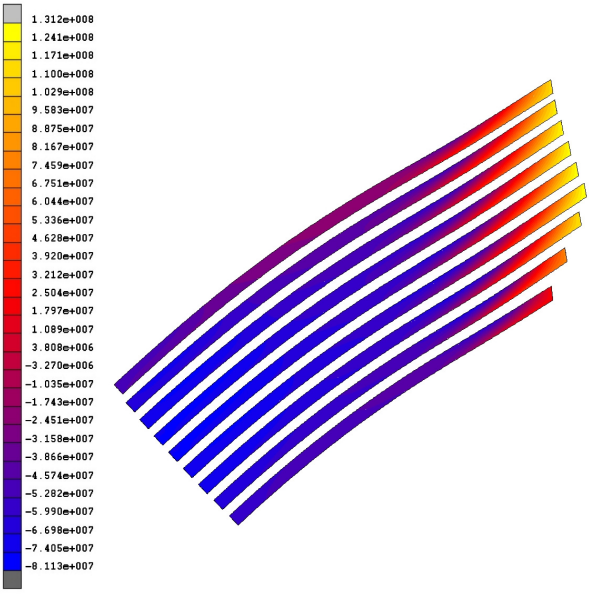


Fig. 10 Mechanical analysis: bending plot of circumferential stress in composite shims.

The rubber layers are subjected to high level of shear strain, especially in proximity of the borders (Fig. 11). In these regions the rubber can elongate up to 520%.

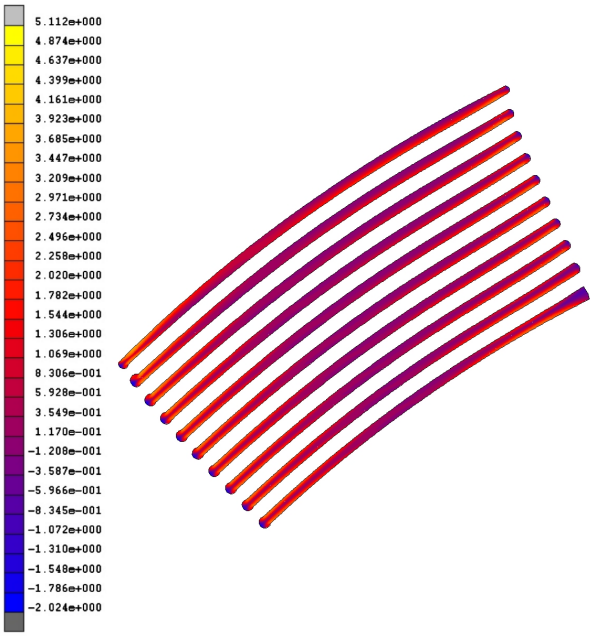


Fig. 11 Mechanical analysis: bending plot of shear stress for rubber layers.

Under actuation load the stress state on the structure increases significantly and it can reach values of 450 MPa for the aluminium (300 MPa in the hot regions) and 120 MPa (in compression) for the composite in circumferential direction.

References

[1] "Solid rocket thrust vector control", NASA SP-8114, 1974

[2] Sutton G. P., Biblarz O., "Rocket propulsion element" John Wiley and Sons, USA, 2001,.

[3] Gjbir Singh, Vankateswara Rao G., "New design formulas for an ex-bearing joint", Journal of Spacecraft and Rockets, 30, 1993, pp. 779-780.

[4] Amin A. F. M. S.,Wiraguna S. I., Bhuiyan A. R., Okui Y., "Hyperelasticity model for finite element analysis of natural and high damping rubbers in compression and shear", Journal of engineering and mechanics, 12, 2006, pp. 54-64.

[5] Bathe K. J., "Finite element procedure", Prentice Hall Inc, Upper Saddle River, New Jersey, USA, 1996.

[6] Ogden R. W., "Non-linear elastic deformations", Dover Publications Inc., Mineola, NY, 1997.

[7] Miller K., "Testing elastomers for hyperelastic material model in finite element analysis", Testing for Hyperelastic, Axel Products, Inc, 2000.

[8] Gent A. N., "Engineering with rubber", Oxford University Press, New York, NY, 1992.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Experimental assessment of not pneumatic wheel for lunar rover technological demonstrator

D. Cardile
diego.cardile@polito.it, *Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy*

N. Viola, A. Rougier and Prof. S. Chiesa
S.P.A.I.C. srl, Corso Einaudi 59, 10129, Torino, Italy

M. A. Perino
Thales Alenia Space – Italy, Strada Antica di Collegno 253, 10146 Torino, Italy

Keywords: rover, elastic wheel, design, test

Abstract

The paper deals with the design, development and test of the elastic not pneumatic wheel of the mobility system of a large Moon rover demonstrator. The study has been accomplished in the framework of STEPS (Systems and Technologies for the ExPloration of Space), which is a research project co-financed by Piedmont Region (Italy), firms and universities of the Piedmont Aerospace District.

The study has started from the conceptual design of the elastic wheel of a large lunar rover (the "nominal wheel"). The activity has proceeded with the design of the elastic wheel of the demonstrator (the "resized wheel"). The manufacturing of a structural similitude model of the technological demonstrator wheel ("scaled wheel") has allowed the revision and validation of the numerical model. Finally, since the numerical model has been subjected to revision, the "resized wheel" design has been revisioned. The paper illustrates the fundamental steps and the main achieved results of both the design of the elastic wheel of the lunar rover demonstrator and the development and test of "scaled wheel".

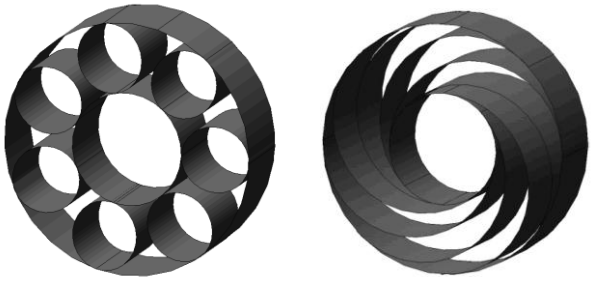
1 Introduction

Lunar rovers will be the key system to perform efficient and safe moon surface operations. Nowadays, the Lunar Roving Vehicle (LRV), the MET and the Lunokhod were the only rovers used on the Moon surface. Their wheels were very different in design and material. The LRV had flexible wheels, constituted by a wire mesh carcass with a stiff inner frame, the MET wheels were made of rubber and the Lunokhod wheels were constituted by a wire carcass connected by spokes. Nevertheless, many studies about possible design solution for moon rover wheels have been carried out. Some very interesting examples are reported in ref. [1]. The present work deals with the conceptual design of the elastic wheel of a large lunar rover demonstrator. The accomplishment of the wheel conceptual design is part of a wider study, which is currently under way, in the framework of STEPS (Systems and Technologies for the ExPloration of Space). STEPS is a research

project, which has been co-financed by Piedmont Region (Italy), firms and universities of the Piedmont Aerospace District. STEPS begun in January 2009 and will end in December 2011 with the ambitious goal of developing and testing new technologies for space exploration including a demonstrator of a future large lunar rover. The study of the wheel of the lunar rover demonstrator has started from the conceptual design of the elastic wheel of the lunar rover (also called the "nominal wheel"). The activity has then proceeded with the design of the elastic wheel of a technological demonstrator of the lunar rover (the "resized wheel"). The manufacturing of a structural similitude model of the technological demonstrator wheel (the "scaled wheel") has allowed the validation of the predicted performance. Eventually, the "resized wheels" have been manufactured, tested and integrated.

2 Nominal wheel

In a previously paper, the authors illustrated and applied a methodology for the conceptual design of the elastic wheel of a lunar rover (the "nominal wheel"). The methodology starts from the top level system requirements and constraints. Through a qualitative and quantitative trade off analysis, based on various parameters, like ride comfort, environment compatibility, stability, motion resistance and thrust developed, the main features of the wheel, as wheel diameter, width and wheel architecture have been selected in order to minimize as much as possible the wheel weight and maximize the drawbar pull. In particular, two wheel architectures have been well investigated, i.e. the ellipse spring wheel and the spiral spring wheel. The ellipse spring wheel consists of several radial elements connected to a central hub and a tread. The central hub is a cylindrical element that provides transfer of axial, radial and circumferential loads, generated by wheel-ground interaction, to the wheel support. The tread is a cylindrical element connected to the radial elements and it shall be deformable enough to ensure the required ground contact area and shall be able to sustain stresses generated by the wheel-ground interaction. The radial elements are necessary to connect the hub and the tread, to provide load bearing and good shock absorption capabilities. The spiral spring wheel is similar to the ellipse spring wheel but the radial elements are a series of bi-tangent semicircles that connect the hub and the tread. In Fig. 1 the two wheel concept are showed. At the end of the design activity the ellipse spring wheel architecture has been selected and the obtained design is summarized in Table 1 and showed in Fig. 2.



(a) (b)

Fig. 1 (a) ellipse spring wheel concept (b) spiral spring wheel concept

Table 1 "Nominal wheel" design summary *wheel stands on a radial element **wheel stands on the gap between two radial elements

Wheel width	*300*	*mm*
Wheel diameter	*1000*	*mm*
Hub diameter	*445.5*	*mm*
Number of radial elements	*7*	*-*
Longer axis of radial elements	*277*	*mm*
Shorter axis of radial elements	*260*	*mm*
Total mass of wheel	*36.5*	*kg*
Maximum sustainable load	*5000*	*N*
Nominal load	*2000*	*N*
Stiffening load	*2500÷3000*	*N*
*Deflection at nominal load**	*70*	*mm*
*Deflection at nominal load***	*71*	*mm*
Vibration intensity at nominal load	*0.68*	*m/s^2*
Material	*Maraging Steel*	

The "nominal wheel" width is 300 mm and the diameter is equal to 1000 mm. The wheel consists of 7 ellipse radial elements, whose longer axis is 277 mm and shorter axis is 260

mm. The ellipse longer axis is aligned with the wheel radial direction. The total wheel mass reaches the value of 36.5 kg. The nominal and maximum sustainable loads are respectively 2000 N and 5000 N.

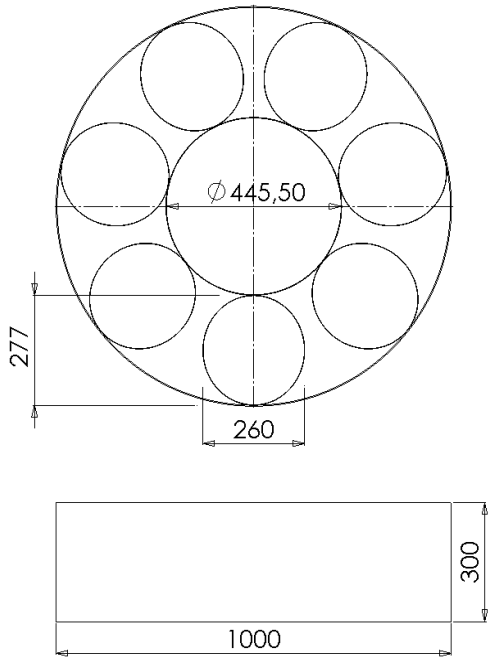


Fig. 2 "Nominal wheel" design

It has been observed that the wheel stiffness behavior changes, if the wheel stands on a radial element (configuration 1 of Table 3) or on the gap between two radial elements (configuration 2 of Table 3). In the first case, the wheel stiffness increases when the deformation leads to a contact between the nearest displaced radial elements. In the second case the wheel stiffness increases continuously with the applied load. For the "nominal wheel", the requirements impose that when the wheel stands on a radial element, the contact between the radial elements occurs for loads bigger than the nominal load. The change of stiffness during rotation causes vibrations transmission to the overall system. The amplitude of acceleration of the final wheel concept is 0.68 m/s^2. This value meets the requirement.

3 Resized wheel

The STEPS (Systems and Technologies for the ExPloration of Space) program is a research project co-financed by Piedmont Region (Italy), firms and universities of the Piedmont Aerospace District of the duration of three years and with the ambitious goal of developing and testing new technologies for space exploration including a demonstrator of a future large lunar rover. In this framework, the S.P.A.I.C. srl has been selected to develop and fabricate six elastic wheels of a lunar Rover Demonstrator. In the first year of activity the authors carried out the studies about the "nominal wheel". They developed and implemented the methodology of design of the elastic wheel. In the second year of activity the authors proceeded with the design of the elastic wheel of a technological demonstrator of the lunar rover and finally in the third year of activity, which is still in progress, the authors engine the production process.

The rover demonstrator is a scaled version of a flight model that has been developed. It includes all the technological demonstrators' technologies coming from all the STEPS work packages. Mainly, the rover demonstrators' dimensions are 3.5 m length, 2.4 wide and 2.4 high. The rover total mass is 1500 kg and it is equipped with 6 motor-wheels with direct drive motors (one per each wheel), manufactured by Sicme Motori.

The configuration of the demonstrator wheel comes from the design activities performed during the first year, even though the sizing is slightly different due to the fact that the wheel requirements have been partially rearranged, thus leading to a new design synthesis, which basically has the same architectural layout of the lunar rover elastic wheel, but slightly different values of the main geometrical characteristics and other materials. The wheel design is showed in Fig. 3 and summarized in Table 2.

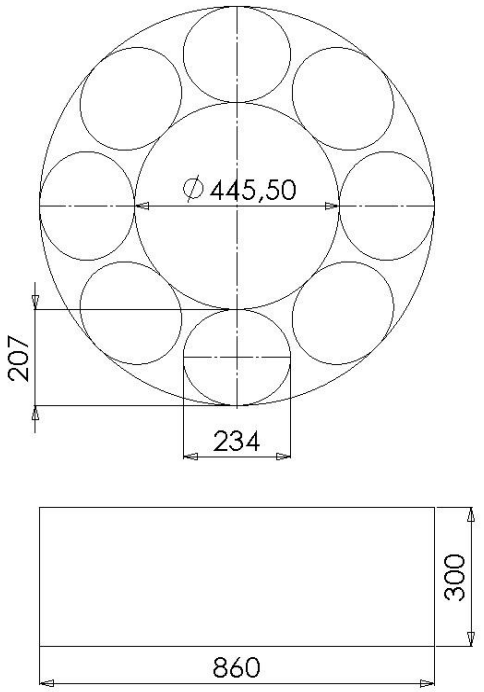


Fig. 3 "Resized wheel" design

The wheel design activity started from the following requirements:

- The wheel shall be able to operate on a soil with physical characteristics similar to the Moon soil.
- The wheel diameter shall be 0,8 m.
- The wheel width shall be 0.3 m.
- The wheel shall be mounted on a motor. The motor diameter is 445.5 mm.
- The wheel shall be able to withstand a nominal vertical load of 2500 N.
- The wheel shall be able to withstand a contingency factor of 2 on the vertical load without permanent deformation. The ultimate load is 5000 N.
- The wheel shall transfer the axial, radial and circumferential loads from ground to the motor and vice versa.
- The wheel shall be able to provide the sufficient traction on the ground. At the nominal load, the minimum contact area shall be equal to 3 dm^2.
- The tread shall be resistant to the ground contact.
- The wheel shall be removable. When the motor-wheel system is still mounted on the suspension, the wheel shall be removable.

After the implementation of the design methodology that has been previously introduced and well described in ref. [2], the wheel design summarized in Table 2 has been obtained. Mainly, the wheel foresees eight radial elements located between the hub and tread. The radial elements are elliptical with longer axis equal to 234 mm and shorter axis equal to 206 mm. The hub diameter is 445.5 mm, the wheel diameter is 860 mm and the wheel width is 300 mm. The material selected is a harmonic steel (C72). The harmonic steel C72 has been selected by analogy with the material considered for the "nominal wheel" and because allows lower production costs. The wheel total mass is 27.5 Kg. The wheel is characterized by a stiffness trend, which increases rapidly, when two adjacent radial elements are in contact as result of wheel deformation. When the radial elements are in contact, the wheel becomes stiffer and the stresses are distributed over a larger number of radial elements. This configuration allows a more efficient use of the wheel material and thus a lighter system with respect to a wheel design that does not allow radial element interaction. The proposed wheel configuration foresees the contact between the radial elements when the applied load reaches 2000 N. This stiffening load has been chosen in order to provide the "resized wheel" with the lower mass possible and in order to ensure the necessary wheel deformation capabilities.

Table 2 "Resized wheel" design summary

Wheel width	*300*	*mm*
Wheel diameter	*860*	*mm*
Hub diameter	*445.5*	*mm*
Number of radial elements	*8*	*-*
Longer axis of radial elements	*234*	*mm*
Shorter axis of radial elements	*207*	*mm*
Total mass of wheel	*27.5*	*kg*
Maximum sustainable load	*5000*	*N*
Nominal load	*2500*	*N*
Stiffening load	*2000*	*N*
Material	*C72*	*-*

4 Scaled wheel

In order to validate the numerical analyses, the authors have proceeded with the study of a structural similitude model of the "resized wheel" according to the structural models theory presented by Giuseppe Gabrielli (ref. [3]). According to this theory, the ratio between the loads applied to the model and the loads applied to real wheel is equal to the square of the scale factor if the material is the same. In particular in the present case, being the "scaled wheel" a 1:3 scale model, the ratio of the applied loads is equal to 1/9.



Fig. 4 A structural similitude model: the "scaled wheel"

The first tests on the "scaled model" have aimed to verifying the capability of bearing the maximum vertical axial and torque loads. The chosen wheel concept has successfully passed these tests, being always beneath the material yield limit, which is the same for the "scaled wheel" and the "resized wheel". As far as the stiffness is concerned, the tests have shown that the "scaled model" stiffness is lower than that numerically estimated by means of the FEM models. The tests have therefore turned out to be quite useful to calibrate the FEM models themselves. The Fig. 5 shows the trend of stiffness calculated by means of the numerical analysis, the trend of stiffness measured by test and the percentage difference amongst them. As can be observed in the Fig. 5, it is evident that that the two stiffness curves have the same trend but also a significant deviation. The average percentage difference is equal to 18.5%. The differences amongst the analytical and test results can be attributed to FEM model simplifies assumptions, numerical approximations, test measurement errors, construction methodology of the "scaled model", material proprieties of the "scaled model", small geometrical differences between the "scaled model" and FEM models.

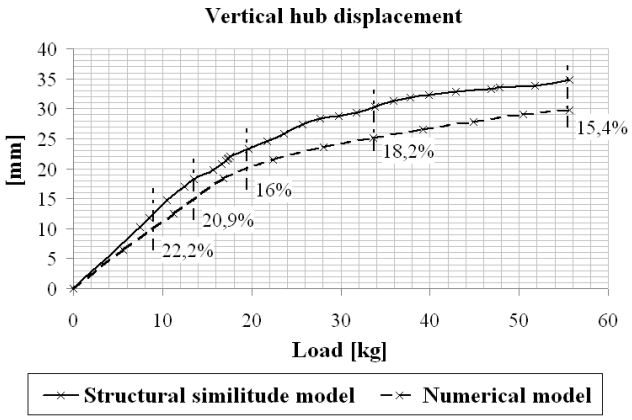


Fig. 5 Vertical hub displacement: test and numerical results comparison

The Fig. 5 and Fig. 6 show respectively the "scaled model" and an image of the tests performed on it. The wheel is mounted on a cylindrical element that reproduces the engine-wheel interface and houses a special beam able to support massive elements at the ends (wheel support). The wheel-wheel support assembly is placed inside a special structure that holds the wheel but allows wheel deflection (holding structure). The Fig. 6 shows the wheel under ultimate load conditions. The test procedure foresees to measure the initial weight of the wheel when it is mounted on the wheel support. Than, the wheel-wheel support assembly is put inside the holding structure so that eventual deflection resistances are annulled. Finally, with a surface gauge, incremental deflection of the wheel due to load increasing can be measured and annotated. The proposed test procedure has been performed several times and the results have been obtained after statistical elaboration.



Fig. 6 Structural similitude model test activity

At the end of the test campaign the numerical model have been revised considering the average difference percentage to correct the results and to obtain as mach as possible the matching of the numerical and test results.

5 Resized wheel design revision

The test results have shown that the developed and implemented analytical model overestimate the wheel stiffness. In particular an average difference percentage of 18.5% has been calculated. Thus considering the test results, a revision of the analytical model and of the wheel design has been considered necessary. In particular, a more stiff design than the previously proposed has been considered necessary. In the following part of this section the system sizing coming from the second design loop is proposed.

The final "resized wheel" design is very similar to the first proposed sizing. The wheel design foresees the same geometry configuration of the first proposed solution but the thickness of the radial elements has been increased. As well explained in ref. [2], the main functions of the radial elements are to connect the hub and the tread, to provide load bearing and good shock absorption capabilities. The thickness and the eccentricity of these elements are the design parameters that allow the desired strength and load absorption features. If the thickness of the radial element increases the wheel become more stiff and the load of increasing stiffness (the wheel become more stiff when the radial elements are in contact due to wheel deformation) as the ultimate load increase. Considering the radial elements eccentricity, if this parameter increases, the radial elements are closer and the load of increasing stiffness decreases. For what concerns the "resized wheel", the increasing of the radial elements metal sheet thickness has increased the wheel stiffness, the load of increasing stiffness and the ultimate load. This has been considered acceptable.

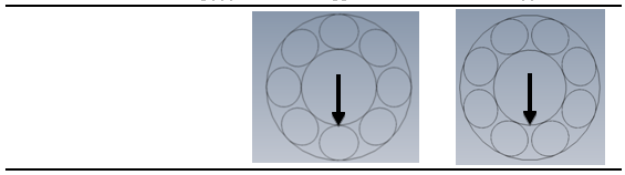
In conclusion the wheel foresees eight radial elements located between the hub and tread. The radial elements are elliptical with longer axis equal to 234 mm and shorter axis equal to 206 mm. The hub diameter is 445.5 mm while the wheel external diameter is 860 mm and the wheel width is 300 mm. The material selected is a harmonic steel (C72). The harmonic steel C72 has been selected by analogy with the material considered for the "nominal wheel" and because allow lower production costs. The wheel total mass is 27.5 Kg.

The following figures show the "resized wheel" vertical displacement for different angles of rotation and load. In particular the Table 3 shows the applied vertical load and the calculated vertical displacement. In Table 3, it is also indicated the nominal load and the load of increasing stiffness when the wheel stands on a radial element (configuration 1). Finally, the Fig. 7 graphically shows the information proposed in Table 3. It can be observed that wheel stiffness behavior changes, if the wheel stands on a single radial element (solid line) or on the gap between two radial elements (dashed line).

In the first case, the wheel stiffness increases when the deformation leads to a contact between the nearest displaced radial elements. In the second case the wheel stiffness increases continuously with the applied load.

Table 3 Stiffness of the wheel when in rotation: vertical hub displacement

Note	Load [N]	Configuration 1 [mm]	Configuration 2 [mm]
	0	0	0
	500	11	26
	1000	22	39
	1500	33	48
	2000	44	55
nominal load	2500	56	60
load of increasing stiffness	3000	66	61
	3500	75	63
	4000	80	64
	4500	85	65
	5000	88	66



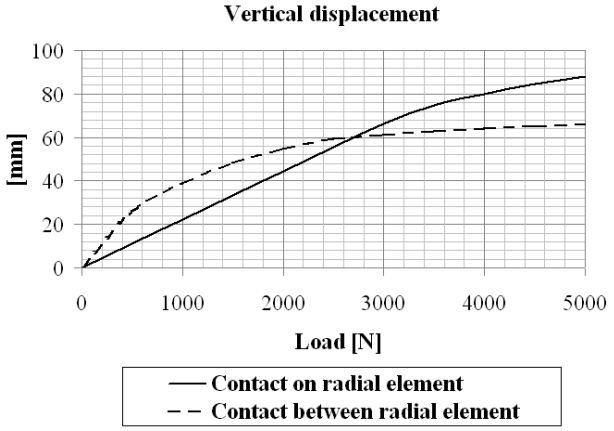


Fig. 7 Stiffness of the wheel when in rotation: vertical hub displacement

6 Conclusion

The paper deals with the design, development and test of the elastic not pneumatic wheel of the mobility system of a medium/large rover for Moon exploration demonstrator. The study has been accomplished in the framework of STEPS (Systems and Technologies for the ExPloration of Space), which is a research project co-financed by Piedmont Region (Italy), firms and universities of the Piedmont Aerospace District.

The study has started from the conceptual design of the elastic wheel of the lunar rover (the "nominal wheel"). Once the design methodology has been briefly illustrated, the most important results about the "nominal wheel" have been summarized. The activity has proceeded with the design of the elastic wheel of a technological demonstrator of the lunar rover (the "resized wheel"). The "resized wheel" has the same architectural layout of the lunar rover elastic wheel, but slightly different values of the main geometrical characteristics and other materials. The rover demonstrator main features have been summarized and the imposed wheel requirements have been illustrated. Once a first loop of design has been completed, the manufacturing of a structural similitude model of the technological demonstrator wheel ("scaled wheel") has allowed test the validation of the wheel performance. Since the test results have shown a significant difference percentage (18.5%) amongst the predicted and the real behavior, the numerical model has been revisioned in order to obtain matching of the performance. Finally, the "resized wheel" has been subjected to a second design loop and the main results have been proposed. The results obtained have been considered encouraging and the "resized wheel" will be manufactured, tested and integrated.

Acknowledgments

The authors wish to thank P. Pellegrino, S. Ferraris, A. Cernusco of Thales Alenia Space, S. Galli of Altran Italia and R. Lamberti, D. Artioli of Sicme Motori for their precious support to wheel system development.

References

[1] Asnani V., Delap D., Creager C., "The Development of Wheels for the Lunar Roving Vehicle", NASA, 2009.

[2] Cardile D., Viola N., Chiesa S.,Rougier A., "Applied design methodology for lunar rover elastic wheel", to be published.

[3] Gabrielli G., *Lezioni sulla scienza del progetto degli aeromobili, Vol. I e Vol. II,* Levrotto e Bella - Torino, 1974.

CEAS 2011

ATM Operations Optimization

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Prediction, Detection and Avoidance of Wake Vortices Using Onboard Systems

S. Schönhals, M. Steen and P. Hecker
Institute of Flight Guidance, Technische Universität Braunschweig, Germany

Keywords: *wake vortex awareness, 4D-trajectory, onboard systems, model/sensor fusion*

Abstract

In current operations, the hazard posed by wake vortices to the following aircraft is mitigated by use of separations. Future concepts including 4D-trajectories and self separations will require an adapted approach including the (partial) delegation of separation responsibility to the flight crew and the creation of wake vortex awareness or, higher levels of automation provided, also avoidance or mitigation to support the pilots. An on-board Wake Vortex Awareness and avoidance System (WVAS) could take this part making use of the information available on-board and the wake prediction and detection approaches.

The concept of WVAS as well as the development and implementation of an experimental system at the Institute of Flight Guidance is presented. Aspects of the development and evaluation of this concept are addressed.

1 Introduction

SESAR in Europe and NextGen in the USA will significantly change the current Air Traffic Management (ATM) system e.g. with the introduction of new operational procedures designed to meet the challenges of future air traffic (need for efficient, green and punctual flights as well as growing demand for en-route and airport capacity). Therefore, it is often referred to as the "Next Generation ATM System". One important characteristic of this system will be the delegation of separation authority and responsibility towards a higher involvement of the airspace users through a collaborative negotiation of agreed 4D-trajectories. This will on the one hand require an improved capability to estimate the own precise position, velocity and other significant states, on the other hand also the information about the states of other airspace users and other potential obstacles will play a significant role. This other obstacles can be terrain, weather but also the wake vortices generated by preceding aircraft.

The separation from wake vortices in the future air space will essentially differ from the procedures applied today that are based on fixed distances depending on aircraft weight categories. A self-separation from wake vortices

is postulated e.g. by the SESAR Master Plan [1]. For this purpose, an increased awareness of hazardous wake vortices is needed on-board the aircraft. This need will even increase when other aspects of 4D Precise Trajectory Clearances are taken into account that will bring about more climbing and descending movements and crossing of tracks thus increasing the probability of wake encounters.

This paper will describe the impact of these changes on wake vortex related procedures and compare them to today's situation. The need for increased wake vortex awareness will be explained and a concept for a dedicated on-board system will be derived.

The work done in this respect by the Institute of Flight Guidance of Technische Universität Braunschweig will be presented and the development and implementation of an on-board wake vortex visualization concept will be introduced.

Finally, further objectives and the development towards an enhanced wake vortex awareness system will be presented.

2 Background

2.1 Wake vortex procedures today

Today's minimum safe separation distance between aircraft is mainly governed by the need to avoid the hazard associated with the wake turbulence generated by the preceding aircraft. There are two main factors defining the potential consequences of a wake encounter that can result in mild disturbance or a catastrophic event. These are the relative sizes of aircraft in-trail and the way an aircraft encounters the wake. Today, only the relative size, for which the main parameter is the weight of an aircraft, is taken into account and results in the current wake separations based on aircraft weight categories.

The recommended ICAO separations [2] for take-off and landing are widely applied, although national variations exist. Currently, these separations are regarded as safe but in many cases overly conservative (e.g. during crosswind or turbulent conditions). Considerable effort is made to develop new procedures [3] or to revise the existing wake vortex categories [4] in order to improve efficiency of flights and to increase capacity of congested airports.

However, there is also a difference between wake vortex procedures during take-off and landing in the terminal area and during en-route flight phases. For example, during the approach the traffic density is high and the aircraft are following each other on the ILS glide path. It is necessary for the pilot to be aware of the situations where wake turbulence may be encountered and take measures to avoid it. Therefore, common procedures during approach and departure phases of flight, where aircraft follow in-trail of other aircraft and the wake turbulence risk increases, depend on the available visibility conditions. In visual meteorological conditions (VMC) when other aircraft can be seen and avoided, the responsibility to maintain safe separations is usually given to the pilot of the follower aircraft. He is able to estimate the location of the wake from the generator aircraft position and try to avoid it. However it should be noted that an accurate estimation is complicated by the circumstance that decay and transport of the wake vary from aircraft to aircraft and also depend on meteorological conditions like wind and turbulence.

During instrument meteorological conditions (IMC) when no visual contact to preceding aircraft is possible and the pilot cannot judge the location of the predecessor's wake, responsibility for adequate spacing is taken entirely by Air Traffic Control (ATC). The controller has no further information about wake vortex location (e.g. whether it has been transported out of the transport corridor by wind, as it is the case under most atmospheric conditions). Therefore, the only means for ATC to ensure wake vortex safety in such situations is establishing the recommended conservative spacing between aircraft in-trail. A change of procedures that might be introduced by new concepts would aim at increasing the capacity but at the same time maintaining at least the same safety level achieved today.

During the en-route flight phase, it is assumed that the radar separation will suffice to prevent critical encounters, as the self-induced downward velocity will cause the wake to sink out of the way of the next aircraft following at the same flight level. However, wake vortex encounters in cruise have been reported and the investigation presented by previous studies [5] has pointed out the increased probability of encounters in Reduced Vertical Separation Minimum (RVSM) airspace. These investigations have shown that encounters were most likely when either the wake generating or wake penetrating aircraft are ascending or descending a cruise altitude.

2.2 SESAR operational concept impact on wake vortex separation and responsibility

Both in Europe and in the USA, the SESAR and NextGen programs have been launched that are aiming at a paradigm shift in management of airspace. In this paper, the SESAR Master Plan [1] will be taken as a reference for ease of use, however, many aspects are valid for the changes anticipated from NextGen as well. With SESAR, a shift from tactical separation to strategic management of traffic flows in high density airspace will take place. In particular, after the ATM Service Level 5 will be available for operations, the following new separation modes will be relevant for wake vortex safety:

- 4D Precision Trajectory Clearances (4D-PTC)
- Airborne Separation Assistance System with Self-separation operations (ASAS-SSEP)
- Airborne self-separation from wake vortices (ASEP-WV)

With the implementation of 4D-PTC, flights will follow their optimised, efficient and green trajectories. The steps towards the implementation of 4D-PTC in Europe is shown in Fig. 1. This will on the one hand optimise the use of airspace but on the other hand also increase the number of flight path crossings. The probability of climb-through and descend-through events will increase which will magnify the risk of encountering the wake vortex of a preceding aircraft and will require actions to at least maintain the safety level achieved today.

When aircraft will achieve ASAS-SSEP capability, more responsibility for the separation will be delegated to the aircraft and will require the flight crew to be aware of the traffic situation around them. Not only the safe separation from other aircraft will be necessary but also the separation from the surrounding wake vortex traffic will have to be established by the pilots.

For this reason, SESAR envisages that an airborne self-separation from wake vortices will be available on-board. In such a system wake vortices will be regarded as just another obstacle that the aircraft has to avoid. For this purpose, the same information as about other flight objects need to be provided to the aircraft, namely the wake vortices' 4D-trajectories. For this reason, research and development of an on-board Wake Vortex Awareness System (WVAS) is needed to provide the required concepts and technology for the implementation of these future ATM concepts.

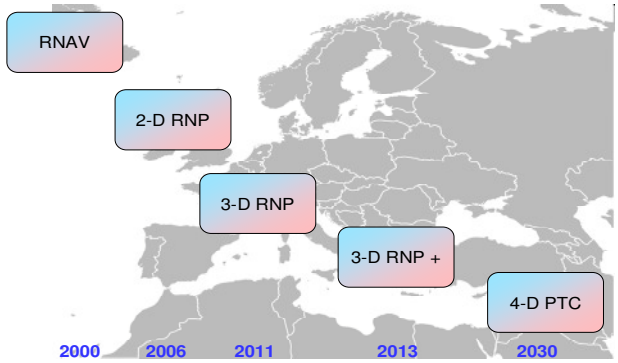


Fig. 1 Timeline of implementation of 4D-PTC in Europe

3 Concept for prediction, detection and avoidance of wake vortices in onboard systems

The central idea of the WVAS concept is to increase the flight crew awareness of wake turbulence hazards and to enable avoidance and/or alleviation of wake encounters. For this purpose, prediction and detection of dangerous wake vortices can be used in order to provide the necessary data for manoeuvres.

Mathematical models exist that describe and predict wake vortex generation, evolution, transport and decay based on atmospheric and aircraft inputs. The issue of modelling wake vortices is extremely complex, so only some of its general characteristics will be briefly explained here for ease of understanding.

The models coming into focus for operational use with a fused system are the so-called fast-time models that reliably predict vortex characteristics as vortex strength (called circulation) and decay as well as position in faster than real-time. They were developed using theoretical concepts according to the underlying physical principles and calibrated with empirical data. Some of the best known representatives are NASA's Aircraft Vortex Spacing System Prediction Algorithm (APA) [6], UCL's Deterministic Wake Vortex Model (DVM) [7] and finally the model developed by DLR called the Deterministic 2 Phase Wake Vortex Decay and Transport Model (D2P) [8]-[10]. DVM and D2P have been further developed to deliver probabilistic predictions to account for the uncertainties of model input. The main model inputs are parameters of the generating aircraft, namely weight, airspeed, wingspan and position, as well as meteorological data (e.g. wind vector over height, turbulence and temperature). Fig. 2 illustrates the aircraft parameters that influence the wake generation.

Wake vortex detection (either ground based or for on-board applications) is realized by means of remote sensing technology. Here, especially the RADAR and LIDAR sensors come into focus, both utilizing radio waves to sense movement of the ambient air masses. For the purpose of this paper, exemplary the wake detection by LIDAR should be explained in more detail.

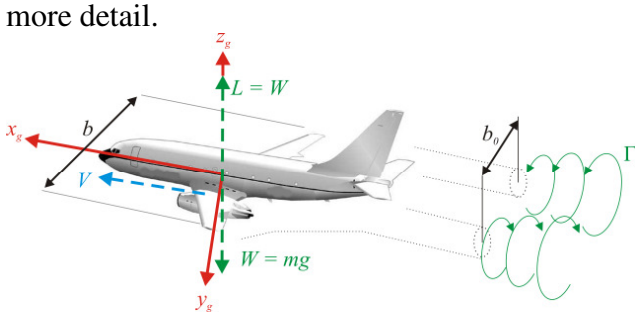


Fig. 2 Wake turbulence generating aircraft with the respective wake vortex parameters

The LIDAR system is measuring the line-of-sight velocity of the aerosols in a certain scanning plane. From the velocity distribution the position of the wake vortex is determined e.g. via the slopes of the tangential velocity distribution. Several methods to determine the wake vortex circulation from this distribution are available.

The major advantage of wake vortex monitoring by LIDAR or RADAR lies in the physical turbulence detection which is complementary to model prediction. As the flow field in an operational environment is very complex, the automated processes of retrieving wake vortex position and strength from the measured quantities are prone to errors like over- or underestimation or even failure of recognition of mature vortices from the velocity field and suffers in addition of the measurement noise.

The general problem of airborne wake vortex detection and tracking is depicted in Fig. 3, where the wake vortex generated by the preceding aircraft is being tracked in a certain scanning plane by the follower by use of a dedicated sensor installed onboard this aircraft. The range r and bearing Θ of the vortex are measured onboard in the body-fixed coordinate system of the follower.

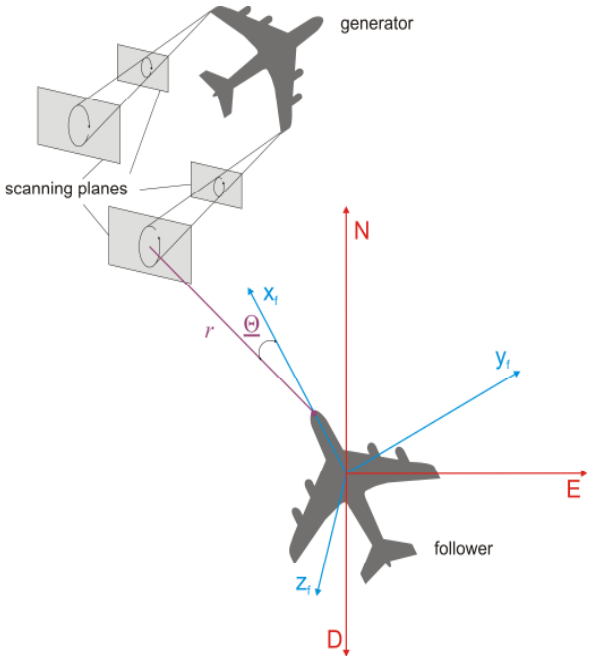


Fig. 3 Detecting and tracking wake vortices onboard

For sole wake alerting, model-based prediction using information about the wake generating aircraft already looks promising compared to the turbulence information the pilots have today (i.e. none). However, the uncertainty of a sole model solution depends on the accuracy of the input parameters, as stated above. As long as important factors, e.g. the generator's weight or the wind vector at the position where the wake is created, are not provided via data links, safety margins will still be required. Therefore, adequate aiding by a dedicated wake vortex on-board sensor could help to increase system accuracy and integrity.

If, in addition to the alerting function, on-board wake surveillance information should be used for purposes of automated wake alleviation through flight control functions (e.g. to reduce the effect of wake turbulence on the own aircraft if an encounter can not be avoided), more characteristics of the wake velocity field are required.

In this case an on-board sensor is preferable, because it is likely that the numerous influence factors needed by prediction models for a detailed prediction of the vortex flow field will not be available onboard.

The principle would be comparable to the Airborne Collision Avoidance System (ACAS) that serves as a safety net for cases when safe separations are not guaranteed by the standard procedures and that directly measures the real time-to-go (also known as the TAU-criterion) until the Closest Point of Approach (CPA). For wake encounter avoidance, the same function could be taken by the onboard sensor.

The system also must be integrated into the aircraft's avionic system in order to receive the parameters it requires for wake prediction and detection as shown in Fig. 4. In this case, the wake vortex model and the sensor data processing is integrated inside the fused airborne wake vortex surveillance module (FAWVS). The sensor would interact with the FAWVS through a dedicated interface allowing delivering measurement information to correct the prediction and in turn would be controlled and aided in acquisition and tracking.

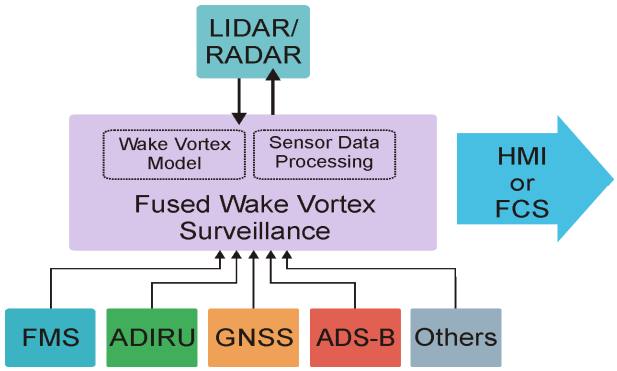


Fig. 4 Scheme of possible interfaces with the aircraft avionics

The system also must be integrated into the aircraft's avionic system in order to receive the parameters it requires for wake prediction and detection as shown in Fig. 4. These data are:

- position and velocity of ownship from the GNSS unit
- traffic information via data link (e.g. ADS-B)
- planned trajectory from the Flight Management System (FMS)
- attitude and heading from the Air Data Inertial Reference Unit (ADIRU)
- wind and weather information from a Weather Information and Monitoring System (WIMS, as the one developed in FLYSAFE [11])

A detailed discussion of the required interfaces is given in [12].

In the system shown above the wake vortex model and the sensor data processing is integrated inside the fused airborne wake vortex surveillance module (FAWVS). The sensor would interact with the FAWVS through a dedicated interface allowing delivering measurement information to correct the prediction and in turn would be controlled and aided in acquisition and tracking.

The output of the system could e.g. contain vortex hazard areas indicating also the remaining circulation intensity level, inflated with a safety factor to account for the remaining uncertainty. This information could be passed on to a Human Machine Interface (HMI) in order to display the no-fly areas to the pilot (comparable to the already existing traffic display used for collision avoidance).

An automated wake alleviation or avoidance of dangerous turbulence zones can be achieved when the FAWVS is connected to the Flight Control System (FCS) of the aircraft.

The advantage of the use of such a fused system adequately integrated into the existing aircraft avionics is the increased reliability of wake prediction and detection results compared to stand alone sensor or model solutions. The pilot would receive additional information on the aircraft environment that would provide better situational awareness. Pilots will be given the possibility to make separation decisions on a more complete information basis which is an important precondition for safe operations as the ones foreseen in the future ATM systems. Also automated wake conflict solution can be enabled.

The concept described above is also applicable to different equipment levels, e.g. if no on-board sensor will be available the system could still deliver wake vortex information based on the prediction function but with a higher level of uncertainty. The same applies to the precision and availability of other information sources listed above.

4 Development and implementation of WVAS

The Institute of Flight Guidance (Institut für Flugführung, IFF) of the Technische Universität Braunschweig is involved in several research activities dealing with the determination and graphical presentation of wake vortices. To achieve a maximum degree of usability, the integration of wake vortex information into existing cockpit displays is one possible approach. Another possibility is to connect the system with the Flight Control System of the aircraft if automated wake alleviation shall be implemented; for conflict resolution, the Flight Management and Guidance System could be addressed. An onboard Wake Vortex Awareness System according to the concept described above is being developed and integrated into an Airbus A320 cockpit simulator that is part of the modular aviation simulation environment of the IFF. This environment will be described in the following.

4.1 Simulation Environment

An essential part during the development process of new concepts and systems deals with the validation and evaluation in order to ensure their suitability for real operation conditions. Because testing in a real environment might be complicated and not feasible in many situations mainly due to safety issues and also costs, simulation is a common approach.

The IFF owns a simulation environment that is used for demonstration of existing and future air traffic management concepts and human machine interface technologies, educational purposes as well as research activities within national and international projects focusing on the development of future assistance systems. It consists of simulators for both cockpit and air traffic control side. While each component is standalone and can be operated individually, it is also possible to collaboratively run the whole system. In the following, the single components are briefly presented.

For the flight deck part, an Airbus A320 cockpit simulator is available. It is fix-based and uses the Microsoft Flight Simulator X (FSX) as simulation software. The FSX is widely-used, inexpensive (but not open source) and meets the compatibility requirements of the used avionic systems simulation software packages. Another advantage is the realistic outside view and the possibility to extend its default scenery set (e.g. by a highly detailed one of a specific airport). Focusing on HMI development requires a preferably realistic visualization of the aircraft environment instead of focusing on a maximum detailed and precise flight physics simulation.

The stationary A320 simulator (Fig. 5) has a triple channel external view using three projectors, each plugged to an individual computer. This allows a viewing angle larger than 180°.



Fig. 5 Displays and instruments of the A320 flight deck simulator

A multi-channel audio system is used to simulate the intercom, radio messaging and warning sounds. While the physical flight model is based on the FSX, the system simulation is realized by a proprietary software package comprising all cockpit displays, input devices and their logic.

For the ATC part, the simulation environment consists of a controller working position and an airport traffic simulator. The controller working position is based on a HMI technology that has been implemented in a former research project. The display has been developed to support airport surface controllers and is currently extended to meet the requirements of an en-route controller working position. The airport traffic simulator is used to generate traffic on an aerodrome surface and for approach and departure flight phases. The position and movement information of all simulated aircraft as well as the complete communication between ATC controllers and the flight crews (e.g. clearance requests and clearances) are provided to all other components of the simulation environment using different data interfaces.

4.2 Developing the WVAS in the simulation environment

The wake vortex awareness system that is currently developed using the simulation environment of the IFF is based on the fused airborne wake vortex surveillance module which integrates the functions of wake vortex prediction and wake detection. The prediction function is performed by a fast-time wake vortex model as described above. The detection is provided by an emulation of an onboard LIDAR sensor which takes into account the characteristics of such a sensor. Both model and sensor information are fused using an estimation filter developed by the IFF for the purpose of wake vortex surveillance and providing a more accurate and reliable information than the sole prediction or detection functions. Possible implementation examples of such fused systems as well as their advantages over only sensor or only model solution were presented by the IFF e.g. in [13].

For the determination of wake vortex position and strength, following parameters are read from the FSX using the mentioned software framework:

- weather data
  - air density
  - air temperature
  - air pressure
  - wind velocity
- aircraft data
  - position
  - orientation
  - velocity
  - weight
  - wingspan

For real operation conditions the system needs to be integrated into the aircraft's avionic systems to receive all required parameters for wake vortex prediction. However, these parameters are available from already existing systems and data interfaces on-board.

In order to increase the pilot's awareness of the wake vortex hazard and to enable the flight crew of adjusting the appropriate wake vortex separation, a representation of the surrounding wake vortex traffic is possible. Fig. 6 and Fig. 7 show a possible visualization solution for the wake hazard information currently investigated at the Institute of Flight Guidance.

This visualization concept has been developed for the existing simulation infrastructure of IFF and is designed for the

cockpit of the Airbus A320. It follows the already existing concepts of the Predictive Windshear System (PWS) and the Traffic Alert and Collision Avoidance System (TCAS). While the Navigation Display (ND) provides a global overview of current wake vortex situation, the Primary Flight Display (PFD) offers tactical information to support the pilot's decision making. In order to prevent a wake turbulence encounter, it is not reasonable to visualize the position and strength of each individual vortex. Instead, it is adequate to display only the overall danger zone of the predicted wake vortices.

Based on the symbolism and colour coding of the TCAS, the wake vortex awareness system consists of the following three warning levels: normal (white), caution (amber) and warning (red). Depending on the time a potential hazardous wake of surrounding traffic can be reached, each danger zone is assigned to one warning level and visualized by its corresponding color. The current overall operating status equates to the most critical warning level. In addition, an aural warning is given when either the caution or the warning level is reached. The acoustic warning "WAKE AHEAD" is played once when the dangerous turbulence zone can be reached in less than 40 seconds, the warning "WAKE" is issued repeatedly when the wake is within 25 seconds reach of the own aircraft. The logic explaining when and how wake warnings are issued is presented in Table 1.

Table 1 Visual and acoustic warnings

Operating Status	Display on ND	Display on PFD	Acoustic Warning
NORMAL	No display or white polygon if requested by the pilot	-	-
CAUTION	Amber polygon	Amber rectangle	WAKE AHEAD played once
WARNING	Red polygon	Red and amber rectangle	WAKE played repeatedly

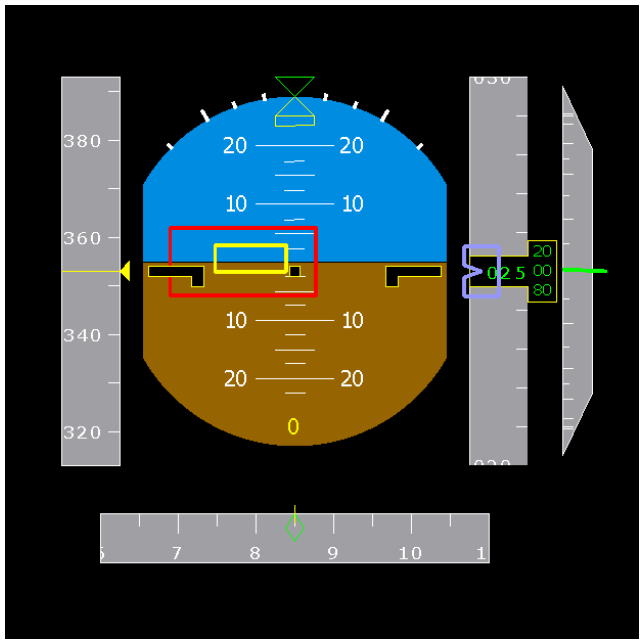


Fig. 6 Representation of wake hazard areas on the PFD

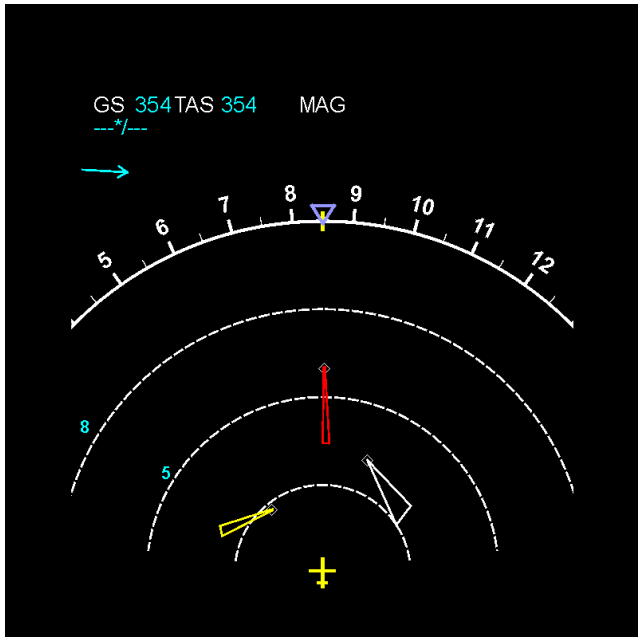


Fig. 7 Representation of wake hazard areas on the ND

For the development of a wake vortex awareness system it is furthermore essential that a close collaboration between research and the stakeholders (e.g. pilots, controllers, airlines and service providers) is ensured. This has been initiated at a very early phase during the concept evolvement of the wake vortex display described above. On the one hand, questionnaires and interview with pilots during the development phase as well as simulator flight tests in the validation phase were an

important part of the work above. On the other hand, considerations of aspects of general human machine interface design as well as the concepts of modern airliner cockpits were taken into account. In the next development phase, it is envisioned to regard the procedural separation aspects and to integrate the WVAS into the ATC environment. For this purpose, an interaction between pilots and controllers shall be simulated and feedback from both sides will be considered.

5 Conclusion

The paper has presented the current wake vortex separation procedures as well as the impact which is anticipated for the future ATM operational concepts introduced by implementation of 4D-PTC, Self-separation and wake vortex spacing ASEP that are envisioned by SESAR. A need to develop dedicated wake vortex awareness systems for onboard use has been identified.

A concept for an on-board wake vortex awareness system has been presented and discussed. An implementation of such an on-board system developed and tested by the Institute of Flight Guidance by means of simulation has been introduced.

Further work is projected in this field including the advancement of the wake visualization and refinement of the prediction and detection functions by using a fused approach. Another important field of study will be the integration of the wake awareness system into the ATC environment in order to test and evaluate new wake separation procedure and how they can be enhanced by use of the WVAS.

Moreover, future research work at the IFF will focus on the combination of different technologies supporting the flight crew during the approach phase. That means that onboard assistance systems for an increased situational awareness (like the presented wake vortex visualization concept) will collaborate with approach guiding systems like Ground Based Augmentation System (GBAS) and ATC assistance systems. The overall goal is to have automated support applications for multiple approach procedures (e.g. different approach angles to avoid preceding wake vortices) to give the cockpit and tower crew assistance without increasing their workload. The different subsystems developed at the IFF will be combined in the presented simulation environment to provide evaluation and demonstration capabilities for these concepts toward higher degrees of automation in ATM.

Acknowledgment

Parts of the work (FAWVS) presented in this paper are part of a Research Grant co-financed by the European Organisation for the Safety of Air Navigation (EUROCONTROL) under its Research Grant scheme. The content of the work does not necessarily reflect the official position of EUROCONTROL on the matter.© 2008-2011, EUROCONTROL and the Technische Universität Braunschweig. All Rights reserved.

References

The papers in the reference list must be cited in the text. In the text the citation should appear in square brackets “[ ]”. In the Reference list the font should be Times New Roman with 10 points size. The authors names list should be terminated by a colon. The citation number should be enclosed in square brackets.

In the reference list, give all authors’ names; do not use “et al.” unless there are six authors or more. Private communications and personal Web sites should appear as footnotes rather than in the reference list.

References should be cited according to a standard publication reference style (for examples, see the “References” section of this template). Names and locations of publishers should be listed.

[1] SESAR Consortium, “SESAR Master Plan: Deliverable 5,” Brussels, 2008, p. 38.
[2] ICAO Doc 4444 PANS-ATM Fifteenth Edition 2007, Chapter 4, Paragraph 4.9.1.
[3] Eriksen P., “SESAR Joint Undertaking wake vortex research (WP6.8.1)“, Presentation at the

Global Wake Vortex Conference II, San Diego, December 2010.

[4] Tittsworth J., "Wake Re-Categorization phase I results", Presentation at the Global Wake Vortex Conference II, San Diego, December 2010.

[5] Nelson R. C., "Trailing vortex wake encounters at altitude - a potential flight safety issue?", AIAA Paper 2006-6268.

[6] Robins R. E., Delisi D. P., "NWRA AVOSS wake vortex prediction algorithm vrsion 3.1.1", NASA/CR-2002-211746, Hampton, Virginia, National Aeronautics and Space Administration, 2002.

[7] De Visscher I., Winckelmans G., Desenfans O., Lonfils T., "Overview of UCL operational tools for predicting aircraft wake vortex transport and decay: the Deterministic/Probabilistic wake Vortex Models (DVM/PVM) and the WAKE4D platform", Louvain-la-Neuve, Belgium, Universite catholique de Louvain (UCL), 2008.

[8] Holzaepfel F., "Probabilistic two-phase wake vortex decay and transport model", Journal of Aircraft, Vol. 40, No. 2, American Institute of Aeronautics and Astronautics, 2003, pp. 323-331.

[9] Holzaepfel F., "Probabilistic two-phase aircraft wake-vortex model: further development and assessment", Journal of Aircraft, Vol. 43, No. 3, American Institute of Aeronautics and Astronautics, 2006, pp. 700-708.

[10] Holzaepfel F., Steen M., "Aircraft wake-vortex evolution in ground proximity: analysis and parameterization", Journal of Aircraft, Vol. 45, No. 1, American Institute of Aeronautics and Astronautics, 2007, pp. 218-227.

[11] Gerz T., "Weather information management systems (WIMS) and onboard applications", Presentation at the FLYSAFE Final User Forum, Amsterdam, 2009.

[12] Schönhals S., Steen M., Hecker P., "Surveillance systems on-board aircraft: predicting, detecting and tracking wake vortices", 8th Innovative Research Workshop & Exhibition (INO 2009), EUROCONTROL Experimental Centre Brétigny-sur-Orge, 2009.

[13] Schoenhals S., Steen M., Hecker P., „Towards wake vortex safety and capacity increase: the integrated fusion approach and its demands on prediction models and detection sensors", 27th Congress of the Aeronautical Sciences (ICAS 2010), Nice, 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Climate impact assessment of varying cruise flight altitudes applying the CATS simulation approach

Alexander Koch[a], Benjamin Lührs[a], Katrin Dahlmann[b], Florian Linke[a], Volker Grewe[b], Markus Litz[c], Martin Plohr[d], Björn Nagel[a], Volker Gollnick[a] and Ulrich Schumann[b]

German Aerospace Center (DLR)
[a] Air Transportation Systems, Blohmstraße 18, 21079 Hamburg, Germany
[b] Atmospheric Physics, Münchner Straße 20, 82234 Oberpfaffenhofen-Wessling, Germany
[d] Simulation and Software Technology, Linder Höhe, 51147 Köln, Germany
[d] Propulsion Technology, Linder Höhe, 51147 Köln, Germany

Keywords: *Climate compatible Air Transport System, Climate impact assessment, Operational measures, Varying cruise altitudes, Distributed Design Environment.*

Abstract

The present paper describes a comprehensive assessment and modelling approach that was developed in the DLR project Climate compatible Air Transport System (CATS) with the goal to analyze different options to reduce the climate impact of aviation.

The CATS simulation chain is applied to assess the climate impact reduction potential (via CO_2, contrail-cirrus, H_2O, NO_x, ozone, methane, primary mode ozone) for the world fleet of a representative long-range aircraft operated on a global route network in the year 2006.

The average temperature response (ATR) and the direct operating costs (DOC) are calculated for flights with varying cruise flight altitudes and speeds.

The obtained results are expressed as relative changes with respect to the minimum DOC trajectory and assessed as cost-benefit ratio (ATR vs. DOC). The results are highlighted for a single route and transferred to the global route network, showing a large potential to reduce the climate impact of aviation for small to moderate increments on costs.

1 Introduction

Aviation has an influence on global warming through the emission of gaseous compounds, aerosols, particle matter and induced cloudiness.

Aviation accounted in 2005 for a total radiative forcing of 43.2 mW/m^2 (median) excluding the impact from linear contrails and aviation induced cirrus clouds [1]. A new study reports model based estimates for the global climate impact from contrail induced cloudiness (CIC) of 31 mW/m^2 for the year 2002, including the impact from linear contrails, contrail-cirrus clouds and resulting changes in natural cirrus cloudiness [2]. This estimate highlights the relevance of induced cloudiness in any aviation climate impact study.

Without any further measures the projected air traffic growth of 4.8% per year till 2036 [3] will largely surpass the estimated annual fuel efficiency improvements of 1-2% [4]. The rise of annual emissions rates will hence further increase the climate impact from aviation.

The Advisory Council for Aeronautical Research in Europe (ACARE) states in this sense that a social and climate compatible air transportation system is required for a

sustainable development of commercial aviation. To achieve such a sustainable air transport system, mitigation strategies have to be developed based on comprehensive assessments of the different impacting factors and reduction potentials.

The global climate impact from air traffic varies not only with the amount and type of emitted species, but also with altitude and latitude of emission and the underlying atmospheric conditions. The atmospheric response results from complex interrelated processes of very different nature, spatial and temporal scales.

The assessment of the aviation related climate impact still holds large uncertainties. However Grewe et al. (2007) [4] and Dahlmann et al. (2009) [6] showed that a comparative evaluation of new aircraft technologies or operational procedures is possible by minimizing the uncertainties in climate impact changes by adequate Monte-Carlo simulations.

The assessment of options to reduce the climate impact from aviation by new technologies and operations requires expert knowledge from different disciplines and adequate models that sufficiently incorporate the driving impact factors. Such a comprehensive simulation and analysis approach was developed within the DLR project *Climate compatible Air Transport System* (CATS) and presented for the first time by the authors in 2009 [7].

In the present paper the authors present an enhanced version of the CATS-simulation chain and its application to evaluate operational measures to reduce the climate impact of air traffic. The analyzed measures are independent from individual weather situations; they are based on climatologic mean values derived from daily variations in weather situations. Different cruise flight altitudes and speeds are assessed for the world fleet of a representative long range twin-jet configuration operated on a global route network containing 1178 routes. The study discusses the potential to mitigate climate impact expressed as changes in global average temperature response (ATR) and direct operating costs (DOC).

2 Climate impact from aviation

The climate impact from air traffic results from induced cloudiness and concentration changes of the atmospheric constituents caused by the emission of carbon dioxides (CO_2), nitrogen oxides (NO_x), sulfur oxides (SO_x), water vapour (H_2O), aerosols and particle matter such as soot [8]. These atmospheric perturbations change the terrestrial radiation balance and cause a radiative forcing (RF) that drives the earth-atmosphere system to a new state of equilibrium through a resulting temperature change.

CO_2 is an effective greenhouse gas and one of the major contributors to the global warming (28 mW/m^2 in 2005) caused by air traffic. The impact of CO_2 is independent from the location of emission due its long atmospheric lifetime.

But also non-CO_2 effects have a large impact on the radiative forcing as displayed in Figure 1.

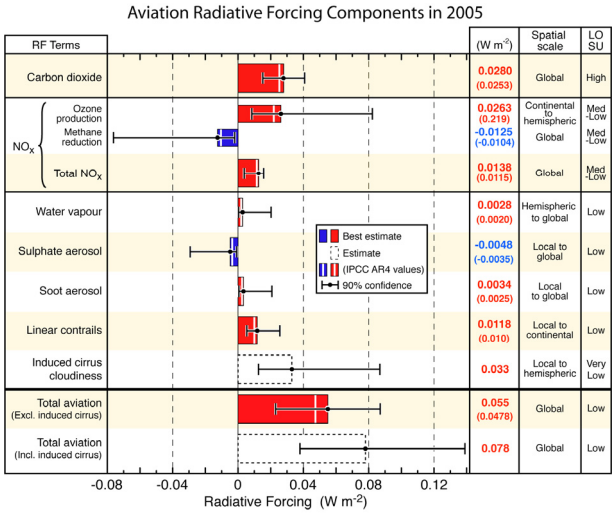


Fig. 1 Radiative forcings from global aviation as evaluated from preindustrial times until 2005 [1].

NO_x emissions from subsonic air traffic in the upper troposphere and lower stratosphere enhance the formation of ozone (O_3) and depletion of methane (CH_4), both compounds being strong greenhouse gases. The increased ozone concentration causes a warming effect (positive RF) whereas the reduction of atmospheric methane has a cooling effect (negative RF). Still, the net radiative forcing of NO_x is positive (13.8 mW/m^2) as the impact of ozone prevails. The impact from NO_x emissions on the concentration change of ozone and

methane is sensitive to altitude and latitude. The maximum net radiative forcing is found at the tropical tropopause and decreases towards lower altitudes and higher latitudes [10]. This spatial dependency needs to be considered in the applied climate response model to correctly reflect the impact from aviation.

Volatile aerosols forming in young aircraft plumes have a small direct impact on the radiation budget through scattering and reflecting shortwave radiation (-4.8 mW/m^2) [12].

Aircraft-emitted soot particles have a small direct warming effect through absorption of shortwave and reemission of long-wave radiation (3.4 mW/m^2). Soot particles are key contrail-forming agents and are suspected to induce cirrus cloud formation. The latter effect is very uncertain because it is not known if these particles nucleate ice efficiently [12].

The impact resulting from H_2O emitted at typical sub-sonic flight levels is comparatively small (2.8 mW/m^2) due to its small impact on the background concentration of H_2O.

Contrail induced cloudiness consists of linear contrails, contrail-cirrus clouds and of changes in the occurrence or properties of natural cirrus clouds. The formation and persistence of linear contrails depends on the atmospheric conditions, which have to be supersaturated with respect to ice, hence sufficiently cold and humid [13][14]. This results in a clear altitude dependency [15]. Linear contrails transform to contrail-cirrus clouds under favourable meteorological conditions, able to cover large areas. For fixed conditions, the impact of contrails depends mainly on the coverage and optical depth [16]. The global climate impact from CIC, including the impact from linear contrails, contrail-cirrus clouds and resulting changes in natural cirrus cloudiness, accounts for 31 mW/m^2 [2]. CIC hence represents besides the impacts from NO_X and CO_2 a major contributor to climate change from air traffic.

Any climate response model applied to assess the impact from aircraft should consider the formation and impact of linear contrails and contrail cirrus, either by considering real weather situations [9] or climatologic averaged probabilities depending on altitude and latitude [10]. Further, any climate impact assessment not only requires the comprehensive modelling of non-CO_2 effects, but also a reasonable climate metric. Studies have shown that metrics like global warming potentials (GWP) or radiative forcing index (RFI) are easily misleading and not always appropriate for aviation studies [17]. Another approach to estimate the climate impact is the change of near-surface air temperature, as presented by Sausen and Schumann (2000) [18] or Grewe and Stenke (2008) [10]. The authors believe that the change of global average temperature response (ATR) [19] is the most suitable metric for the purpose of the present study.

$$ATR_H = \frac{1}{H}\int_{t}^{t+H} \Delta T(t)dt \quad (1)$$

The presented metric integrates the surface temperature change $\Delta T(t)$ over a chosen time period H (for this study 100 years) considering thus impact of short-lived (e.g. contrails) and long-lived (e.g. CO_2) forcing agents in appropriate way.

3 Modelling the climate impact from aviation with the CATS approach

As outlined above, the climate impact from aviation depends among other factors within the earth-ocean-atmosphere system primarily on the emitted amount of the different species, the geographical position and altitude of emission as well as on the atmospheric background conditions.

The chemical and physical atmospheric processes that follow the emission of aircraft exhaust gases and lead to radiative forcing and temperature changes are partly interdependent and occur on very different spatial and temporal scales [3]. The identification and quantification of dependencies and driving factors within these highly complex processes requires large disciplinary expertise, simulation and experimental resources. Also the assessment of operational flight procedures, unconventional aircraft configurations and novel technologies asks for disciplinary expertise and proper model capabilities in the fields of air traffic

management, flight performance and aircraft design with its related sub-disciplines aerodynamics, structural design, systems, propulsion, etc. Altogether the required expert knowledge and model capabilities goes beyond the means of a single research engineer. Instead a group of experts and disciplinary models is required to reliably answer multidisciplinary research questions with such scale of complexity.

Within the DLR research project *Climate compatible Air Transport System* (CATS) such a comprehensive approach was developed to analyze the climate impact from air traffic, allowing the environmental and economical evaluation of new aircraft technologies, designs and operational procedures [7]. The CATS approach comprises models for appropriate physical simulation of aircraft design and performance, propulsion and emissions, mission calculation, climate response and operating costs. All models are developed and supervised by experts from the related fields. The setup of simulations and analysis of results is done in a collaborative way ensuring reliable conclusions.

3.1 Collaborative analysis capability

As argued above, the assessment of complex multidisciplinary topics requires a range of experts and models. Commonly the experts from research and/or industry entities are not located at the same location but regionally distributed.

This leads to the need of a distributed design and analysis environment that links the integrated disciplinary analysis models and provides means for the remote triggering, overall process control, convergence and optimization. The chosen model-based architecture and the underlying software engineering techniques influence the efficiency of the resulting analysis processes [20]. With increasing complexity of the model chain the number of interfaces is the critical factor for the flexibility of a design environment. Therefore a common namespace (central data model) is required for the efficient exchange of data and increased flexibility in model substitution [20].

Activities to establish such a collaborative analysis capabilities at DLR have started in 2005. Based on these activities, the central data model *Common Parametric Aircraft Configuration Scheme* (CPACS)[ii] and modelling environment *Remote Component Environment* (RCE)[viii] have been developed and applied in various projects. Both, RCE and CPACS will be made open source in 2011.

RCE [22] allows the linking of models provided by DLR entities and industry partners. The workflows can be started locally on the computer of the integrator or transferred to another platform. The models run on their server at the respective partner sites. CPACS [20][21] is based on XML technologies and includes the parametric description of the atmosphere, aircraft and engine performance, mission profiles, airports, fleet network as well as the resulting trajectories, climate impact and cost break down. Figure 2 shows the principle of collaborative design and analysis process as it is applied in CATS.

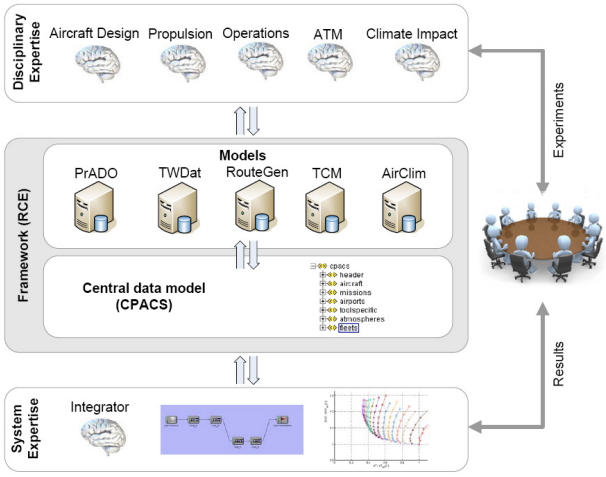


Fig. 2 DLR collaborative design and analysis process as applied in CATS.

3.2 CATS simulation chain

The integrated analysis models are provided by several DLR institutions and academia as listed at the end of this document (index *i-viii*). Plausibility of single model and overall simulation results are ensured by the involved experts in collaborative way. The outlined simulation process is displayed in Figure 3 and shows the simulation sequence of models.

The multi-disciplinary aircraft design tool *Preliminary Aircraft Design and Optimization*[viii] (*PrADO*) developed by Technical University Braunschweig is applied to calculate the flight performance and technical characteristics of actual and novel aircraft configurations. *PrADO* comprises physical models with semi-empirical extensions for aerodynamics, structural analysis, weight prediction, flight performance incl. trim calculations and geometry description [22]. *PrADO* is further applied to determine the influence of aircraft subsystems on engine performance through bleed air and shaft power extraction.

The surrogate database model *TWdat*[iii] provides engine performance maps for several actual engines and possible future propulsion concepts, which are pre-calculated by the well-established thermodynamic cycle program *Varcycle*[iii] [24] and fitted to real engine data where possible. The performance maps contain emission indices (i.e. CO, NO_x, UHC, soot), thrust and fuel flow characteristics.

Models for preliminary flight preparation (*RouteGen*[ii] and *FuelEstimator*[ii]) provide relevant data concerning the mission profile (airports, lateral flight path, yearly frequencies) and estimation of required mission fuel and resulting payload limitations for all analyzed routes.

The *Trajectory Calculation Module* (*TCM*)[ii] is applied to calculate detailed emission inventories with 4D trajectories. *TCM* performs a fast-time simulation integrating the relevant flight conditions based on the total energy model [25]. It reads as input mission parameters, aircraft weight breakdown, engine and aerodynamic performance tables for different high-lift configurations provided by *PrADO* and *TWDat*. This capability enables the flight performance simulation and evaluation of novel aircraft concepts.

The model *FlightEnvelope*[ii] checks each calculated flight trajectory if aircraft specific operating constraints (stall, buffet and altitude limits) are violated. In such case the concerning trajectories are removed from the dataset.

Atmospheric data such as e.g. annual mean temperature, pressure, relative humidity, wind vectors as function of latitude and altitude are provided by the model *Atmos*[i].

The climate impact of each flight is assessed with the climate response model *AirClim*[i] [10][11]. The model comprises a linearization of a climate-chemistry model from the emission to radiative forcing, resulting in an estimate in near surface temperature change, which is presumed to be a reasonable indicator for climate change. AirClim is designed to be applicable to aircraft technology, considering the altitude and latitude of emission for the climate agents CO_2, H_2O, CH_4, O_3 and primary ozone mode (PMO) (latter three resulting from NO_x emissions), line-shaped contrails and contrails cirrus clouds. Combining aircraft emission data with a set of previous calculated atmospheric perturbations, AirClim calculates the temporal evolution of radiative forcing and resulting global near surface temperature change. The pre-calculated data are derived from 78 steady-state simulations for the year 2000 with the DLR climate-chemistry model *E39/CA*, prescribing normalized emissions of nitrogen oxides and water vapour at various atmospheric regions [26].

The economical impact is calculated as direct operating costs (DOC)[ii] of each flight including the costs for fuel, crew, maintenance, navigation and landing fees and financing [27].

Depending on the scope of study, different process control scripts are integrated for the variation of routes or aircraft configurations.

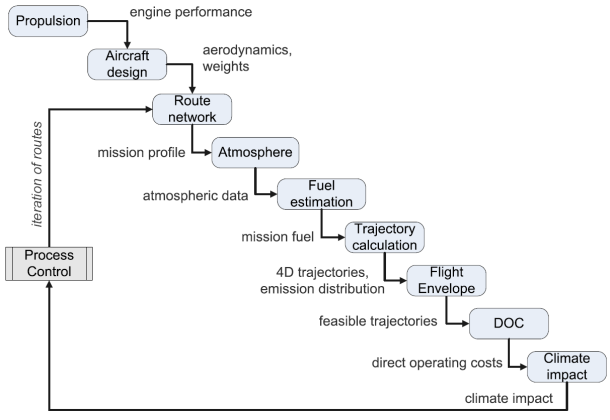


Fig. 3 Simulation sequence as applied in CATS.

4 Calculating the climate impact of varying cruise flight altitudes and speeds

The goal of the present study is to apply the CATS approach to identify the potential of climate impact reduction and related DOC changes through varying cruise flight altitudes and speeds based on the world fleet of a representative long range aircraft.

4.1 Setup of experiment

Therefore all routes operated in the year 2006 by an Airbus A330-200 are derived from OAG data [28] resulting in a set of 1178 global distributed city pair connections with corresponding flight frequencies.

The reference aircraft is modelled to fit the geometry, weights and performance characteristics of the real aircraft equipped with CF6-80E1A3 engines. Table 1 specifies the basic model characteristics and Figure 4 shows the resulting model geometry. Figure 5 displays the model payload-range performance in comparison to the real aircraft [29]. Note that the different slope in the payload range diagram comes from a slightly better SFC in the engine model. This leads to a 3.5% lower max take-off weight and an increased ferry range of 3.7%. Scaling of the model SFC is considered not appropriate because of the not fully quantifiable impact on emission indices, which are considered of higher priority to the study goal than a perfect fit of the reference aircraft.

Design conditions		
Design range	7861	km
Design cruise Mach	0.82	-
Design payload	253 PAX + 23.75 t	
Design ICA	10000	m
Resulting model characteristics		
Operational empty weight	116.2	t
Max take-off weight	221.5	t
FAR Take-off field length	2389	m
FAR Landing field length	1690	m
Approach speed	75.5	m/s
L/D @ design conditions	19.9	-
SFC @ design conditions	0.05876	kg/N/h

Table 1 Model performance characteristics

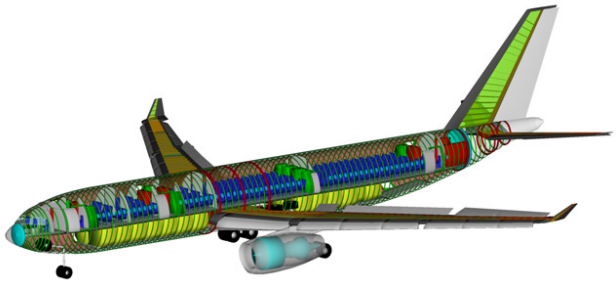


Fig. 4 Geometry model of reference aircraft.

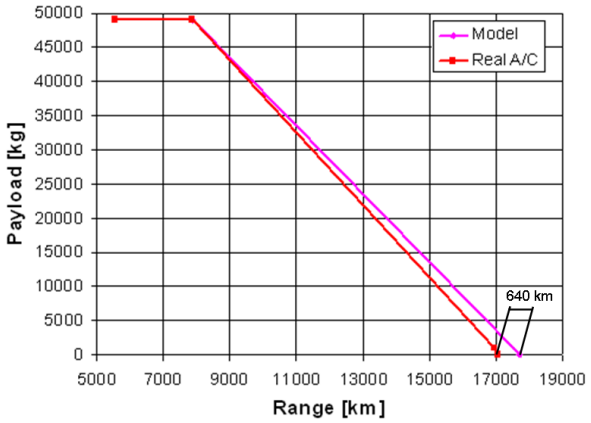


Fig. 5 Payload range comparison of model vs. real A/C

The vertical flight profile is based on typical flight phases and respects (where possible) the common air traffic control (ATC) speed and altitude constraints during climb and descend. The cruise phase is modelled as continuous climb cruise at constant lift coefficient. This is assumed an appropriate approximation to a real cruise flight with step climbs, which are normally optimized by flight planning tools for the specific route considering the actual transport conditions and ATC restrictions.

The operative payload is considered with an average passenger load factor of 0.76 and additional 5000 kg of cargo [3].

The DOC [$/cycle] are calculated without the cost of ownership (depreciation, interests, insurance) in order to asses the increment on flight operation costs only (cash operating costs). The average fuel price in 2006 (0.595 USD/kg) is derived from [30], cost values for labour (25 USD/h) and fees are based on [27] and scaled by the average US inflation rate (2.66%) from 1993-2006 [31].

The climate impact is calculated for sustained emissions over 32 years, which corresponds to the average lifetime of the reference aircraft [3]. The evolution of background emissions of CO_2 and CH_4 follow the IPCC scenario A1B [32] while the development of background air traffic for contrails-cirrus assessment is based on data from the *QUANTIFY* emission inventory [33] scaled by IPCC scenario Fa1 [8]. The average temperature response is evaluated for the time frame of H=100 years (ATR_{100}) starting in 2006.

For each route (index i), cruise Mach numbers from 0.4 to 0.85 in steps of 0.025 and initial cruise altitudes (ICA) from 3962-12497m (=13kft to 41kft) in steps of 305m (=1kft) are analyzed with the described settings. Each Ma_{cr}–ICA combination represents one cruise operating point. The authors would like to emphasize that the wide ranges of ICA and Ma_{cr} values are chosen to identify the maximum potential of climate impact reduction, reflecting less the actual ATC practice.

4.2 Evaluation methodology

For each cruise operating point and resulting trajectory (index *k*) the reduction potential is expressed as cost-benefit ratio (ATR_{rel} vs. DOC_{rel}) relative to the route-specific cruise operating point for minimum DOC (DOC_{min}).

There are still many uncertainties in the calculation of the climate impact of air traffic [1]. To account for these uncertainties a Monte-Carlo simulation is performed in such way that AirClim calculates a minimum and maximum temperature change for each species and each uncertainty specified in [10]. For each species 10000 random values α_j are picked within the given distribution of the uncertainty range for each uncertainty parameter (Radiative Forcing and climate sensitivity for each species, as well as stratospheric and tropospheric lifetimes). The respective climate impact for each species is calculated for each step j of the Monte Carlo simulation

(2) $ATR_{spec}(j) = \alpha_j ATR_{min}(j) + (1 - \alpha_j) ATR_{max}(j)$

Based on this, the relative difference (φ) between the trajectory *k* and reference trajectory for each step j of the Monte Carlo simulation can be determined.

$$(3) \qquad \varphi_k(j) = \frac{ATR^k_{spec}(j) - ATR^{ref}_{spec}(j)}{ATR^{ref}_{spec}(j)}$$

Where $ATR^k_{spec}(j)$ is the average temperature change resulting from the impact of a certain species on trajectory *k* and $ATR^{ref}_{spec}(j)$ is the temperature change for the reference trajectory, in this study the route-specific cruise operating point for minimum DOC (DOC_{min}). The total climate impact $ATR^k_{all}(j)$ for one trajectory and route is given by the sum of impacts from all species. For the different $\varphi_k(j)$ resulting from the Monte Carlo simulation, the median and the 2.5, 25, 75 and 97.5% percentiles are calculated and serve as uncertainty range. This specific Monte-Carlo calculation takes advantage of the lower uncertainty of the relative difference between the climate impacts of two scenarios compared to their individual absolute values [11].

The ATR_i and DOC_i changes relative to $DOC_{min,i}$ are calculated for all feasible operating points (Ma_{cr}, ICA) on route (*i*).

$$(4)\ DOC_{rel,i}(Ma_{cr}, ICA) = \frac{DOC_i(Ma_{cr}, ICA)}{DOC_{\min,i}}$$

$$(5)\ ATR_{rel,i}(Ma_{cr}, ICA) = \frac{ATR_i(Ma_{cr}, ICA)}{ATR_i(Ma_{cr}, ICA)_{DOC_{\min,i}}}$$

Plotting $DOC_{rel,i}$ vs. $ATR_{rel,i}$ for all calculated operating points in one diagram per route yields a Pareto front of best solutions that displays the climate impact reduction potential as function of the related increment in DOC relative to the route specific operating point for minimum DOC. As such relation is only given for computed Ma_{cr}–ICA combinations, the information about the corresponding operating point is lost for interpolations between Pareto elements. To determine hence the largest relative climate impact reduction for an accepted maximum cost increment *x* (e.g. *x*=1.10) on each route, the route-specific Pareto front is intersected at the defined *x* value and

evaluated for the next smaller available operating point on the curve. The resulting $DOC_{rel,i}(\tilde{x})$ and $ATR_{rel,i}(\tilde{x})$ values have thus to be considered as best solution $\tilde{x}$ that fulfils the specified cost limit x on a specific route.

The $DOC_{min,i}$ operating point and corresponding absolute ATR_i vary for each route depending on fuel burn, flight time and latitude of emission.

To estimate the reduction potential resulting from all routes ($ATR_{rel,all}(\tilde{x})$), every Pareto front is intersected at the specified maximum DOC increment x and evaluated at the next smaller Pareto Element $\tilde{x}$. The related $ATR_i(\tilde{x})$ values are weighted by the route specific flight frequency (f_i) and summed for all routes.

$$(6)\quad ATR_{rel,all}(\tilde{x}) = \frac{\sum_{i=1}^{n\_routes} f_i \cdot ATR_i(Ma_{cr}(\tilde{x}), ICA(\tilde{x}))}{\sum_{i=1}^{n\_routes} f_i \cdot ATR_i(Ma_{cr}, ICA)_{DOC\,min,i}}$$

The same approach is applied to determine the overall $DOC_{rel,all}(\tilde{x})$ and uncertainty range for $ATR_{rel,all}(\tilde{x})$, which is based on the percentiles of each specific route.

5 Discussion of results

To provide an overview about the study results and trends the route Detroit-Frankfurt (DTW-FRA) is discussed exemplarily. Figures 6 and 7 show DOC_{rel} and ATR_{rel} respectively for varying cruise Mach number with fixed ICA [a] and vice-versa [b]. The operating point for minimum DOC on route FRA-DTW was found at Ma=0.85 and ICA=10973m.

Figure 6a shows the evolution of DOC_{rel} with varying Ma_{cr} for constant ICA. For this case the total DOC_{rel} curve shows a variation of 43% over the plotted Ma range and follows largely the trend of the time dependent costs (crew, maintenance). The fuel costs are proportional to the fuel burn and show in this case a minimum for Ma_{cr}=0.675, rising to lower and higher speeds due to the increased drag and required thrust.

Figure 6b displays DOC_{rel} for fixed Ma_{cr} and varying ICA, which experiences a variation of 9% over the analyzed ICA range. The impact of ICA on the time dependent costs is in this case very small due to subtle changes in flight time. This leads to a total DOC evolution that follows the fuel costs curve, which has a minimum at ICA=10973m.

These two graphs indicate that the evolution of DOC is primarily driven by the cruise Mach number through its impact on maintenance and crew costs. Note that this trend (decreasing total DOC towards higher speeds) is influenced by the costs assumptions in the study scenario [27], which assumes a relative low fuel price compared to the labor costs.

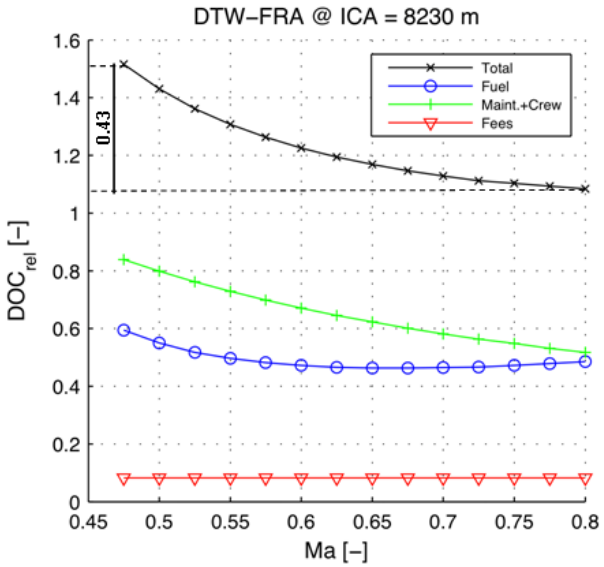


Fig. 6a Relative DOC changes for ICA=8230m and varying cruise Mach numbers on route DTW-FRA.

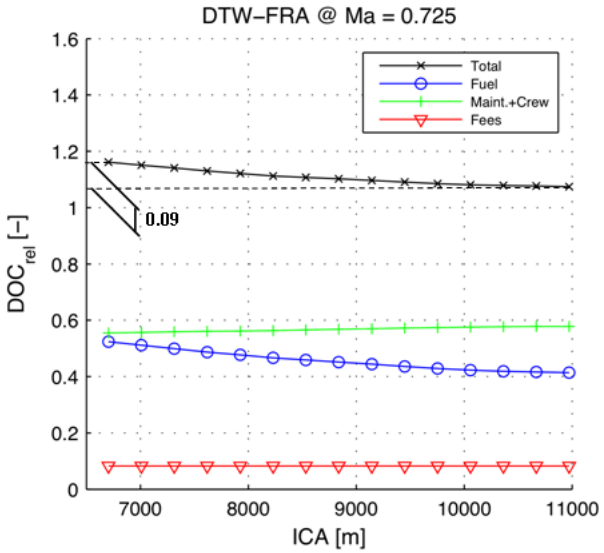


Fig. 6b Relative DOC changes for Ma=0.725 and varying ICA's on route DTW-FRA.

Figure 7a displays the evolution of ATR_{rel} with varying Ma_{cr} for constant ICA, showing a variation of 12% between the minimum at Ma=0.625 and maximum at Ma=0.475. In this case the amount and related impact of emitted CO_2 is proportional to the fuel burn (3.15 kg CO_2/kg fuel) and follows the trend displayed in Fig. 6a, whereas the impacts from O_3 and CH_4 are related to the emitted amount of NO_x (at constant altitude). At both ends of the fuel burn curve the NO_x emission index (kg NO_x/kg fuel) rises and provokes thus an increased impact from O_3 (warming) and from CH_4 (cooling). The formation and related impact from contrails and contrail-cirrus (index *cont*) is not sensitive to speed changes as the formation of contrails is mainly dependent on altitude. The impact resulting from PMO is proportional to CH_4. The impact from H_20 is almost independent from Ma_{cr}, showing only a very small dependency from fuel burn at constant altitude.

Figure 7b shows ATR_{rel} for fixed Ma_{cr} and varying ICA, exhibiting a variation of 32% over the analyzed ICA range with increasing reduction potential for decreasing altitudes. Here the impact from O_3 dominates at higher altitudes, increasing steeply above approx. 9500m. This increase is largely caused by the rising sensitivity of O_3 production with higher altitudes, whereas the increase in emitted NO_x contributes also but is not the dominant effect. CH_4 displays a light dependency from altitude, its cooling impact diminishing slightly with increasing altitude. The same applies to PMO, being proportional to CH_4, only at smaller magnitude. The formation of contrails and contrail-cirrus has its maximum on this route at ICA=9449m, decreasing to lower and higher altitudes. The CO_2 impact decreases for higher altitudes as fuel burn decreases (compare Fig. 6b). The emitted amount of H_20 is proportional to fuel burn (1.25 kg H_20/kg fuel) and thus decreases with altitude, whereas the impact from H_20 instead increases due its increasing atmospheric lifetime. PMO shows a relative small dependency from altitude.

Figures 7a-7b highlight that the sole minimization of CO_2 (fuel burn) does not lead to the minimum climate impact. In fact, the minimization of CO_2 leads to higher altitudes while the minimum climate impact is obtained at lower altitudes (Fig. 7b).

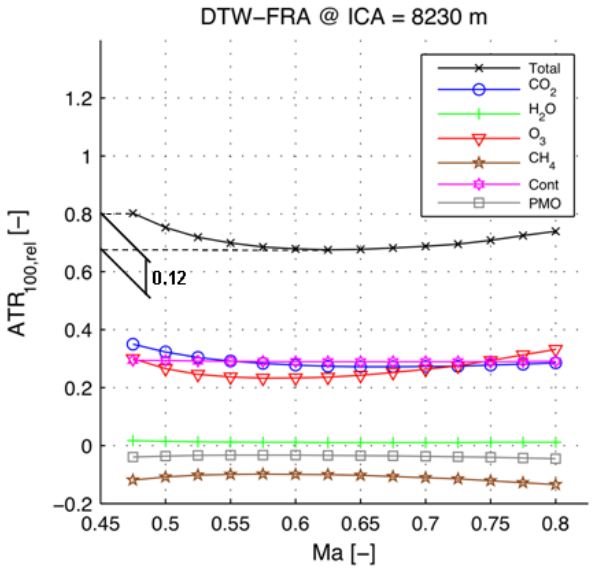


Fig. 7a Relative ATR changes for ICA=8230m and varying cruise Mach numbers on route DTW-FRA.

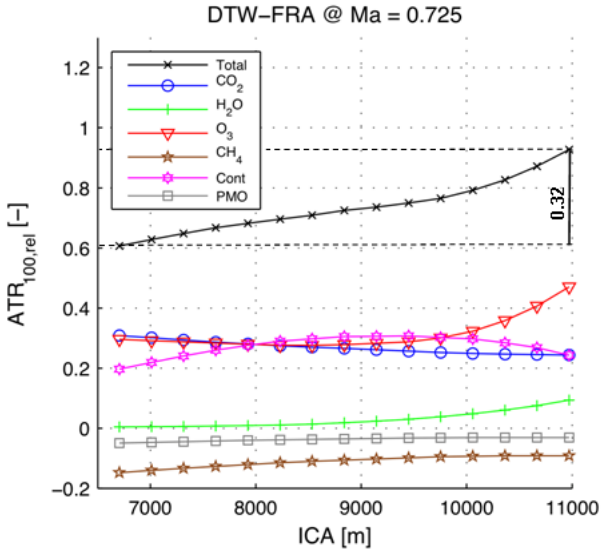


Fig. 7b Relative ATR changes for Ma=0.725 and varying ICA's on route DTW-FRA.

When comparing the relative importance of the different forcing agents it shows that impacts resulting from induced cloudiness, CO_2 and O_3 are having the largest contribution to the total climate impact. Among those, the impacts from O_3 and induced cloudiness show the largest variation due to ICA changes.

The reduction of cruise altitude therefore leads to a reduction in climate impact, even with slightly increased fuel burn and related CO_2

impact. Resulting from the discussion of Figures 6a-6b the sensitivity of DOC with altitude changes is small in comparison to speed changes, which leads to the conclusion that a reduction of cruise altitude will have only a moderate impact on DOC.

This relation can also be observed in Figure 8 that shows the resulting ATR_{rel} and DOC_{rel} for all calculated operating points for the route DTW-FRA. The operating points are indexed for constant ICA values with increasing Mach numbers moving down the curves for constant ICA. The Pareto front contains all the Pareto optimal Ma_{cr} – ICA combinations between the reference point (DOC_{min}) and the operating point for maximum achievable ATR_{rel}.

Intersecting the DTW-FRA Pareto front exemplarily for a maximum DOC increment of 10% (x=1.10), the closest calculated operating point that fulfils this condition is found at ICA=8534 (light brown triangles) and Ma_{cr}=0.75. At this point a relative climate impact reduction of 28% (median), $ATR_{rel}(\tilde{x})$=0.72, is achieved with a corresponding DOC increment of 9.3%.

This ratio shows that a relatively large climate impact reduction is achievable for a comparably small cost increase.

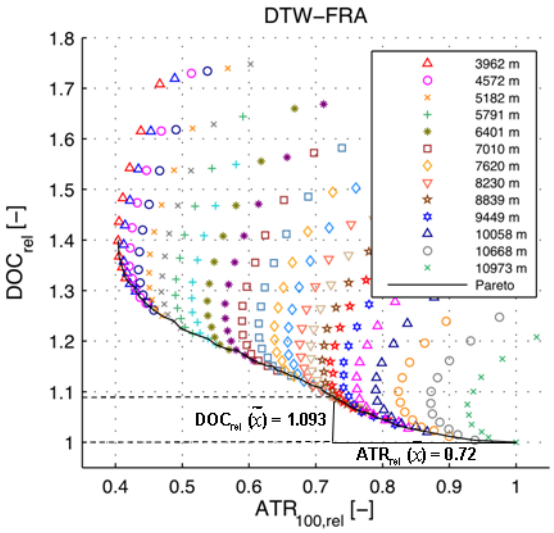


Fig. 8 Pareto front for route DTW-FRA obtained from all feasible operating points with resulting ATR_{rel} and DOC_{rel} values. Note that only every second ICA is displayed in the legend. ICA changes follow steps equivalent to 1000ft. Ma increases moving down the curves for constant ICA.

The DOC_{min} operating point and absolute ATR vary for each route, depending on the flight distance and feasible trajectories. Due to this, a different Pareto front and corresponding potential to reduce the climate impact is obtained for each route.

In order to identify the overall potential resulting from all routes operated by the reference aircraft in 2006, the approach described by Eq. 6 is applied. Figure 9 displays the $DOC_{rel,all}$ and $ATR_{rel,all}$ values based on all Pareto fronts. The error bars are calculated with Eq. 6 and indicate the percentile ranges for 25%-75% (blue) and 2.5%-97.5% (red), serving as measure of uncertainty for $ATR_{rel,all}(\tilde{x})$.

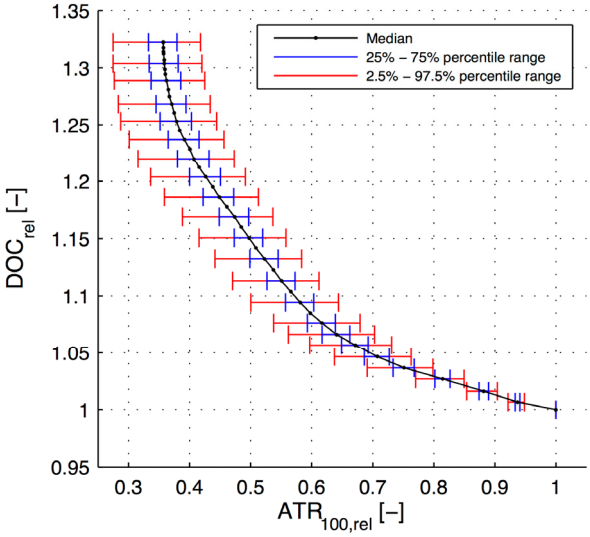


Fig. 9 $DOC_{rel,all}$ vs. $ATR_{rel,all}$ Pareto front for all routes operated with A330-200 in 2006. Error bars correspond to 25%-75% (blue) and 2.5%-97.5% (red) probability range.

The graph shows that the potential to reduce the climate impact by reduced cruise altitudes on all routes is still favourable, even better than on the route DTW-FRA. A maximum DOC increment of 9.5% relative to the DOC_{min} operating conditions on each route would yield a reduction of $ATR_{rel,all}$ between 36% and 50% (median 42%) when considering the 2,5%-97,5% percentile range.

The operating points corresponding to a given $DOC_{rel,all}(\tilde{x})$-$ATR_{rel,i}(\tilde{x})$ combination differ for each route, depending on the distance and payload. In order to infer from a given $DOC_{rel,all}(\tilde{x})$ back to the corresponding distribution of operating points, the Ma_{cr}-ICA

combinations at the route-specific $DOC_{rel,i}(\tilde{x})$ are plotted in a diagram Ma_{cr} vs. ICA for all route (Figures 10a-10b). The occurrence of operating points is expressed relative to the number of analyzed routes.

Figure 10a displays the distribution and relative occurrence of Ma_{cr}-ICA combinations with minimum DOC ($DOC_{rel,all}(\tilde{x})=1$) for all analyzed routes. As in Figures 6a-6b, the trend towards high Mach numbers and altitudes for DOC_{min} is clearly visible also in this graph. The majority of operating points are found at the highest possible Mach number 0.85 and ICA=12000m, while the rest is located in the near vicinity showing only a small variation.

Figure 10b shows the distribution of operating points for the exemplarily discussed $DOC_{rel,all}(\tilde{x})=1.095$. As described above, the decrease of ATR_{rel} and related increase of DOC_{rel} are connected to lower cruise altitudes and speeds. In the analyzed case the majority of operating points are located between $0.75 \leq Ma \leq 0.775$ and $8000m \leq ICA \leq 8500$, showing a larger variation than for DOC_{min}.

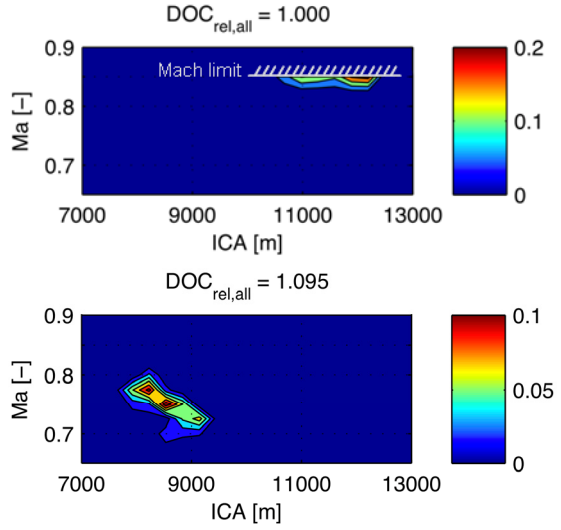


Fig. 10a-b Distribution and relative occurrence of all operating points for $DOC_{rel,all}$=1.00 (DOCmin) [a] and for $DOC_{rel,all}$=1.095 [b]

6 Conclusions

The present paper describes a comprehensive simulation and analysis approach developed at DLR to analyze different options to reduce the climate impact from aviation. The CATS simulation chain is applied to identify the potential to mitigate climate impact through reduced cruise flight altitudes on a global route network operated by a representative twin jet long-range aircraft in the year 2006. The study results are analyzed relative to the cruise operating conditions for minimum direct operating costs. The different impacts and trends on direct operating costs and average temperature response are discussed on the basis of a single route and transferred to the global route network. The distribution of relative occurrence of operating points for a given DOC increment is shown.

The present study found that the climate impact is mainly driven by altitude changes while the DOC are driven instead by speed changes. The study results further showed that there exists a large potential to reduce the climate impact from aviation for relative small to moderate increments on costs when the fleet of reference aircraft would be operated at lower cruise flight altitudes and speeds on all routes.

Future studies will focus on the potential that can be achieved when re-designing the reference aircraft for operating points with lower climate impact. Therefore a new design point will be derived from actual data. The reduction potential of the new aircraft will be discussed in comparison to the reference aircraft.

Acknowledgments

The following partner institutions contribute to the project *Climate compatible Air Transport System* (CATS) with expertise and models.

i. DLR - Atmospheric Physics
ii. DLR - Air Transportation Systems
iii. DLR - Propulsion Technology
iv. DLR - Combustion Technology
v. DLR - Flight Guidance
vi. DLR - Aerospace Medicine
vii. DLR - Simulation and Software Technology
viii Technical University Braunschweig - Aircraft Design and Lightweight Structures

Authors contact

Alexander Koch
German Aerospace Center (DLR)
Air Transportation Systems
Blohmstrasse 18
21079 Hamburg
Alexander.Koch@dlr.de
+49-40-42878-3634

References

[1] Lee, D.S., et al., *Aviation and global climate change in the 21st century*, Atmospheric Environment, 2009.

[2] Burkhardt, U., and Kärcher, B., *Global radiative forcing from contrail cirrus,* Nature Climate Change 1 (2011), 54–58, 2011.

[3] International Civil Aviation Organization – Committee on Aviation Environmental Protection, *FESG CAEP/8 Traffic and Fleet Forecasts*, CAEP-SG/20082-IP/02, 2008.

[4] Intergovernmental Panel on Climate Change, *Climate Change 2007 - Synthesis Report*, Cambridge University Press, 2007.

[5] Grewe, V., et al., *Climate impact of supersonic air traffic: an approach to optimize a potential future supersonic fleet – results from the EU project SCENIC*, Atmos. Chem. Phys., 7, 5129 5145, 2007.

[6] Dahlmann, K., at al., *Climate impact evaluation as part of aircraft pre-design*, 9th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA 2009-6957, 2009.

[7] Koch, A, et al., *Integrated analysis and design environment for a climate compatible air transport system*, 9th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA 2009-7050, 2009.

[8] Penner JE, Lister DH, Griggs DJ, Dokken DJ, McFarland M., *Aviation and the Global Atmosphere*, Cambridge University Press, 1999.

[9] Schumann, U., Graf, K., Mannstein, H., *Potential to reduce the climate impact of aviation by flight level changes*, 3rd AIAA Atmospheric Space Environments Conference, AIAA 2011-3376, 2011.

[10] Grewe, V., Stenke, A., *AirClim: an efficient climate impact assessment tool*, Atmospheric Chemistry and Physics, 8, 4621- 4639, 2008.

[11] Dahlmann, K., Bewertung neuer Flugzeugtechnologien hinsichtlich ihrer Klimawirkung, PhDthesis, Ludwig-Maximilians-Universität München, to be published.

[12] Kärcher, B., Möhler, O, DeMott, P.J., Pechtl, S., Yu, F. *Insights into the role of soot aerosols in cirrus cloud formation*. Atmos. Chem. Phys. 7 4203-4227, 2007.

[13] Appleman, H. *The formation of exhaust contrails by jet aircraft*, Bull. Am. Meteorol. Soc. 34, 14-20, 1953.

[14] Schumann, U., *On conditions for contrail formation from aircraft exhausts*, Meteorol. Z., 5, 4 –23, 1996.

[15] Fichter, C., Marquart, S., Sausen, R., Lee, D.S., *The impact of cruise altitude on contrails and related radiative forcing*, Meteorologische Zeitschrift, 14, S. 563 – 572, 2009a.

[16] Meerkötter, R. et al. *Radiative forcing by contrails*, Ann. Geophys. 17, 1080-1094, 1999.

[17] Wuebbles, D.J., Yang, H., Herman, R., Climate *Metrics and Aviation: Analysis of Current Understanding and Uncertainties,* Aviation-Climate Change Research Initiative, 2008

[18] Sausen, R., Schumann, U., *Estimates of the climate response to CO2 and NOx emission scenarios*, Climatic Change 44, 27–58, 2000.

[19] Schwartz, E., Kroo, I. M., *Metric for Comparing Lifetime Average Climate Impact of Aircraft*, AIAA Journal, Vol. 49-8, 2011.

[20] Böhnke, D., Nagel, B., Gollnick, V., *An Approach to Multi-Fidelity in Conceptual Aircraft Design in Distributed Design Environments*, IEEE Aerospace Conf., 2011.

[21] Liersch, C., Hepperle, M., *A Unified Approach for Multidisciplinary Aircraft Design*, CEAS European Air and Space Conference, 2009.

[22] Bachmann, A., Kunde, M., Litz, M., Schreiber, A., *Advances in Generalization and Decoupling of Software Parts in a Scientific Simulation Workflow System,* 4th International Conference on Advanced Engineering Computing and Applications in Sciences (ADVCOMP) , 2010.

[23] Heinze, W., Ein Beitrag zur quantitativen *Analyse der technischen und wirtschaftlichen Auslegungsgrenzen verschiedener Flugzeugkonzepte für den Transport großer Nutzlasten*, PhD-thesis TU Braunschweig, ZLR Forschungsbericht 94-01, 1994.

[24] Deidewig, F., *Ermittlung der Schadstoffemissionen im Unter- und Überschallflug*, PhD-thesis, 1998.

[25] Linke, F., *Trajectory Calculation Module*, DLR internal report, LK-IB 001-2008, 2008.

[26] Fichter, C., *Climate impact of air traffic emissions in dependency of the emission location and altitude*, Manchester Metropolitan University, 2009b.

[27] Liebeck, R. H., et al., *Advances in subsonic airplane design & economic studies*, NASA CR 195443, 1995.

[28] OAG Aviation Solutions, *Airline schedule and aircraft fleet databases,* 2006.

[29] Airbus S.A.S., *A330 Airplane Characteristics for Airport Planning,* 1993.

[30] ESG Aviation Services, *The Airline Monitor, Vol. 20 No. 6*, Edmund S. Greenslet, 2007.

[31] International Monetary Fund, *World Economic Outlook Database*, Sept. 2006, accessed 2011.

[32] Nakicenovic, N., et al., *Special Report on Emissions Scenarios: A special report of Working Group III of the Intergovernmental Panel on Climate Change*, 2000.

[33] Borken-Kleefeld, J., et al., *QUANTIFY transport emission scenarios up to 2100*, Environmental Science and Technology, 2010.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Emission-optimal flight trajectories with weather hazard avoidance and 4D wind modeling

Rafael Fernandes de Oliveira
EADS Innovation Works, Germany

Prof. Dr. Christof Büskens
Universität Bremen, Germany

Keywords: *flight trajectory optimization, optimal flight planning, emission abatement, weather-hazard avoidance*

Abstract

With the aim of reducing environmental impact caused by aviation, not only related to pollutant emissions, but noise annoyance on the population surrounding airports, and to reduce economic factors related to flying, being either fuel consumption or delay management, optimal control techniques have been investigated to obtain optimized trajectories that are able to minimize a given target or combination of them. Special care need to be taken when formulating the problem, as there are multiple limitations on how any given aircraft is able to fly, related to its own safety boundaries, ATM-imposed or operational restrictions. This research presents the application of such techniques in the development of a trajectory modeling tool that shall be capable of delivering near-real-time optimal flight profiles while still respecting any necessary constraint. It should also respond to any changes in weather conditions during flight, suggesting new paths that will avoid e.g. weather-related hazards. In order to illustrate the potential of the approach, a flight from Marseille to Munich is investigated, together with 4D winds and weather avoidance regions.

1 Introduction

Current flight planning and flight guidance procedures date back to the age of radio-navigation with VOR stations as waypoints. With the gradual deployment of GPS-assisted navigation, it became possible to define custom waypoints, providing greater freedom in the flight planning and improved departure and arrival procedures that reduced the aviation impact on the environment. Nevertheless, a high-fidelity trajectory that is optimal for a given goal is not yet feasible within this context.

Free-flight trajectories, such as studied by the Single European Sky Air Traffic Management Research (SESAR) can provide greater flexibility for the aircraft to define and follow its own optimal-trajectory without the need of following the current waypoint navigation approach. In either case, one approach for obtaining an optimal flight plan is the use of numerical trajectory optimization methods, which date back to the development of the digital computer and to the space exploration program. Its use have been successfully demonstrated in many flight planning problems, such as fuel consumption reduction, and emissions or noise-abatement, as seen in previous work on the subject by *Visser & Wijnen* [1-3], *Atkins & Xue* [4,5] and *Prats & Puig* [6], where novel emissions and noise-abatement procedures were investigated.

The current work introduces a tool being developed in the context of the CLEANSKY European research project that integrates trajectory optimization concerning economic and environmental goals, capable of flight

planning and replanning, and taking into consideration atmospheric conditions, such as wind and weather-related avoidance zones.

Overall, it translates the flight planning problem into an optimal control problem (OCP), an optimization method used to obtain control inputs for a given system that satisfies a set of constraints (physical, geometrical, dynamical) and at the same time minimizes (or maximizes) a certain optimality criterion. The OCP to be solved is usually described following the Continuous Bolza problem [7], where the Bolza cost-function to be optimized is defined as:

$$J(t, \boldsymbol{x}, \boldsymbol{u}) = \int_{t_0}^{t_f} L[\boldsymbol{x}(t), \boldsymbol{u}(t)]dt + \Phi[t_0, \boldsymbol{x}(t_0), t_f, \boldsymbol{x}(t_f)]$$

Here, $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u}(t) \in \mathbb{R}^m$ is the control vector, t_0 is the initial time, and t_f is the final time (both can be either fixed or free, according to the problem definition). This function is then optimized subject to a set of constraints described in Table 1. The OCP can then be solved numerically by transcribing it into a non-linear programing (NLP) problem. The main idea is to discretize time into multiple nodes, define a set of states, control inputs and constraints at each of these nodes, and then to interpolate the original continuous state and control functions. How this interpolation is obtained (linear, polynomial …) defines the class of optimal control method used. Here, a collocation method based on pseudospectral polynomials is used, as implemented by the MATLAB toolbox called GPOPS [8] (which stands for "General Pseudospectral Optimal Control Software"), that implements the Radau Pseudospectral method. For solving the resultant NLP, both the SNOPT [9] and WORHP [10] solvers are used.

The main target of the ongoing work is to investigate how a *near real-time* solution could be implemented on-board and could assist the pilot in taking the decision on how to plan the flight should be, and at the same time following all required ATM restrictions. It should be capable of providing faster and more refined solutions when compared to current planning approaches, and of optimizing and re-optimizing the current flight plan based on new sudden and unperceived events (like environmental condition changes or new ATM orders) by integrating in real-time information from external sources (weather radar, changes in aircraft dynamics …). Other expected usage scenarios are to help managing delays by providing time-of-arrival optimized trajectories, or to quickly provide landing trajectories to nearby airports in case of an emergency.

Table 1 – Set of constraints for an optimal control problem

Dynamic constraints	$\dot{\boldsymbol{x}}(t) \equiv \boldsymbol{f}[\boldsymbol{x}(t), \boldsymbol{u}(t), t], \quad t \in [t_0, t_f]$
Path constraints	$\boldsymbol{p}[\boldsymbol{x}(t), \boldsymbol{u}(t), t] \in [\boldsymbol{p}_{min} \; ; \; \boldsymbol{p}_{max}]$
Boundary conditions	$\boldsymbol{x}(t_0) \in [\boldsymbol{x}_{min,0} \; ; \; \boldsymbol{x}_{max,0}]$ $\boldsymbol{x}(t_f) \in [\boldsymbol{x}_{min,f} \; ; \; \boldsymbol{x}_{max,f}]$ $\phi[t_0, \boldsymbol{x}(t_0), t_f, \boldsymbol{x}(t_f)] = 0$
State and control boundaries	$\boldsymbol{x} \in [\boldsymbol{x}_{min} \; ; \; \boldsymbol{x}_{max}]$ $\boldsymbol{u} \in [\boldsymbol{u}_{min} \; ; \; \boldsymbol{u}_{max}]$

2 Modeling

2.1 Aircraft modeling

The aircraft is considered here to be a point-mass model, as it is not in the interest of this work to consider the full body dynamics (see for instance [11] for a detailed explanation). It can be described by the following state-space equations:

$$\dot{e} = V \cdot cos(\gamma) \cdot cos(\psi) + u_{wind}$$

$$\dot{n} = V \cdot cos(\gamma) \cdot sin(\psi) + v_{wind}$$

$$\dot{h} = V \cdot sin(\gamma)$$

$$\dot{V} = g \cdot \left[\frac{(Thrust - Drag)}{m \cdot g} - sin(\gamma)\right]$$

$$\dot{\psi} = \frac{g}{V} \cdot \left[N_Z \cdot \frac{sin(\phi)}{cos(\gamma)}\right]$$

$$\dot{\gamma} = \frac{g}{V} \cdot [N_Z \cdot cos(\phi) - cos(\gamma)]$$

$$\dot{\phi} = \dot{\phi}_C$$

$$\dot{m} = -FuelFlow(Thrust, V, h)$$

Where e, n and h represents the aircraft center of gravity coordinates in the local coordinate system, V the true airspeed, ψ the Yaw Angle, γ the flight path angle, ϕ the roll angle and m the aircraft mass. u_{wind} and v_{wind} are the local wind speeds, used to translate true airspeed into ground speed. Control inputs are: Load Factor N_Z, *Throttle* and Roll Angle Rate $\dot{\phi}_C$.

The Aircraft Performance Model (APM)is derived from the model introduced by EUROCONTROL in its *Base of Aircraft Data* (BADA) [12]. It consists of multiple aircraft-specific parameters describing aerodynamics and engine performance, identified from actual flight tests, and fitted to the underlining assumed APM. In the trajectories described later, the coefficients pertaining to the Airbus A320 are used.

Assuming that the aircraft is always capable of providing the required lift to fly and maneuver within the bound defined by load factor, speed and height limits, the aerodynamic drag can be obtained from the commanded Load Factor (N_Z) as in:

$$C_L = \frac{2 \cdot m \cdot g \cdot N_z}{\rho \cdot S \cdot V^2 \cdot \cos(\gamma) \cdot \cos(\phi)}$$

$$\mathcal{D} = \frac{1}{2} \cdot \left(C_{D_0} + C_{D_2} \cdot {C_L}^2\right) \cdot S \cdot \rho \cdot V^2$$

Here, S is the reference wing area, ρ is the air density, V is the true airspeed, C_{D_0} is the zero-lift drag coefficient and C_{D_2} is the lift-induced drag coefficient.

As for the total engine thrust, the APM specifies maximum thrust levels ($\mathcal{T}_{max}$) for take-off, climb and cruise, and together with the throttle control. Therefore, thrust is simply calculated as $Thrust = \mathcal{T}_{max} \cdot Throttle$. Finally, fuel consumption is defined for jet engines as a thrust specific fuel consumption factor η (given in $\frac{kg}{min \cdot kN}$), function of the true airspeed V in knots, is used to obtain the nominal fuel flow in $\frac{kg}{min}$, which is then multiplied by the total aircraft thrust $\mathcal{T}$ to obtain the instantaneous fuel consumption. A minimum fuel flow for when the engine is operating in idle is also defined.

$$\eta = C_{f1} \cdot \left(1 + \frac{V}{C_{f2}}\right) \qquad f_{min} = C_{f3} \cdot \left(1 - \frac{H_P}{C_{f4}}\right)$$

$$f_{nom} = max(\eta \cdot \mathcal{T}; f_{min})$$

2.2 Emissions Modeling

Aircrafts emissions occur directly in the upper troposphere, altering the concentration of atmospheric greenhouse gases and therefore contributing to global climate change. Carbon monoxide (CO), incompletely burnt hydrocarbons (HC) and water vapor (H_2O) all contribute directly to the global warming, as they have a positive radiative forcing [13]. Studies performed by the Intergovernmental Panel on Climate Change (IPCC) [14] have shown that aviation is responsible for between 2% (1992) and 2.5% (2005) of global CO_2 emissions, having mostly increased during the years (see Figure 1).

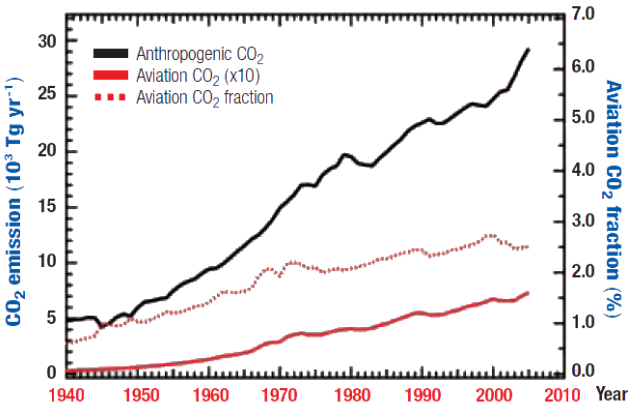


Figure 1 - History of aviation CO_2 emissions, from [13]

Although not too high, these emissions are concentrated on the higher atmosphere levels, where their impact can be higher. Also, just in the European Union, greenhouse gas emissions traced back to aviation were estimated to be 87% higher in 2006, when compared to the levels in 1990 [15].

In addition to the CO_2 released by burning fuel, there are also other emissions that locally affect air-quality and are thought to affect climate change [16]. Emissions of nitrogen oxides (NO_X) influence directly the ozone formation rate, contributing to deplete the radiation-protective ozone layer. Water vapor

and SO_X can trigger formation of condensation trails (*contrails*), which have a significant contribution to climate changes [13,17] and therefore will be included as a minimization goal in future developments of this tool.

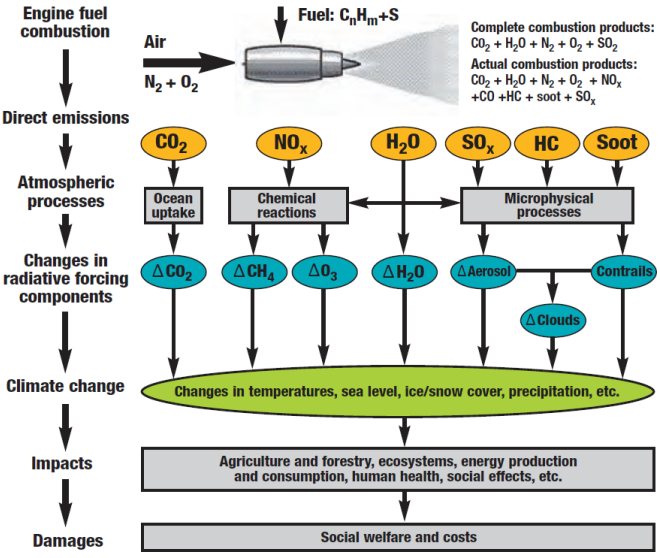


Figure 2 - ***Aircraft emissions and their causal linkages with potential climate changes, from*** [13]

Environmental emissions are modeled using the Advanced Emission Model 3 (AEM3) from EUROCONTROL [18]. As a modified version of the Boeing Method 2 (BM2), it estimates emissions by approximating the reference emission index of each pollutant regarding fuel-flow. Each of these indexes is defined as grams of pollutant per kilogram of burnt fuel (g/kg) at a reference sea-level International Standard Atmosphere (ISA), and interpolated from the data presented at the ICAO databank for engine rated thrusts of 7%, 30%, 85% and 100%. The values are further corrected for changes related both to pressure, temperature, pressure and humidity, and also deviations from ISA.

Emissions of H_2O, CO_2 and SO_X are estimated based on fixed emission indices from the amount of fuel burnt (1237, 3155 and 0.8 g per kg of burnt fuel, respectively).

2.3 Atmospheric and Weather Model

Previous studies by the FAA [19] have shown that up to 70% of all flight delays in the United States airspace are related to weather conditions. Therefore, it is useful to integrate weather avoidance into the trajectory optimization and provide optimal solutions that take into account not only environmental objectives and ATM constraints, but also avoids hazardous airspace zones. As it is not possible to control the weather, and safety must be ensured no matter what weather-related hazards, the ability of performing online trajectory replanning is of major interest to respond to any weather conditions changes that were not predicted before the flight took place. Another way the trajectory optimization can be used is to support improved airspace utilization (e.g. as described by [20-22]). Depending on how the weather conditions affect arrival or departure routes, they can become coupled. Therefore, it is not only useful to perform a single aircraft optimization process, but instead to optimize both routes accounting for the number of aircrafts that are going to fly them and for the minimum distance that is required between each.

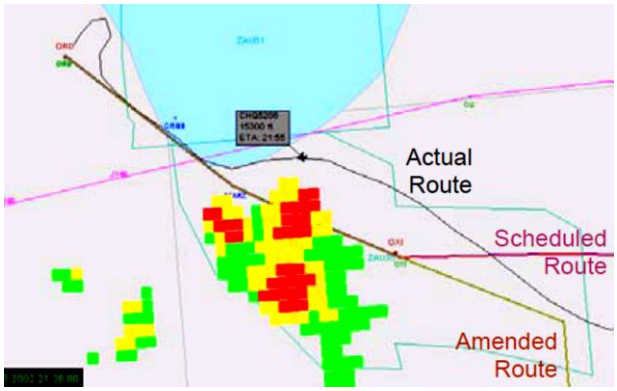


Figure 3 – Example of a weather cell affecting an arrival route, causing the aircraft pilot to request a deviation due to weather (Chicago O'Hare International Airport, 9/18/02) – from [20]

The standard atmospheric model is here augmented with data obtained from an improved version of the atmospheric modeling system SKIRON, developed by University of Athens [23]. It covers the European region, the Mediterranean Sea and north of Africa. Among many different outputs, it provides predicted wind patterns in multiple atmospheric layers, and expected cloud coverage, with hourly updates. Wind intensity is estimated during flight using multivariate interpolation (against latitude, longitude, pressure altitude and time).

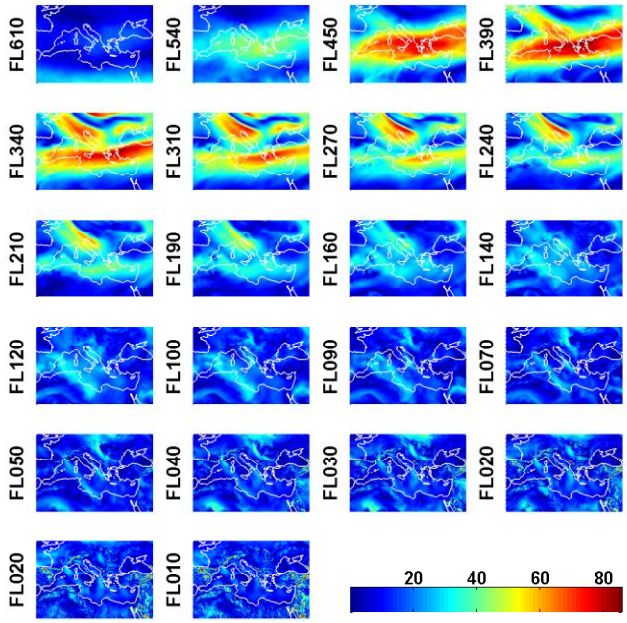


Figure 4 – Sample wind intensity map over Europe at multiple flight levels, data from July 2nd 2011 at 12am.

To illustrate how weather avoidance can be implemented, data from the total cloud coverage ratio (percentage of the sky hidden by all visible clouds.) was used. Although it does not represent directly the hazard level for an aircraft to fly through, usually defined by the precipitation level estimated by the weather radar in its cloud reflectivity map, it resembles how such a map might look like, and was used as representative enough for the studied scenario.

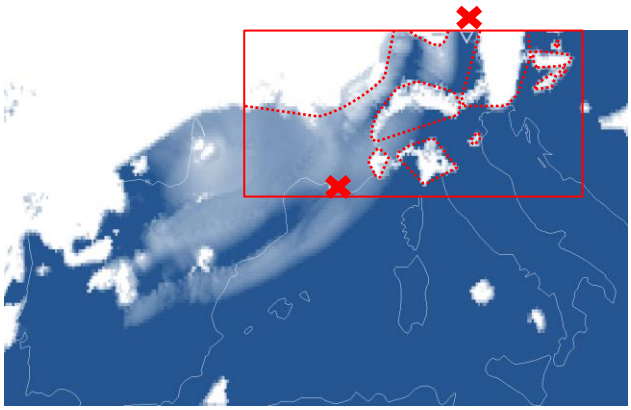


Figure 5 – Sample cloud coverage map over Europe (Marseille and Munich shown in red) and convex hull of no-fly zones

Cloud avoidance is implemented by thresholding the map into no-fly polygonal zones. Then, the signed distance between each node in the trajectory and the closest point in the polygon is required to bigger than a given desired value by adding a new path constraint to the OCP formulation.

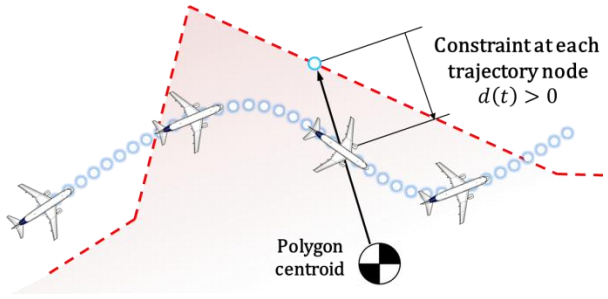


Figure 6 - Arbitrary-shaped polygonal no-fly regions

2.4 Modeling requirements

While translating the optimization problem into an OCP, special care needs to be taken in the implementation of each model. Forward-mode automatic differentiation is used, which also introduces new special requirements – see [24]. Also, in order to facilitate convergence, functions should be at least one time continuous differentiable (second derivatives are approximated with a BFGS method), so that special care has to be taken with stepwise changes This can be approximated with the use of an arc-tangent function, as in Table 2. The parameter β controls how smooth the transition is, with smaller values producing a smother function at the cost of lower fidelity.

Table 2 – Smooth step-function definition

	Original	Smoothed function
$f(x) =$	$\begin{cases} a, x \le k \\ b, x > k \end{cases}$	$\left\{\text{atan}[\beta \cdot (x-k)] + \frac{\pi}{2}\right\} \cdot \frac{b-a}{\pi} + a$

As an example, the maximum true airspeed defined for the Airbus A320 has a step change at FL250. Under this flight level, the aircraft is limited to fly at 250 KIAS or less, while above it, it is constrained by a maximum Mach number of 0.82. Using the above formulation with $\beta = 0.01$, one can obtain the approximated function shown in Figure 7.

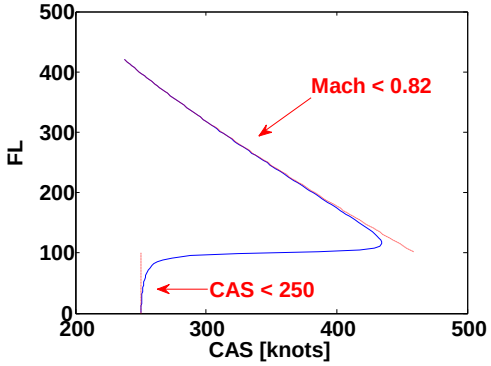


Figure 7 – A320 true airspeed limits approximated as a continuous function

3 Optimization Criteria

Fuel Consumption

As fuel consumption is already calculated from the aircraft performance model, the optimization function is simply defined as

$$J_{Fuel}(t, \boldsymbol{x}, \boldsymbol{u}) = \int_{t_0}^{t_f} -\dot{m}[\boldsymbol{x}(t), \boldsymbol{u}(t)]\, dt = m(t_0) - m(t_f)$$

Flight Time

Flight time is optimized by the following function

$$J_{Time}(t) = t_f - t_0$$

Emissions

Gaseous emissions are optimized by integrating their generation rate (defined by the Emissions Index explained before) over the whole flight, as in

$$J_{Emissions}(t, \boldsymbol{x}, \boldsymbol{u}) = \int_{t_0}^{t_f} EI(t) \cdot \dot{f}_{engine}[\boldsymbol{x}(t), \boldsymbol{u}(t)]\, dt$$

3.1 Pareto optimization

To evaluate the tradeoff between time or emission-optimal trajectories, the concept of Pareto efficiency is used. Due to the single-objective nature of the OCP formulation, a weighted sum of multiple objectives is necessary to find the Pareto front, with k_t being changed to reflect a lower or higher focus on optimizing emissions, as in:

$$J(t, \boldsymbol{x}, \boldsymbol{u}) = J_{Emissions}(t, \boldsymbol{x}, \boldsymbol{u}) + k_t \cdot J_{Time}(t)$$

4 Scenario Formulation

4.1 Flight Plan

A flight plan is modeled in the OCP by transcribing each pair of waypoints into a different phase with a local coordinate system. Constraints linking the final state of each phase with the initial state of the following are added by converting both start and end coordinates back to a Latitude/Longitude coordinate system.

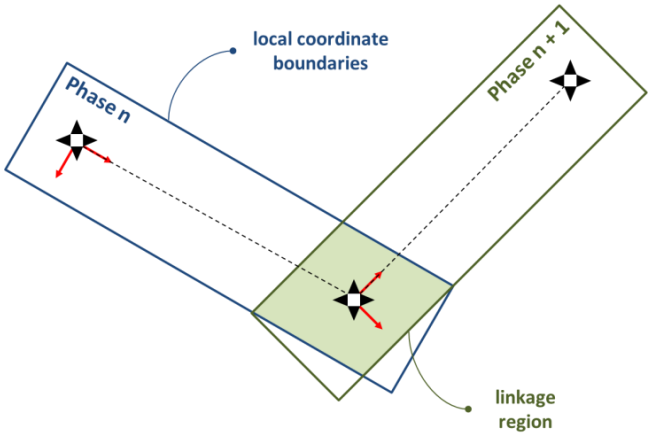


Figure 8 - Flight plan modeling in multiple phases

4.2 Constraints

Different requirements, such as aircraft operational limits, trajectory restrictions, weather conditions can lead to constraints in how the aircraft should fly. Following the OCP modeling, they can be implemented either by limiting state bounds – e.g. roll rate less than $15\,°/s$ – or as path constraints, function of the aircraft states and controls.

4.2.1 Flight Envelope

The following flight envelope constraints are defined together with the aircraft performance model, restricting maximum and minimum speeds and altitude ceiling attainable by the aircraft while still maintaining a required minimum rate of climb. The aircraft maximum speed is constrained by both the maximum operational calibrated airspeed (V_{MO}) and

maximum operational Mach number (M_{MO}), while the minimum speed is defined as a function of the aircraft stall speed, defined for different aircraft configurations and altitude boundaries.

$$V_{min} = \begin{cases} C_{Vmin,TO} \cdot V_{stall}, & if\ in\ takeoff\ cfg. \\ C_{Vmin} \cdot V_{stall}, & otherwise \end{cases}$$

The maximum altitude for any given aircraft mass is defined by:

$$(h_{max})_{actual} = \min \begin{pmatrix} h_{MO} \\ h_{max} + G_t \cdot \max(0, \Delta T - C_{Tc,4}) + \cdots \\ \cdots\ G_W \cdot (m_{mto} - m) \end{pmatrix}$$

Where h_{MO} is the maximum operational altitude, h_{max} is the maximum altitude at MTOW and ISA conditions (defined by a required residual rate of climb of $300\,\frac{\text{ft}}{\text{min}}$), G_t is the temperature gradient at h_{max}, ΔT is the deviation in temperature from an ISA atmosphere, G_W is the mass gradient at h_{max}, m is the actual aircraft mass and m_{mto} is the maximum take-off mass.

4.2.2 Low Speed Buffeting Limit

Buffeting is associated with high-frequency instability, usually present at high altitudes and low Mach, caused by airflow separation or shock wave oscillations. A second speed limit is set when an aircraft is flying above FL150, by limiting the Mach to the lowest positive solution of the following cubic function:

$$k \cdot M^3 - C_{Lbo(M=0)} \cdot M^2 + \frac{m \cdot g}{S \cdot p \cdot 0.583} = 0$$

Where k is the lift coefficient gradient, $C_{Lbo(M=0)}$ is the initial buffeting onset lift coefficient at Mach = 0, S is the aircraft wing reference area, m is the actual aircraft mass and p is the air pressure.

4.2.3 Operational restrictions

A subset of operational constraints, identified for the Airbus A320 aircraft (from [25-28]) is currently taken into consideration.

The aircraft roll angle should not exceed the boundaries defined by the aircraft manual and the ICAO procedures, and the roll rate is limited to $|\dot{\phi}| \leq 15\,^{\circ}/_{s}$. The aircraft airspeed is constrained to 250 KIAS when flying below FL100 and to Mach 0.82 when above (see Figure 7). Aircraft rate of climb ($\dot{h}$) is limited according to the current flight level as

$$\dot{h} \leq \begin{cases} 2400\,^{ft}/_{min}, & below\ FL200 \\ 1800\,^{ft}/_{min}, & between\ FL200\ and\ FL300 \\ 1200\,^{ft}/_{min}, & above\ FL300 \end{cases}$$

5 Numerical Results

The scenario demonstrated here describes a commercial flight from Marseille to Munich (scenario formulation from [29-31]). The A320 take-off from LFML runway 13, climbs up to 300ft, then crosses waypoint TRETS, climbs until FL120 at 6.5%, and then climbs to FL300 and starts the cruise phase, through waypoints NASIK, RAPED; GEN, OSKOK, DIBAX, TUTAL and IRBIR (the full list was simplified to remove unnecessary waypoints that are in a straight line on the trajectory).

Then, on the arrival phase, the aircraft crosses waypoint AMGAS under FL130, and then under FL110 through ANDEC and BETOS. From there on, it follows the standard terminal arrival charts (STAR), flying under FL080 through MUN, NAPSA, DM448 and NELBI. Finally, the aircraft descends at -3º and lands at runway 26L.

5.1 Optimal departure from LFML

Results for the departure from the Marseille Airport are shown in Figure 10. As the procedure is not so complex, the main difference found was when optimizing against NO_X emissions. The aircraft climbs slower, and has a later acceleration phase, obtaining a reduction of about 10% in NO_X. CO and HC optimal trajectories are identical, climbing as fast as allowed, and then using the available thrust to accelerate to the final speed – resulting in a reduction of about 28% and 17% each pollutant emission, respectively.

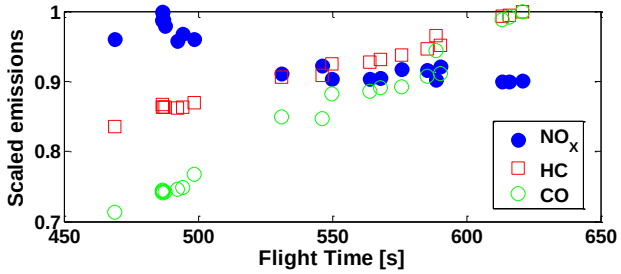


Figure 9 – Pareto front for emissions-optimal departure trajectories, with $k_t \in [10^{-5}; 10^{-3}]$

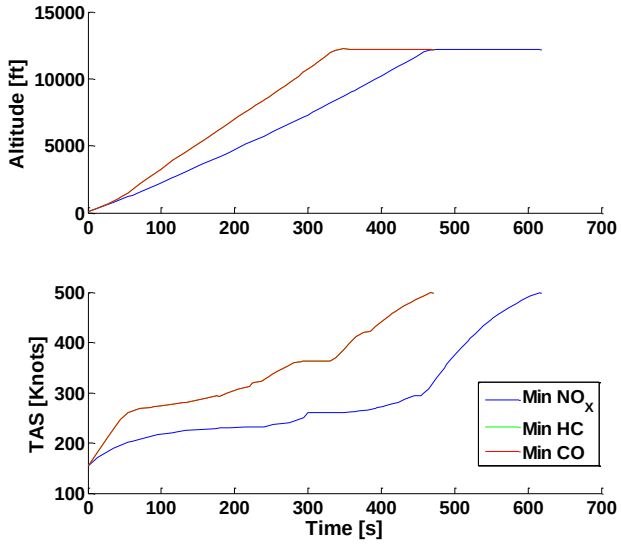


Figure 10 - Departure emission-optimal trajectories

5.2 Optimal arrival at EDDM

During the arrival phase, different trajectories were found for each pollutant. NO_X has a very slow descent profile that reduces earlier the airspeed, and then gradually descents until destination, having a 75% lower emission than other arrival trajectories. The HC-optimal trajectory, instead, descent and decelerate until it reaches FL065, and then maintain a constant airspeed of about 295 knots until it arrives to the NELBI waypoint, decelerating further to the required landing speed, which results in a reduction of about 25% in incompletely burnt hydrocarbons. Finally, the CO optimized trajectory has a later-descent profile, flying at FL100 for half of the arrival, and then further descending and decelerating to reach the runway. Doing so, it reaches a reduction of about 55% in the amount of CO produced during fuel combustion.

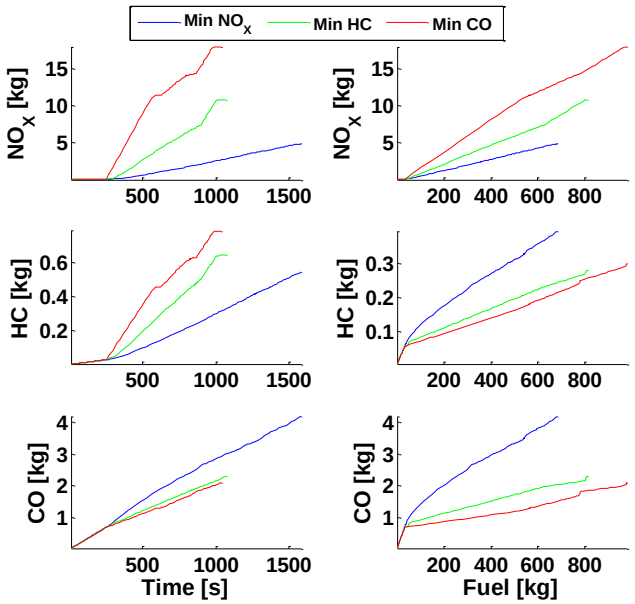


Figure 11 - Emissions history against time and fuel consumption

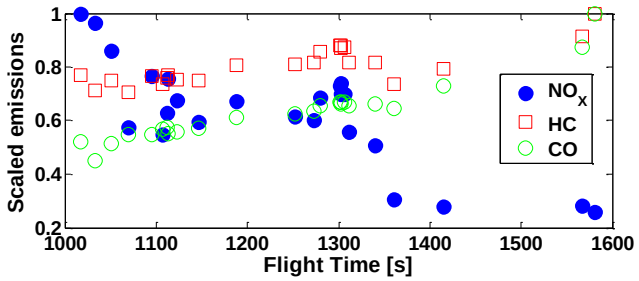


Figure 12 – Pareto front for emissions-optimal arrival trajectories, with $k_t \in [10^{-5}; 10^{-3}]$

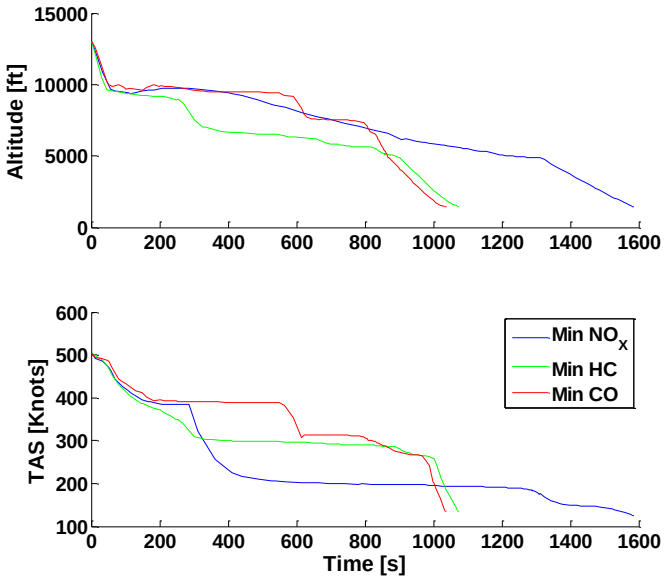


Figure 13 – Vertical flight profile and TAS history for arrival emission-optimal trajectories

5.3 Cruising through bad weather

Cloud avoidance is shown here with the full flight plan, departing from LFML and landing at EDDM. After converting the cloud information into no-fly zones, each flight phase in the OCP is augmented with a new path constraint for each zone. In the cruise phase, conflicting waypoints in the bad-weather region were also removed, so that the aircraft had the freedom to fly around the clouds.

Original and new trajectories are shown in Figure 14, where it can be seen that the aircraft successfully avoids the regions and converges to the original flight plan at waypoint TUTAL: Cloud data is cropped at Latitude 48º, due to availability of the weather model.

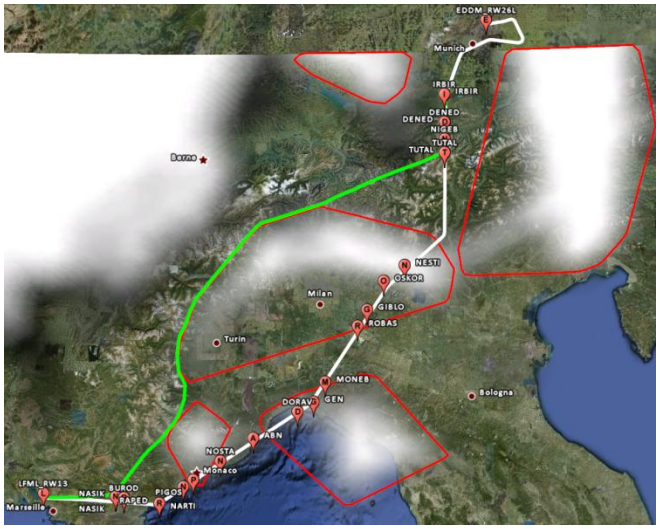


Figure 14 – Avoidance of weather-related no-fly zones in the full flight between Marseille and Munich. Original flight plan shown in white, and replaned trajectory in green.

5.4 Wind influence on optimal trajectories

The influence of wind in the aircraft trajectory can be seen in Figure 15 (departure from LFML) and Figure 16 (arrival at EDDM), in both cases showing a time-optimal trajectory without the influence of wind (in black) and with (in green). It can be seen that a significant deviation from both original ground tracks results from winds of up to 80 m/s during flight.

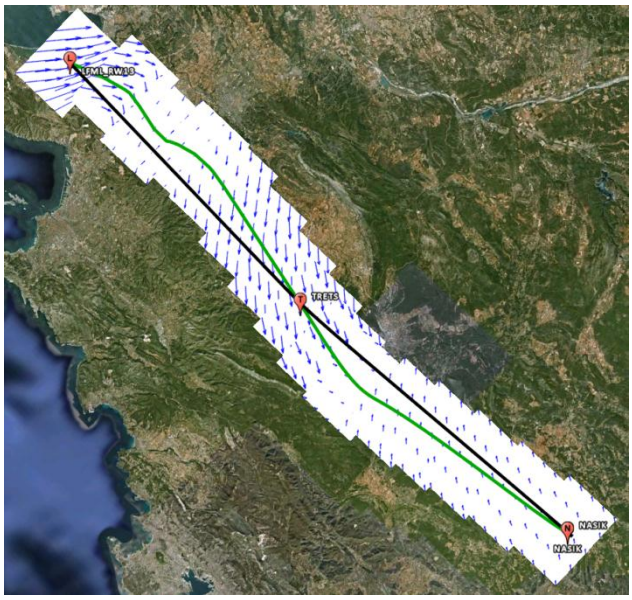


Figure 15 – Original departure (shown in black) and wind-optimal trajectory (in green). Wind vectors are overlaid in 5km intervals, showing how wind is evolving with time and altitude.

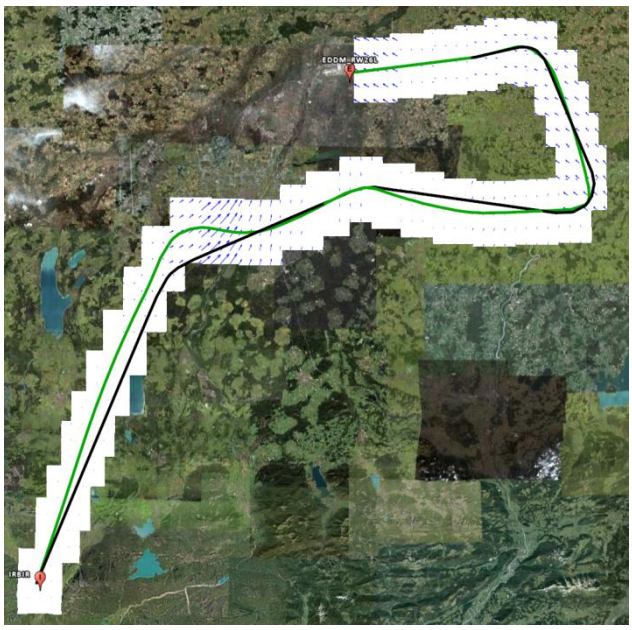


Figure 16 – Original arrival (shown in black) and wind-optimal trajectory (in green). Wind vectors are overlaid in 5km intervals, showing how wind is evolving with time and altitude.

Concerning emissions, the influence of wind can be seen when comparing the new flight profiles and Pareto fronts for departure and arrival. New trajectories are considerably faster, due to favorable wind, but while departure had similar optimal emissions, arrival trajectories were found to be worse for NO_X *(+14.3%)* and CO (*+15.2%*) emissions, and better for incompletely burnt hydrocarbons *(-6.6%)*.

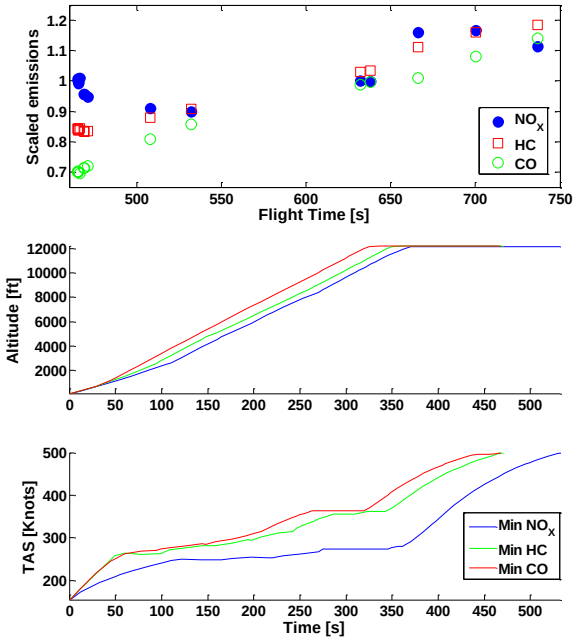


Figure 17 - Emissions-optimal departure trajectories with presence of wind (Pareto front normalized against no-wind trajectories)

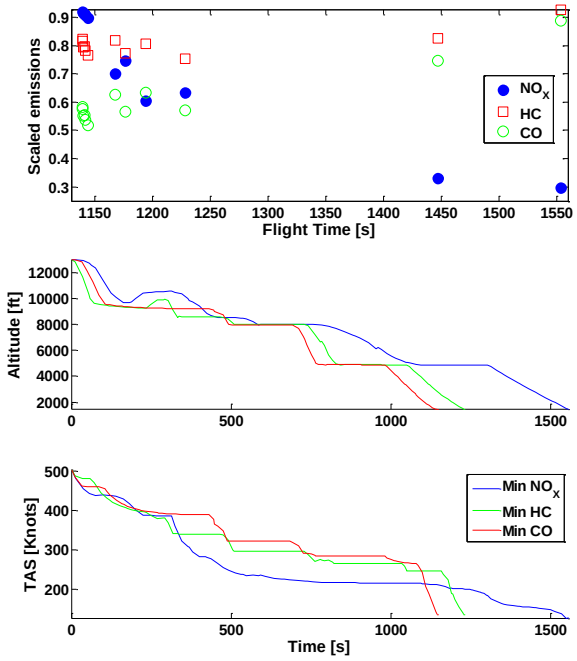


Figure 18 - Emissions-optimal arrival trajectories with presence of wind (Pareto front normalized against no-wind trajectories)

Flight profiles are also significantly different for the arrival. For example, the original NO_X-optimal trajectory had a gradual descent from FL100 to destination, while the new one uses a small climb to reduce airspeed, and then descents in a three-step profile, gradually reducing decelerating to landing speed. Meanwhile, HC and CO optimal trajectories have now a similar vertical profile following also three-step descent, but with the final deceleration before landing occurring about 200 seconds later.

6 Conclusion

The research presented here has shown how weather effects can affect optimal flight trajectories when emissions are taken into account. The approach not only includes the modeling of a given flight plan with reasonable accuracy, but also identified restrictions imposed on the aircraft operation from the aircraft manual and aviation authorities. In the flight between Marseille and Munich, Pareto fronts for optimal departure and arrival trajectories were obtained for different pollutant emissions, showing how RNAV-based standard instrument departures (SID) and standard terminal arrival routes (STAR) can be improved according to the used criteria. Concerning the future near-real-time goal of the tool, optimal trajectories are already obtained within minutes, and future improvements are expected with the introduction of better heuristics for the initial OCP guess, and also with model simplifications and improvements in the NLP solver

References

[1] H.G. Visser and R.A.A. Wijnen, "Optimisation of noise abatement arrival trajectories," *The Aeronautical Journal*, 2003, pp. 607-615.

[2] H.G. Visser and R.A.A. Wijnen, "Optimization of Noise Abatement Departure Trajectories," *Journal of Aircraft*, vol. 38, 2001.

[3] R.A.A. Wijnen and H.G. Visser, "Optimal departure trajectories with respect to sleep disturbance," *Aerospace Science and Technology*, vol. 7, 2003, pp. 81-91.

[4] E.M. Atkins and M. Xue, "Noise-Sensitive Final Approach Trajectory Optimization," *Journal of Aerospace Computing, Information and Communication*, vol. 1, 2004.

[5] M. Xue and E.M. Atkins, "Three-Dimensional Segmented Trajectory Optimization for Runway-Independent Aircraft," *Aerospace Engineering*, 2005, pp. 1-15.

[6] X. Prats, V. Puig, and J. Quevedo, "Optimisation of site-specific noise abatement departure procedures,"

9th Innovative Research Workshop & Exhibition, 2010.
[7] G.T. Huntington, "Advancement and Analysis of a Gauss Pseudospectral Transcription for Optimal Control Problems," Massachusetts Institute of Technology, 2007.
[8] A.V. Rao, C.L. Darby, and M. Patterson, *User's Manual for GPOPS Version 3.3: A MATLAB Software for Solving Optimal Control Problems Using Pseudospectral Methods*, 2010.
[9] P.E. Gill, W. Murray, and M.A. Saunders, "SNOPT : An SQP Algorithm for Large-Scale Constrained Optimization," *Science*, vol. 47, 2005, pp. 99-131.
[10] C. Büskens, T. Nikolayzik, and M. Gerdts, *Nonlinear large-scale Optimization with WORHP*, Bremen: 2010.
[11] D.G. Hull, *Fundamentals of Airplane Flight Mechanics*.
[12] A. Nuic, *User Manual for the Base of Aircraft Data(BADA) Revision 3.8*, 2010.
[13] *Aviation's Contribution to Climate Change*, 2007.
[14] J.E. Penner, D.H. Lister, D.J. Griggs, D.J. Dokken, and M. McFarland, "Aviation and the Global Atmosphere," 1999.
[15] *Climate change : Commission proposes bringing air transport into EU Emissions Trading Scheme*, 2006.
[16] R. Sausen and U. Schumann, "Estimates of the climate response to aircraft CO2 and NOx emissions scenarios," 2000.
[17] U. Burkhardt and B. Kärcher, "Global radiative forcing from contrail cirrus," *Nature Climate Change*, vol. 1, 2011, pp. 54-58.
[18] "AEM - Advanced Emission Model - EUROCONTROL Experimental Centre."
[19] *Weather Forecasting Accuracy for FAA Traffic Flow Management*, Washington, DC: 2003.
[20] J. Krozel, B. Capozzi, A.D. Andre, and P. Smith, "The future National Airspace System: Design requirements imposed by weather constraints," *Architecture*, 2003, pp. 1-14.
[21] "Traffic Flow Management (TFM) Weather Rerouting Decision Support," *Atmospheric Research*, 2000.
[22] G. Davidson, J. Krozel, S.M. Green, C.K. Mueller, A.T. Control, A. Route, T. Control, A.T. Management, C. Convective, F. Product, C.D. Making, S. Units, F.C. Area, G.D. Program, G. Stop, N.W. Service, S. Weather, A. Program, T.F. Management, V.I. Liquid, S. Analyst, D. Division, C. Scientist, and D. Division, "Strategic Traffic Flow Management Concept of Operations," 2004, pp. 1-10.
[23] C. Spyrou, C. Mitsakou, G. Kallos, P. Louka, and G. Vlastou, "An improved limited area model for describing the dust cycle in the atmosphere," *JOURNAL OF GEOPHYSICAL RESEARCH*, 2010.
[24] Y. Chen, J. Gu, J. Liang, and R. Fullmer, "RIOTS + AD : Integrating Automatic Differentiation into Computational Optimal Control."
[25] *Aircraft Operations - Volume II - Construction of Visual and Instrument Flight Procedures*, 2006.
[26] *Aircraft Operations - Volume I - Flight Procedures*, 2006.
[27] W. Schott and C. Miley, *Pilot Resource Document - Airbus A320 Checklist*, 2007.
[28] *Airbus A320 Pilot's Operating Handbook and Checklists*.
[29] *Airport charts for EDDM - Franz Josef Strauß Munich Airport*, Munich: 2010.
[30] *Departure charts for LFML - Marseille Provence Airport*, Marseille: .
[31] "VatRoute.net - LFML to EDDM route."

3rd CEAS Air&Space Conference
21st AIDAA Congress

Multi-Objective Optimisation of a constrained 2000 km Trajectory using Genetic Algorithms

Kenneth Chircop
Department of Electronic Systems Engineering, University of Malta, Malta

William Camilleri, Vishal Sethi
School of Engineering, Cranfield University, UK

David Zammit-Mangion
Department of Electronic Systems Engineering, University of Malta, Malta
School of Engineering, Cranfield University, UK

Keywords: *trajectory optimization, aircraft performance, NOx emissions*

Abstract

This paper presents a study whereby a typical 2000 km trajectory flown by an aircraft of the Airbus A320 and Boeing 737 family, is imposed operational constraints such as altitude restrictions and operational speeds, and optimised for the reduction of fuel, nitrous oxides (NOx) and time.

The trajectory is divided into two phases, the first encompasses the initial climb following take-off up to 3,000ft, and it is optimised for NOx and fuel, while the second encompasses the en-route climb starting from 6,000ft, the cruise phase, and descent to 8,000ft and it is optimised for fuel and time. The resulting pareto frontiers and the corresponding extremal trajectories are analysed and discussed from an aircraft performance point of view, from which salient conclusions are drawn on the optimal trajectories generated.

Nomenclature

AGL	=	Above Ground Level
ATM	=	Air Traffic Management
CAS	=	Calibrated Airspeed
FL	=	Flight Level
NOx	=	oxides of nitrogen
TAS	=	True Air Speed

1 Introduction

Air transport currently depends entirely on fossil fuels and mainly on gas turbine technology for propulsive means. This implies that every flight contributes to the usage of a finite resource, to the emission of green-house gases and secondary emissions (mainly NOx) and also to noise pollution.

The study in this paper addresses the optimisation of a 2000 km trajectory bound by operational constraints in altitude and speed. Flights of this sector length are common for 150-180 seat category air transports such as the Boeing 737 and Airbus A320 family of aircraft and can be exemplified by routes such as Paris-Athens, Dubai-Islamabad and Los Angeles-Dallas. As part of the problem definition, the altitude throughout the trajectory was constrained to a maximum value of 34,000 feet, which may not be typical of real flights since aircraft are normally capable and allowed to reach altitudes of 36,000 feet and higher. The minimum cruise altitude was selected at 20,000ft to facilitate the optimisation set-up. The aircraft operational speed was constrained (limited) in terms of calibrated air speed below 20,000ft and Mach above. Constraining the cruise altitude to above 20,000ft facilitated the problem setup and avoided potential complications that could result from multiple

transitions about this boundary. Consequently, this constraint effectively introduced pseudo-ATM and operational constraints to the problem. Although non-representative of real flights since ATM constraints are normally imposed on relatively short segments of the trajectory, it is, nonetheless, useful in this initial study on a generic trajectory.

A genetic algorithm was used to optimise the trajectory as this class of algorithms is not prone to converge on a local minimum. This is a significant capability in the application of trajectory optimisation with discrete trajectory constraints, where other types of optimisation techniques may be unable to perform the optimisation in a satisfactory or reliable manner. Genetic algorithms, however, require large computational resources, which normally translate to slower processing. This is particularly due to the fact that genetic algorithms are heuristic in nature, requiring the algorithm to cover a relatively large section of the problem space compared to classical techniques. However, genetic algorithms are less sensitive to the initial conditions, although they require careful setup in terms of, amongst others, problem space definition, population size and crossover techniques if they are to converge to a meaningful result.

2 The Problem Setup

The overall trajectory was divided into two separate stages to facilitate different multi-objective assessment. The first stage encompasses the initial climb following take-off, initiating from a screen height of 50 ft at the runway threshold up to 3,000 ft. This trajectory is optimised for NOx and fuel, and it extends only up to 3,000 ft because this is generally the altitude up to which NOx emissions are considered. Normally, this phase of flight involves several configuration changes, allowing the transition from the take-off to the clean configurations. Configuration changes were not incorporated in the assessment in order to limit the complexity of the problem setup and the optimisation time of the genetic algorithm. Whilst this did not afford the analysis of the impact the optimisation criteria have on strategies associated with configuration changes, it provided a preliminary understanding of the impact NOx considerations may be expected to have on the trajectory.

The second stage encompasses the en-route climb starting from 6,000ft, the cruise phase and descent to 8,000 ft. In this case, fuel-optimal trajectories are generated for an aircraft weighing 60 tonnes. The resulting trajectories are compared, analyzed and a number of conclusions are drawn.

The optimisations were performed using a multi-objective optimisation algorithm and a number of relevant models, which included an aircraft performance model, an engine model, and an emissions model. A standard atmosphere was used and still air assumed in all optimisations. The aircraft performance model is representative of a generic, single aisle, twin-engine jet airliner of the Airbus A320 / Boeing 737 category. The model determines the aerodynamic state of the aircraft based upon a reduced set of aerodynamic derivatives as well as certain geometric characteristics typical to the class of aircraft being replicated. The model outputs the thrust requirement and flight time for the aircraft when required to travel between initial and final pre-defined states, which are the inputs to the model. The initial and final states are defined in any of a number of ways and include parameters such as altitude, range and velocity vectors. Due to the complexity of the model, which also takes into account factors such as the effects of variations of flight path angle and atmospheric conditions between the initial and final states, the solution is obtained iteratively and is based upon an integration routine.

The engine model is representative of a two spool, high bypass, turbofan engine of the 12 tonne static thrust class such as the CFM56 series engine typically found on A320 and Boeing 737 family aircraft. This model computes the steady-state performance of the engine based upon design-point data, generic component characteristic maps and fundamental thermodynamic relationships. The thrust requirement as well as altitude and Mach

number are provided as inputs to the model and the solution is obtained through a number of iterations. The suite of data output by the model includes specific fuel consumption, turbine entry temperature, air mass flow rate and fuel flow rate. The aircraft performance model interfaces with the engine model through the provision of thrust requirement data while fuel flow rate obtained from the engine model is used to update the aircraft weight.

The emissions model is used to determine the fuel burn-related emissions generated by the afore-mentioned engine type. The model is based upon empirical data, reactor models as well as widely reported chemical reaction mechanisms. The model determines the emission indices for Oxides of Nitrogen, Carbon Dioxide and water vapour. The latter two emission types are determined through basic chemical equilibrium while more complex rate chemistry is employed to model oxides of Nitrogen. In order to determine these indices, knowledge of atmospheric conditions and combustor fluid characteristics is required. The former is provided by the atmospheric model, whilst the latter is provided by the engine model.

The model setup, whilst not being numerically precise in terms of specific aircraft performance, afforded the observation of optimal flight trajectory strategies for aircraft typically of the size and configuration referred to in this paper.

2.1 The First Stage – Initial Climb

The initial climb has been defined by 3 segments as shown in Table 1, starting at 50ft up to an altitude of 3000ft. A constant configuration setting (Slats + Flap2) was assumed with gear up. Initial conditions set were:

- weight 60.5 Tonnes
- altitude 50ft above ground level (agl)
- speed 165kts CAS.

Segment	Final Velocity range (CAS)	Altitude range (ft)
1	165 kts	50 - 3000
2	165-250 kts	50 - 3000
3	165-250 kts	3000

Table 1: First stage segment constraints.

2.2 The Second Stage – Climb, Cruise and Descent

The climb, cruise and descent trajectory was defined by 10 segments as shown in Table 2. The initial conditions, not shown in the table, were:

- weight 60 Tonnes
- altitude FL60
- speed 250kts CAS.

Segment	Segment Length	Final Velocity range	Altitude range (FL)
1	63km	270-370kts CAS	100 - 200
2	254km	270-350kts CAS	190 - 290
3	327km	0.75 – 0.82 Mach	200 - 340
4	219km	0.75 – 0.82 Mach	200 - 340
5	123km	0.75 – 0.82 Mach	200 - 340
6	301km	0.75 – 0.82 Mach	200 - 340
7	260km	0.75 – 0.82 Mach	200 - 340
8	241km	0.75 – 0.82 Mach	200 - 290
9	132km	270-370kts CAS	100 - 200
10	14km	256kts CAS	80

Table 2: Second stage segment constraints.

The defined problems were designed such that an aircraft the size and performance of that represented in this paper will have the performance to fly within these boundaries.

3 Results

3.1 The First Stage – Initial Climb

The pareto frontier of the two-objective (fuel-NOx) optimisation is presented in Figure 1. The frontier exhibits moderate smoothness, indicating a good trend line of where the actual boundary can be expected to be. NOx formation in gas turbines is accelerated by higher operating temperatures associated with higher thrust settings (and therefore higher fuel burn) during take-off and initial climb. However, higher thrust settings afford greater efficiency and a quicker climb to 3000ft agl. Consequently, it could be expected that higher thrust levels would be more advantageous in terms of optimising (minimising) fuel burn, whilst for the limitation of NOx emissions, reduced thrust and associated lower engine operating temperatures would be more advantageous.

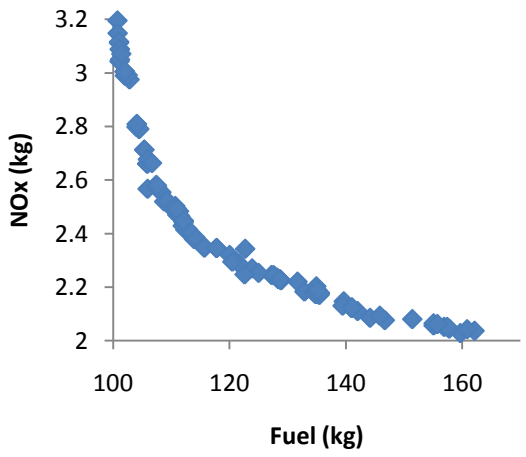


Figure 1: The Pareto Frontier for NOx and fuel burn for the first stage

Figures 2, 3 and 4 present the profiles of the two extreme operating conditions, namely those minimising NOx and fuel burn respectively. As can be observed, the strategy for minimum fuel burn to 3000ft requires a more rapid climb than that for minimum NOx emission, and this is achieved through the application of higher thrust and greater fuel flow (burn) rate, attaining 3000ft within 5km as compared to just under 15km for the minimum NOx trajectory. The rate of climb in the third segment is approximately 2800ft/min for the minimum fuel burn case and 800ft/min for the minimum NOx trajectory. The latter strategy, which exhibits a very low climb rate, is currently not normally adopted operationally as aircraft tend to maintain high rates of climb in the initial climb. It is worthwhile noting, however, that the optimisation strategy did not take into account factors such as noise abatement procedures, engine out performance limitations and obstacle clearance.

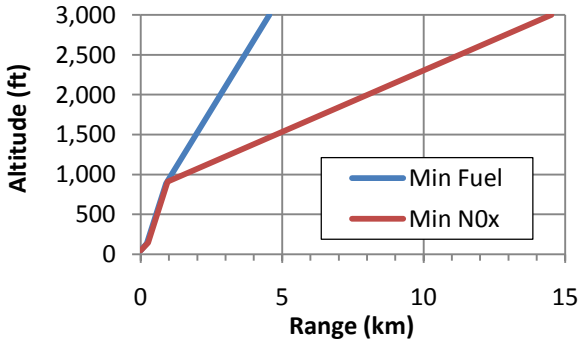
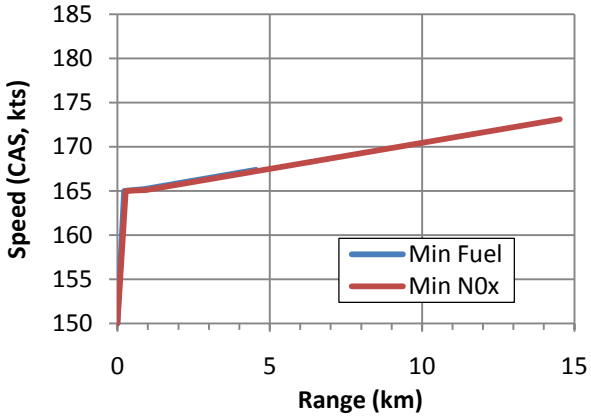
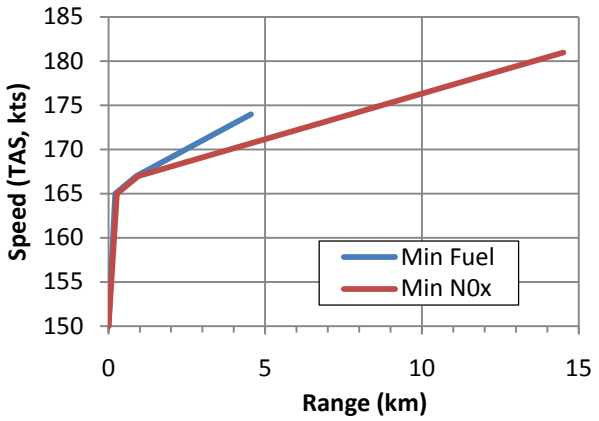


Figure 2: Trajectory profiles for minimum fuel burn and minimum NOx emissions during initial climb



(a)



(b)

Figure 3: Speed profiles for minimum fuel burn and minimum NOx emissions during initial climb (a) CAS, (b) TAS

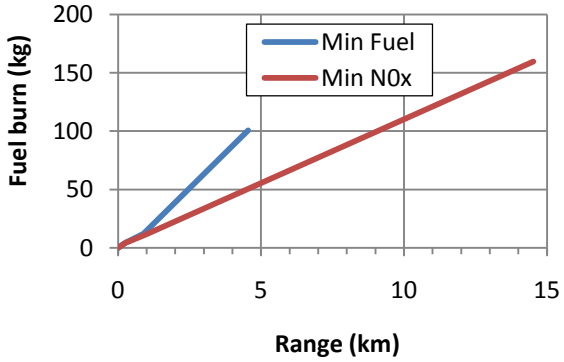


Figure 4: Fuel burn for minimum fuel burn and minimum NOx emissions during initial climb

3.2 The Second Stage – Climb, Cruise and Descent

The two-objective (fuel-time) optimisation of the trajectory profile presented resulted in the pareto frontier presented in Figure 5. The frontier exhibits good convergence with a smooth characteristic throughout the operational regime. As expected, the frontier demonstrates how reduced flight time is achieved at the cost of higher fuel burn. The relative weighting given to fuel burn and flight time will define the optimal operating point on the frontier. This operating point is, as for the case of the NOx/Fuel Burn front consideration, identified by finding the point at which the slope of the tangent defines this relative weighting. For example, if flight time and fuel burn are given an equal weighting, the optimal operating point would be the point where the slope is equal to -1, which is in the region of 7,200kg fuel burn and a flight time of 2 hours and 13 minutes (8,000 seconds) for the trajectory constraints defined in this work.

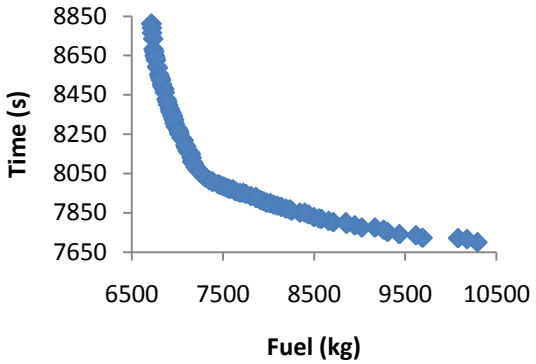


Figure 5: The Pareto Frontier for Time and fuel burn for the second stage

Although the models are generic in nature, the pareto frontier suggests good representation of aircraft of the class, as a similar trajectory profile for a 60 tonne aeroplane would involve a fuel burn of about 6.6 tonnes over 2½ hours and a cruise altitude of 38,000ft. Fuel burn at 38,000ft can be expected to be about 10% less than that at 34,000ft at Mach 0.78, which, due to the higher true airspeed at lower altitude, will translate to about a 5% greater fuel efficiency at the higher altitude over a 2000km sector.

It is interesting to analyse the trajectory strategies determined by the optimisation algorithm for minimum fuel burn and minimum time of flight. Figures 6 and 7 present the flight profiles in terms of altitude vs. range and speed at the start and end of each segment (which represent pseudo-waypoints).

3.2.1 Minimum Fuel Burn Trajectory

For the case of minimum fuel burn, the strategy adopted is the expected, namely climbing to cruise at maximum altitude for fuel

efficiency. However, the results show that the climb is constrained, as the aircraft reaches the maximum altitude in all the first eight segments (Table 2). Indeed, an unconstrained aircraft of the class represented in this work should achieve FL340 within about 200km at a weight of 60 tonnes. This indicates that greater efficiency would be achieved if the aircraft were allowed to climb quicker, which it is capable of doing. It is interesting to note that the selected climb speed is higher than the cruise speed, as shown in Figure 6. Whilst this results in a greater fuel burn in the initial part of the climb, it affords the aircraft to achieve higher altitudes earlier and thus save fuel through prolonged flight there. In essence, this implies a trade-off in fuel burn between segments of the trajectory for global optimisation. This strategy is also evident in Figure 8, which illustrates a high fuel flow (burn) rate in the first segment (close to that for minimum flight time, which disregards fuel burn considerations) in order to climb away from the lower altitudes as quickly as possible. This is also understandable in the context that, for a given CAS, TAS increases disproportionately with altitude. In still air, TAS is what defines the ground speed and hence the flight time. As a result, it is expected that it will be advantageous to expedite climb beyond that providing minimum fuel burn during climb in order to reach higher altitudes quicker. It will be interesting, in future work, to analyse the effect of the altitude constraint during climb on the optimal climb speed.

In cruise, the aircraft would be expected to climb or to slow down gradually as the aircraft becomes lighter through progressive fuel burn. However, in this test case, the aircraft flies at the constraints of maximum altitude and minimum mach number, thus denying the optimisation process from adopting one of these strategies. Consequently, the fuel flow rate will be rising gradually during cruise (Segments 3-8) as a result of the aircraft progressively flying further away from its maximum range operating point.

The descent of the minimum fuel burn trajectory commences towards the end of the 8th segment and this is only because the trajectory constraint required the aircraft to be at or below FL280 by the end of this segment. Consequently, the majority of the descent, that from FL280 down to FL80, was performed in the 146km of the last two segments. The average rates of descent in these two sectors are computed to be just under 2000 ft/min and 1500 ft/min respectively. The descent in segment 9 is also constrained in that the initial altitude is at FL290, the highest possible to maintain efficient cruise, whilst the final altitude attained is FL10. Considering that segment 9 extends 132km, the average glide ratio will be just under 23:1, which will need thrust above idle, since a glide ratio of about 20:1 is achieved with idle thrust. This implies that, due to limitations in the problem setup, the optimal trajectory is limited by the constraint of the initial and final conditions of the descent segment, thus not affording the aircraft to fly longer (or higher) in segment 8, which then results in the higher than idle thrust setting during the descent in segment 9. It is interesting to observe that the end speed of segment 9 rises to a CAS of 337kts and then the excess kinetic energy is traded off into potential energy, resulting in a relatively low rate of descent during segment 10, as the aircraft decelerates to the final CAS of 250kts at idle thrust. This strategy is probably the same as that adopted during climb, where the aircraft flies faster at lower altitudes to reduce the time spent at lower altitudes, where fuel burn is higher.

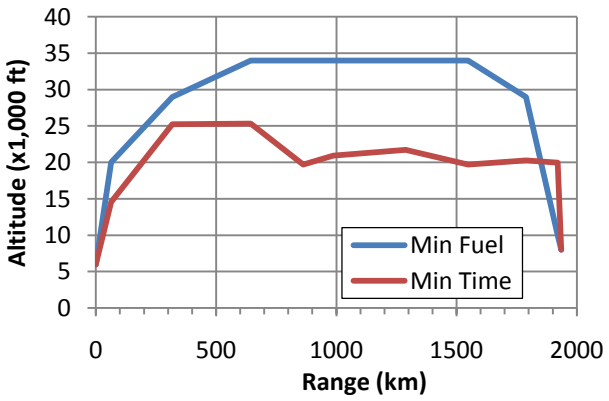
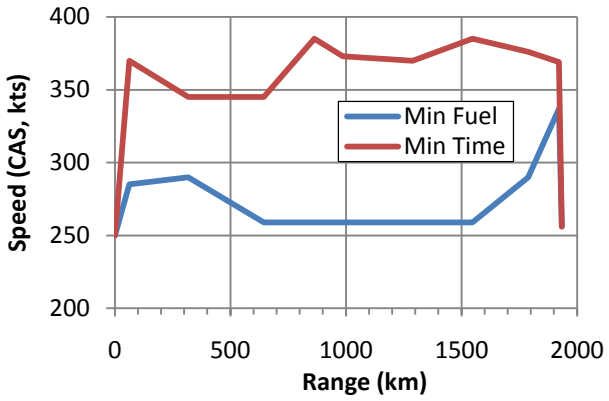
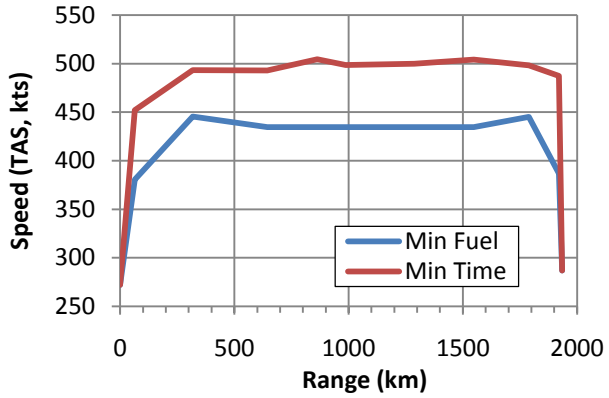


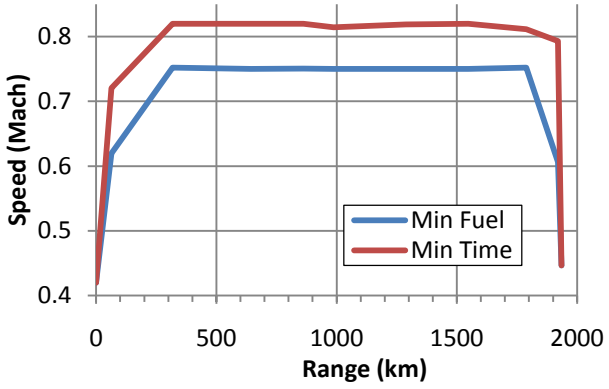
Figure 6: Trajectory profiles for minimum fuel burn and minimum time of flight



(a)



(b)



(c)

Figure 7: Speed profiles for minimum fuel burn and minimum time of flight (a) CAS, (b) TAS, (c) Mach.

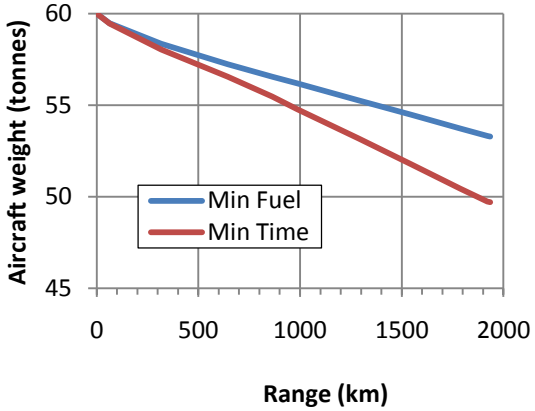


Figure 8: Weight profiles resulting from fuel burn during flight for minimum fuel burn and minimum time of flight

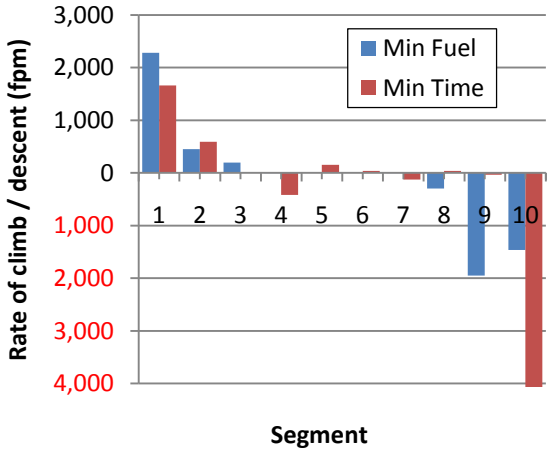


Figure 9: The average rate of climb and descent in each segment for minimum fuel burn and minimum time of flight

3.2.2 Minimum Flight Time Trajectory

The minimum flight time trajectory exhibits a lower climb gradient (Segments 1 and 2) than that in the trajectory for minimum fuel burn, in order to afford a quicker TAS during the climb. The aircraft reaches FL250 at about the same time as when the minimum fuel burn trajectory is flown, but in this case it will have covered a greater range. This implies that, for minimum flying time, the climb gradient is traded for

faster climb speeds. In this case, the aircraft accelerates to the maximum constraint speed of 370kts CAS in the first segment and this affords the aircraft to climb to FL145 by the end of the segment. The result suggests that it may be more advantageous for the aircraft to climb faster than 370kts CAS with a lower climb gradient if it were not constrained to fly at this speed.

Clearly, however, a balance between climb rate and speed needs to be found. A higher climb rate affords the aircraft to fly a higher TAS (and thus a higher ground speed) quicker and achieves higher altitudes earlier, until the flight becomes mach limited above FL200.

Since the intended cruise segments (3 to 8) were constrained in terms of mach number (M0.75-0.78), the trajectory understandably settled at the lowest altitude allowed (FL200), which results in the highest possible ground speed.

The climb to FL250, essentially overshooting the optimal altitude of FL200 may not be expected theoretically. This strategy may be due to the relatively large length of segment 2 (254km), the fact that FL200 is traversed (which leads to the discrete change over from CAS to Mach constraints) and the limitation of the models used in the optimisation, where the trajectory parameters change only gradually between the initial and final states of the segment. In such circumstances, it is probable that the preferred (faster flight) strategy was to climb above FL200 to enable the aircraft to fly at higher TAS. This is plausible, given that at FL200, M0.82 results in a CAS of 383kts, which is higher than the CAS limit defined in the constraint below FL200. This means that, if the flight in the segment remained below FL200, the aircraft would have flown at the CAS limit of 370kts, which would result in a TAS below 490kts, which is lower than that achieved by the selected strategy, although only marginally so (approximately 1 to 2%).

The strategy of varying altitude in cruise is not evident to the authors and indeed, they are of the opinion that this may be due to the optimisation process not fully converging to the actual minimum point of the mathematical function. Indeed, with the aircraft at FL250 at the end of segment 2, the trajectory would have been expected to descend to FL2000, maintaining that altitude until the end of the cruise, flying at Mach 0.82. This would have resulted in a constant TAS of 503.7kts, affording an improvement of the order of 0.3% in the overall flight time.

The graphs indicate that descent strategy adopted for the minimum flight time trajectory is to maintain FL200 until the end of segment 9, prior to rapid descent and deceleration to 256kts CAS in segment 10. This allowed the aircraft to fly at the highest TAS during cruise and to then descend rapidly, again flying at the highest CAS of 370kts. Thus, a rapid descent from FL200 down to FL10 was achieved with an average rate of descent of 9300ft/min, which, although mathematically possible from a performance point of view with the models used in this work, would not be practical in practice for several operational reasons. Nevertheless, the result clearly illustrates the strategy adopted for minimum flight time. It is also relevant to observe that the aircraft is constrained to slow down during segment 8 because of the speed constraint of 370kts CAS, which is Mach 0.79 at FL200. Being constrained by CAS rather than Mach justifies the strategy of maintaining the highest allowed altitude, which is again constrained at FL200.

4 Concluding Remarks

This paper has presented the results of an optimisation of a 2000km trajectory that is bound by operational constraints in terms of altitude and speed. The results have illustrated the strategies adopted by the optimisation process with such constraints and has highlighted the limitations introduced in the problem definition and by the limited number of segments in the trajectories. It has formed the basis of a preliminary study against which more complex problems can be compared in future.

5 Future Work

It will be interesting to assess the effect a number of characteristics of the problem definition may have on the output of the optimization process. The performance of genetic algorithms is affected by factors such as population size, number of generations and cross-over algorithms and these may have an impact on convergence as well as processing time. Given the heuristic nature of the optimisation technique, it is also interesting to assess the repeatability of such results. Another point of significant interest is the consideration of the impact the number of segments into which the trajectory is partitioned for the optimisation process has on the accuracy of the results. Whilst it is obvious that a larger number of segments will generate higher accuracy results, it is interesting to study the nature of the errors (or approximations) brought about by the segmentation process. This is particularly of interest in the light that the performance model has been designed and implemented to cater for dynamic and non-linear effects within a single segment, including changes in local atmospheric conditions (which are particularly relevant during climb and descent). The complexity of the model has been designed to offset some of the limitations introduced by the large segments and it is therefore of interest to assess how the overall process performs in this respect.

Finally, it is also interesting to assess how the strategy will vary with changing operational conditions such as dispatch weight and altitude constraints.

CEAS 2011

CFD Applications 2

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Preliminary evaluation and optimization of civil aircraft sonic-boom properties through Carlson's method

Gennaro Scarselli and Elisa Castorini
Dipartimento di Ingegneria dell'Innovazione
Università del Salento, Italy

Keywords: Sonic-boom evaluation, Carlson's method

Abstract

In this paper a simple approach based on the Carlson's method will be presented to define a supersonic configuration optimized in terms of sonic boom properties. The Carlson's method provides a simplified procedure for the calculation of sonic boom characteristics for supersonic airplane configurations and spacecrafts. The information required for the signature predictions are: aircraft shape factor K_S, aircraft operating conditions and atmospheric data. Unfortunately, there is not an analytic expression of the shape factor. Nevertheless, a graphic representing the relationship between K_S and parameters related to the aircraft geometry, can be used. In this paper K_S have been approximated through a linear and a quadratic interpolation as a function of other parameters and its minimization problem has been formulated. Computational results show the optimal values to be assigned to the aircraft geometry parameters in order to obtain the minimal value of the shape factor and in consequence of the sonic boom overpressure.

Nomenclature

α	=	angle of attack (deg)
$A(x)$	=	cross sectional areas (m^2)
$A_e(x)$	=	total effective area (m^2)
$A_{e,max}$	=	maximum effective area (m^2)
$A_{e,1}(x)$	=	total effective area at midpoint of effective aircraft length (m^2)
a_v	=	speed of sound at aircraft altitude (m/s)
$B(x)$	=	equivalent cross sectional area due to lift (m^2)
h	=	altitude of aircraft above ground (km)
h_e	=	effective altitude (km)
K_p	=	pressure amplification factor
K_R	=	reflection factor, assumed to be 2.0
K_S	=	shape factor
K_t	=	signature duration factor
l	=	aircraft characteristic length (usually the fuselage length) (m)
l_e	=	effective length (m)
M	=	aircraft Mach number
M_e	=	aircraft effective Mach number
Δp_{max}	=	incremental pressure at N-wave bow shock (Pa)
p_g	=	atmospheric pressure at ground level (Pa)
p_v	=	atmospheric pressure at aircraft altitude (Pa)
Δt	=	time increment (s)
W	=	aircraft weight (kg)
x	=	distance from aircraft nose measured backward along flight path (m)

1 Introduction

In the last years there has been a renewed interest in the supersonic transport due to recent research advances in the field of the low sonic boom configurations and to the increasing demand from business people of aircraft able to cover medium and long distances in very short times. These two circumstances have promoted many research projects worldwide aimed to the development of business jets able to travel at speeds above Mach 1.0. The supersonic flight for civil application represents one of the most challenging fields in the aeronautic research since different conflicting requirements must be satisfied in order to make viable the supersonic aircraft both from a commercial and technical standpoint. For instance, the airframe designer wish to drop the weight by thickening the wings but the aerodynamic designer wish to thin the wings to decrease the drag. The noise team wish to increase the By Pass Ratio (BPR) but the aerodynamic team wish to diminish the BPR to lower the drag in supersonic cruise. The engine designer wish to increase the turbine inlet temperature, but the engineers in charge of environmental issues wish to decrease the combustion temperature which is a mean of decreasing NOx emissions. Supersonic propulsion must cope with two conflicting constraints: one requirement of low drag during transonic and supersonic flight, which implies a small fan diameter, and a requirement of acceptable noise at take off obtained by increasing the BPR, which drives towards large fan diameters. These are few elements that clarify why a supersonic civil aircraft is a challenging technical product, final goal of a MultiDisciplinary Optimization process able to manage these conflicts and concerns "right first time" rather than through a high number of design loops. One of the main concerns is the sonic boom that accompanies travel beyond the speed of sound. Successful supersonic civil aircraft must be capable of supersonic flight over land. In order to achieve this end result, questions associated with sonic boom noise must be addressed. Existing knowledge of the impact of sonic boom on the community is based primarily on experiments conducted during the 1960s (Concorde and the U.S. SST) and the 1980s NASA's High Speed Research (HSR) Program. In the former work it was concluded that high-amplitude sonic booms (Concorde flight was able to product about 2 psf of overpressure at ground level) were clearly unacceptable to a large segment of the population, and overland flight was prohibited. Although much progress was made in modeling the sonic boom and its effects during HSR, the boom levels achieved were still unacceptably loud. Worldwide different Research Projects are developing technologies that potentially will lower the boom to acceptable levels. There is very little data on the effects of such booms both on the population and on the civil structures: a reasonable approach should face three fundamental technical items: the atmospheric effects, transmission into structures, and human reaction to these booms. The ability to model sonic boom propagation from the aircraft to the ground is a necessity, with all relevant physical phenomena included in such models. This will enable the accurate prediction of sonic boom levels on the ground under realistic atmospheric conditions and for all flight conditions. In addition, in order to fully understand human reaction to these low-intensity sonic booms, it is necessary to be able to predict the transmission of booms into buildings. Both interior noise levels and structural vibration are of interest since both are important characteristics of the indoor environment. Studies of human response to sonic booms heard indoors and outdoors have identified several factors that may contribute to an increased annoyance indoors, including reverberation and rattle noise. Investigation of these factors through laboratory and field studies is critical to understanding and developing prediction models for human response to low-intensity sonic booms. For a complete and numerically

accurate description of sonic boom phenomena, 3D CFD [1] models should be applied and the simulation of the signal at the ground needs to couple CFD equations and acoustic propagation equations. Thus, the resolution is performed in two steps. Firstly, the flow close to the aircraft, named aerodynamic near field region, is computed by solving the conservative Euler equations for gas dynamic. The pressure distribution obtained at the lower boundary of the near field region is then used as initial condition for the propagation of the acoustic wave to the ground. The near field perturbations are propagated to the ground using the waveform parameter method which describes the pressure wave. Nevertheless, modeling the CFD flow is a tri-dimensional and non-linear whereas the propagation model is mono-dimensional. In order to have a valid coupling between both models, the near field solution must be locally axisymmetric. This condition is necessary to take into account all elements of the geometry of the aircraft (body, wings, etc.). Consequently, an accurate near field solution sufficiently far from the aircraft is required to obtained a good prediction of the sonic boom. It is reasonable to perform this kind of analysis when the aircraft design is already at a mature stage. For a preliminary definition of the aircraft geometry more simple approaches should be used to design a distribution of cross sectional areas and planform able at the same time to achieve low boom effects. In the following paper sonic boom is approached through a simple model allowing to relate the sonic boom maximum overpressure to the main geometric and operative parameters of the civil aircraft. This, indeed, does not reveal an accurate description, as CFD could be, of the pressure field on the ground but provides a reliable estimation of the maximum levels expected generated by an assigned aircraft geometry. A relation between the maximum sonic boom overpressure and an aircraft shape factor exists that can be managed in order to define an optimal preliminary design point for the supersonic civil aircraft. This relation, based on the results of previous works, is mathematically approximated and an optimization problem is formulated in order to provide technical guidelines suggesting an optimal preliminary aircraft geometry from a sonic boom standpoint.

2 Carlson's method

In this paragraph the method is briefly outlined and the basic steps for the sonic boom evaluation are described and discussed: every further technical detail not reported hereinafter can be easily found in the [2], [3]. The information required for the calculations and the pressure-signature predictions provided by the simplified method are: aircraft shape factor; aircraft operating conditions; atmospheric data. The method provides the following information: N-wave bow-shock overpressure; the signature duration; the location of the ground impact point relative to the aircraft position at the time the boom was generated. Three basic steps must be followed: determination of an aircraft shape factor, evaluation of atmospheric propagation factors, and calculation of signature shock strength and duration. For aircraft covered by the charts available for some specific families of airplanes, the shape factor is evaluated by using aircraft weight and operating conditions to calculate lift parameter K_L and to read K_S directly. The lift parameter is given by the following equation in which parameters should be known concerning the operative conditions of the supersonic aircraft:

$$K_L = \frac{(\sqrt{M^2-1})W\cos\gamma\cos\theta}{1.4\,p_v M^2 l^2} \qquad (1)$$

Instead for aircraft not covered by the before mentioned charts the aircraft shape factor is computed by the aircraft geometry. At first the cross sectional areas are evaluated: for this evaluation the areas normal to the flight path should be used but also the areas normal to the aircraft longitudinal axis are acceptable except for high

values of angle of attack. Then the equivalent area due to lift should be evaluated: a reasonable approximation of the lift distribution may be defined by the planform area distribution. The equivalent area due to lift is defined by the equation:

$$B(x) = \frac{(\sqrt{M^2-1})W\cos\gamma\cos\theta}{1.4\,p_v M^2 S}\int_0^x b(x)dx \tag{2}$$

The third step is the combination of these two area contributions to obtain a distribution for the total effective area of the aircraft illustrated in the Fig. 1. Every family of supersonic air-

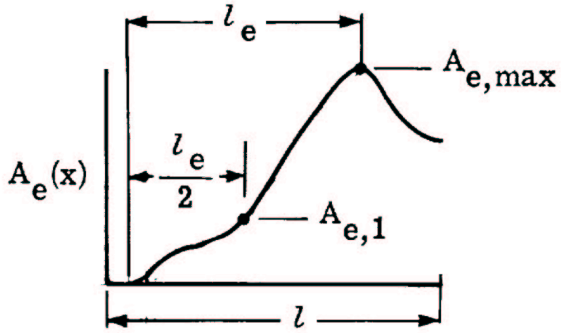


Figure 1: Total effective area

plane exhibits its own specific profile of equivalent areas. From this curve the maximum effective area, $A_{e,max}$, and its location can be determined. Then, as shown in Fig. 1, the effective length l_e and the effective area $A_{e,1}$ required for evaluation of the aircraft shape factor may be selected. Finally, in the fourth step, the aircraft shape factor may be found from a reading of the shape-factor parameter curve of Fig. 2 and insertion of the appropriate areas and length. It can be noted that the aircraft shape factor is primarily dependent on the maximum effective area: the location of the maximum area and shape of the effective area development play lesser roles. Atmospheric propagation factors for sonic boom calculations have been obtained by a repeated use of a dedicate code for a matrix

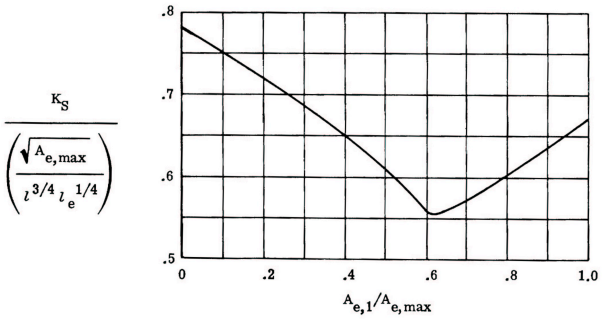


Figure 2: Relationship among shape factor and other parameters

of altitudes, Mach numbers, ray-path azimuth angles, and flight path angles [4], [5]. In order to extend the applicability of these atmospheric propagation factors and to use them for on-track and off-track locations and for different flight path angles, the concepts of effective Mach number, M_e, and effective altitude, h_e, are introduced. Effective Mach number is the Mach number for level flight which would have the same ray-path angle in the flight track plane. Effective altitude is the ray-path distance measured perpendicular to the aircraft flight path. The general equations for effective Mach number, propagation distance and effective altitude are:

$$M_e = \sqrt{1 + \frac{\left[C\left(1 - D\tan\gamma\right)\right]^2}{\left[\left(\tan\gamma + D\right)C\right]^2 + \left[D\sin\theta(\tan^2\gamma + 1)\right]^2}} \tag{3}$$

$$d = K_d\left(\frac{h}{\sqrt{M_e^2 - 1}}\right) \tag{4}$$

where

$$C = 1/(\cos\gamma\sqrt{M^2-1})$$
$$D = 1/(\cos\theta\sqrt{M^2-1})$$
$$h = h_v - h_g$$

and K_d is available from Figure 7b of the [2].

$$\phi = \tan^{-1}\left[\frac{\tan\theta\cos\gamma\left(1+\tan^2\gamma\right)}{\tan\gamma + D}\right] \tag{5}$$

$$h_e = \sqrt{d_y^2 + (h\cos\gamma + d_x \sin\gamma)^2} \qquad (6)$$

where $d_x = d\cos\phi$ and $d_y = d\sin\phi$. From the knowledge of the parameters M_e and h_e it is possible to evaluate the sonic boom overpressure and the signature duration by the following equations:

$$\Delta p_{max} = K_p K_R \sqrt{p_v p_g}\left(M^2-1\right)^{\frac{1}{8}} h_e^{-\frac{3}{4}} l^{\frac{3}{4}} K_S \qquad (7)$$

$$\Delta t_{min} = K_t \frac{3.42}{a_v} \frac{M}{\left(M^2-1\right)^{\frac{3}{8}}} h_e^{\frac{1}{4}} l^{\frac{3}{4}} K_S \qquad (8)$$

The coefficients K_p, K_t (atmospheric factors) are available as a function of M_e in charts that can be found in [2]. The equations (7) and (8) refer to the assumption of N-wave that can be described simply by the overpressure and the duration: it is clear from equation (7) that the maximum overpressure due to sonic boom, fixed the operative conditions in terms of altitude, Mach number, depends primarily on the shape factor K_S: therefore minimizing the sonic boom means to define a geometry able to minimize the value of K_S. Obviously this finding derives from a simple method with the before mentioned limitations and a more accurate prediction is requested for taking in account effects here neglected.

3 Approximation of Shape factor function

An evaluation of shape factor can be obtained using the graphic in Fig. 2. To this aim, the curve in Fig. 2 has to be approximated such that K_S can be expressed as function of other dimensional parameters of the aircraft. Therefore, an optimization problem can be formulated. Two kinds of approximation have been taken into account and, in particular, a linear interpolation has been compared with a quadratic one.

3.1 Linear approximation function

Three points belonging to the curve in Fig. 2 have been evaluated; a piecewise linear function has been found and K_S can be expressed as function of $A_{e,1}$, $A_{e,max}$ and l_e:

$$K_S = \begin{cases} K'_S = & a_1\, A_{e,1}\,(A_{e,max})^{-\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} + \\ & +b_1\,(A_{e,max})^{\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} \\ & \text{if } \; A_{e,1}/A_{e,max} \in [0,c] \\ K''_S = & a_2\, A_{e,1}\,(A_{e,max})^{-\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} + \\ & +b_2\,(A_{e,max})^{\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} \\ & \text{if } \; A_{e,1}/A_{e,max} \in [c,1] \end{cases} \qquad (9)$$

with $a_1 = -0.029\, m^{-\frac{3}{4}}$, $a_2 = 0.025\, m^{-\frac{3}{4}}$, $b_1 = 0.062\, m^{-\frac{3}{4}}$, $b_2 = 0.029\, m^{-\frac{3}{4}}$, $c = 0.62$. These coefficients result from the linear interpolation of the before mentioned curve with the two segments passing through the three points considered. Therefore, a reference geometry for the supersonic airplane configuration [7], suggesting reasonable values for the upper and lower bounds of the variables, has been assumed. This geometry is shown in Fig. 3. Hence, the optimization problem can be formulated as follows:

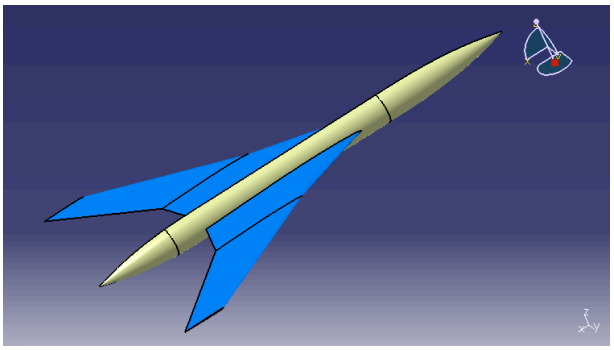


Figure 3: Reference geometry

$$\begin{aligned} &\min \; K_S \\ &s.t. \;\; A_{e,1} - d\, A_{e,max} \le 0 \\ &\quad lb_1 \le A_{e,1} \le ub_1 \\ &\quad lb_2 \le A_{e,max} \le ub_2 \\ &\quad lb_3 \le l_e \le ub_3 \\ &\quad A_{e,1},\, A_{e,max},\, l_e \in \mathbb{R}_+ \end{aligned} \qquad (10)$$

where $d = 0.7$, the lower bounds are: $lb_1 = lb_2 = 5.0\, m^2$, $lb_3 = 15.0\, m$ and the upper bounds are:

$ub_1 = ub_2 = 45.0\,m^2$, $ub_3 = 30.0\,m$. The problem (10) has been solved with MATLAB and the optimal solution is the following:

$$A_{e,1} = 5.0\,m^2$$

$$A_{e,max} = 7.1\,m^2$$

$$l_e = 30.0\,m$$

and the optimal value of the shape factor is $K_S = 0.0531$.

3.2 Quadratic approximation function

A quadratic approximation of the curve in Fig. 2 has been built through the second order divided differences [6] and the optimization problem can be written as follows:

$$\begin{aligned} \min\ & K_S = a_3\,(A_{e,1})^2\,(A_{e,max})^{-\frac{3}{2}}\,(l_e)^{-\frac{1}{4}} + \\ & +b_3\,A_{e,1}(A_{e,max})^{-\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} + \\ & +c_3\,(A_{e,max})^{\frac{1}{2}}\,(l_e)^{-\frac{1}{4}} \\ s.t.\ & A_{e,1} - d\,A_{e,max} \leq 0 \\ & lb_1 \leq A_{e,1} \leq ub_1 \\ & lb_2 \leq A_{e,max} \leq ub_2 \\ & lb_3 \leq l_e \leq ub_3 \\ & A_{e,1},\ A_{e,max},\ l_e \in \mathbb{R}_+ \end{aligned} \tag{11}$$

with $a_3 = 0.053\,m^{-\frac{3}{4}}$, $b_3 = -0.062\,m^{-\frac{3}{4}}$ and $c_3 = 0.061\,m^{-\frac{3}{4}}$. The optimal solution provided by MATLAB of the problem (11) is:

$$A_{e,1} = 5.0\,m^2$$

$$A_{e,max} = 7.1\,m^2$$

$$l_e = 30.0\,m$$

with an optimal value of K_S equal to 0.0498. Therefore, the solution obtained with quadratic approximation is better than one achieved with linear approximation. Moreover, it would be desirable to understand the dependence of the shape factor by the other parameters. To this aim, one of the variables has been fixed to some established values and the related optimization problems have been solved. In particular, 5 different values have been assigned to $A_{e,max}$ and the 5 optimal solutions have been found, as shown in Table 1.

Case	$A_{e,1}\,m^2$	l_e (m)	K_S
$A_{e,max} = 8.0\,(m^2)$	5.0	30.0	0.0515
$A_{e,max} = 11.0\,(m^2)$	6.7	30.0	0.0598
$A_{e,max} = 14.0\,(m^2)$	8.3	30.0	0.0681
$A_{e,max} = 17.0\,(m^2)$	10.7	30.0	0.0731
$A_{e,max} = 25.0\,(m^2)$	15.5	30.0	0.0893

Table 1: Optimal solutions for fixed values of $A_{e,max}$.

Finally, the shape factor K_S has been plotted as function of $A_{e,1}$ and l_e as shown in Figs. 4, 5, 6, 7 and 8.

In order to better understand the behavior of

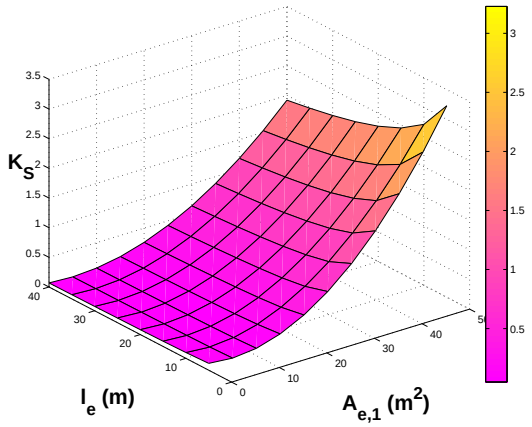


Figure 4: $A_{e,max} = 8\,m^2$

K_S respect to the other parameters $A_{e,1}$, $A_{e,max}$ and l_e, two of three parameters have been fixed, time by time and K_S has been studied as function of only one parameter. The simple supersonic reference geometry before mentioned has been used for this sensitivity study focusing the attention on one specific working configuration in terms of operating conditions and flight parameters: $M = 1.4$, $W = 46313\,kg$, flight altitude of $13.7\,km$ and with an angle of attack of $3\,deg$. Taking into account this configuration, the following cases have been examined:

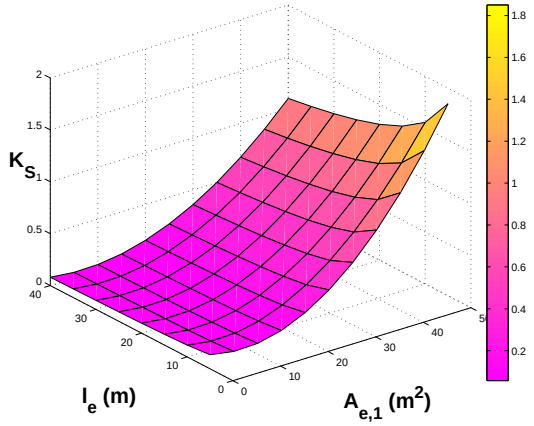


Figure 5: $A_{e,max} = 11\,m^2$

- case a: $A_{e,max} = 12.7\,m^2$ and $l_e = 25.5\,m$;
- case b: $A_{e,1} = 5.0\,m^2$ and $l_e = 25.5\,m$;
- case c: $A_{e,1} = 5.0 m^2$ and $A_{e,max} = 12.7 m^2$.

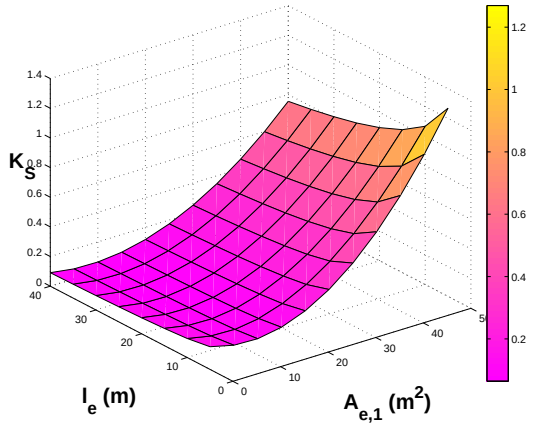


Figure 6: $A_{e,max} = 14\,m^2$

In Figs. 9, 10 and 11 the shape factor K_S as function of $A_{e,1}$, $A_{e,max}$ and l_e respectively has been represented.
In case a K_S has a parabolic behavior, so decreases slightly until the minimum point, evaluated for $A_{e,1} = 7.5\,m^2$, and then increases. In case b K_S has a quick decrease until minimum point near $A_{e,max} = 6.0\,m^2$ and then very slowly increases. In the last case K_S has a monotonic decrescent behavior. Therefore, the minimal values of K_S are obtained for small values of $A_{e,1}$ and $A_{e,max}$ and big values of l_e. These findings suggest technical guidelines in order to find the optimal configuration in terms of minimization of the sonic boom effects with the limitations on which the Carlson's method rely: more refined CFD approach could support a further improvement of this preliminary configuration at a later stage of the supersonic aircraft design.

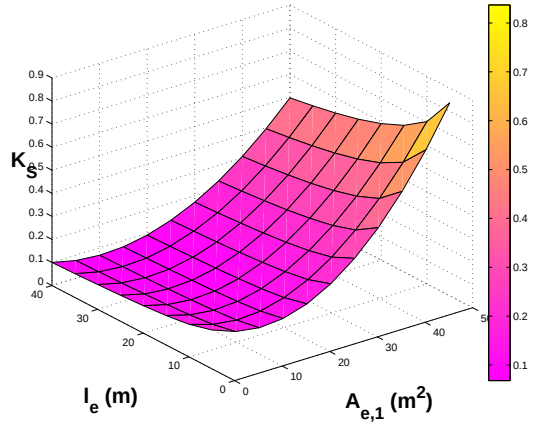


Figure 7: $A_{e,max} = 17\,m^2$

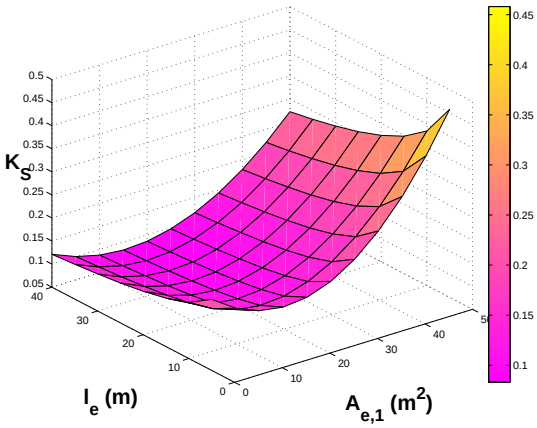


Figure 8: $A_{e,max} = 25\,m^2$

4 Conclusion

In this paper, sonic boom has been evaluated through a simple approach that allows to estab-

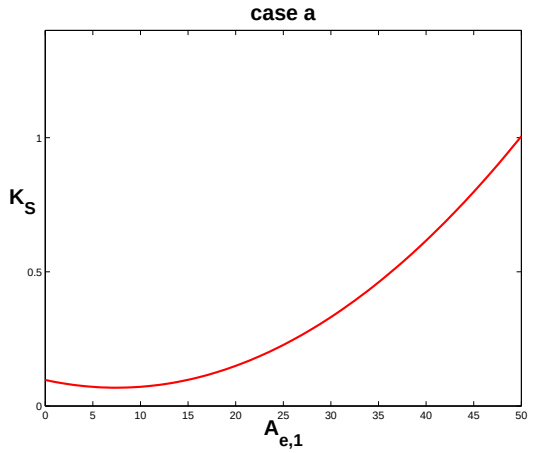


Figure 9: K_S as function of $A_{e,1}$

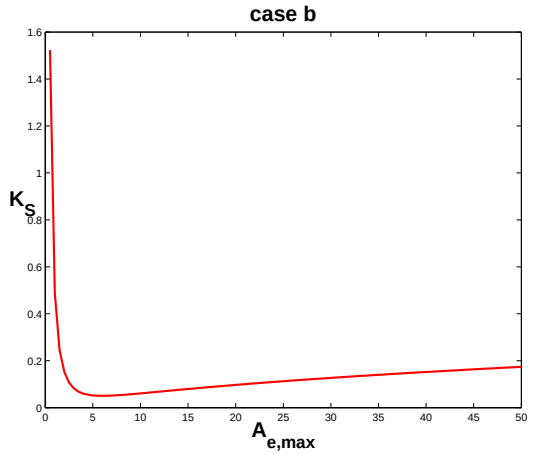


Figure 10: K_S as function of $A_{e,max}$

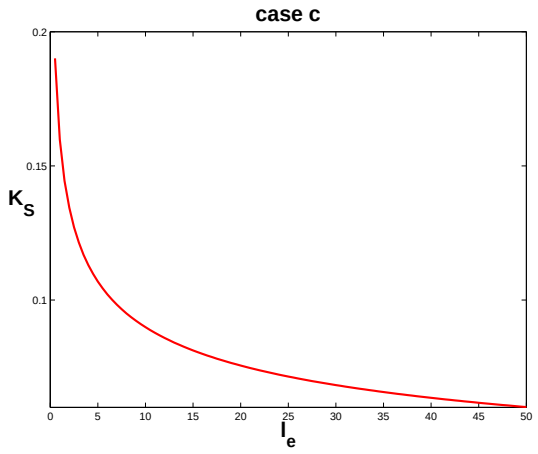


Figure 11: K_S as function of l_e

lish a relation between the sonic boom effects and the operative parameters of the civil aircraft. This relation enables to define an optimal preliminary design point for the supersonic civil aircraft. Applying the Carlson's method, a procedure to evaluate the minimum value of the shape factor and, in consequence, of the sonic-boom effects in terms of overpressure and signature duration has been presented. The outlined procedure can be adopted also for complex variations of the aircraft geometric parameters (wing positions along the fuselage, distribution of cross-sectional and planform areas, etc.) in order to obtain optimal sonic-boom properties.

5 Bibliography

References

[1] Milne-Thomson, L. M., "Theoretical Aerodynamics", *Dover Publications*, 1973.

[2] Carlson, H. W., "Simplified Sonic-Boom Prediction", *NASA Technical Paper 1122*, Langley Research Center, Hampton, Virginia, 1978.

[3] Carlson, H. W., and Maglieri, D. J., "Review of Sonic Boom Generation Theory and Prediction Methods", *Journal of the Acoustic Society of America*, Vol. 51, No. 2, pp. 675-685 3, 1972.

[4] Hayes, W. D., and Runyan, H. L., Jr., "Sonic boom propagation through a stratified atmosphere", *Journal of the Acoustic Society of America*, Vol. 51, No. 2, pp. 695-701, 1972.

[5] Hayes, W. D., Haefeli, R. C., and Kulsrud, H. E., "Sonic boom propagation in a stratified atmosphere, with computer program", *NASA CR-1299*, 1969.

[6] Abramowitz, M. and Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York, 1972.

[7] Scarselli, G. and Marulo, F., "Sonic boom minimization through a simplified approach for the preliminary design of civil supersonic aircraft", *Proceedings of the 16th AIAA/CEAS Aeroacoustics Conference*, Stockholm, Sweden, 2010.

3rd CEAS Air&Space Conference
21st AIDAA Congress

An Application of CFD Simulation to Investigate Compressible Angle of Attack Flow Fields around the Finned Missile

A. Ridluan, W. Charubhun, and P. Chusilp

Aerodynamics Group, Design and Simulation Laboratory, Defense Technology Institute, Nonthaburi, Thailand 11120

Keywords: *Compressible flow, Missile design, Missile Aerodynamics*

Abstract

The Computational Fluid Dynamics (CFD) was conducted to investigate the flow structures and aerodynamic characteristics of finned missile operating within a wide range of Mach numbers (0.6 to 5.5) and angles of attack (0° to 16° degrees). The Steady Reynolds-Averaged Navier-Stokes (SRANS) equations with standard k-ε turbulence model were used to predict the turbulence flow behaviors. The numerical results of aerodynamic coefficients were compared against semi-empirical data. The results revealed the development of complicated vortices and their interaction with fin at the rear part of the finned missile. The behavior of aerodynamics coefficients at angle of attack greater than 12 showed the nonlinear trends. At angle of attack less than 8 degrees, the vortices are symmetric, while asymmetric pattern of the primary vortices is observed angle of attack greater than 12° degrees. The secondary vortices were observed and formed underneath the primary ones, which tend to move away from the missile surface. In addition, the expansion of the longitudinal shock along the missile surface was also observed and named shock lift-up flow pattern. The increasing complexity of vortex flow pattern results in the nonlinear behavior of aerodynamics coefficients.

1 Introduction

Missile instability can be introduced by angle of attack flow. Therefore in modern missile design, consideration of a missile aerodynamics with angles of attack is needed to serve a demand on the maneuverability. The loss of maneuverability is attributed to complex flow physics generated around a missile body, which is characterized by vortices development, curvature effect, turbulent flow separation and reattachment, and local wake flow (downstream).

Even though the missile airflow with varying Angle of Attack (AOA) was investigated and documented by a number of researchers [1-5], however, the Mach numbers were limited only to subsonic and supersonic flow regimes (Mostly Mach number below 3). The AOA flows beyond these regimes need to be investigated since the high performance missile is required to operate at Mach number much higher than the currently-published investigations. In order to effectively control the missile stability, the Aerodynamics group, Defense Technology Institute (DTI) is interested in developing the insight into the aerodynamic structure of fined missile and its influence on the static characteristics of force and moment coefficients at a wide range of flow Mach numbers and angles of attack.

Although experimental fluid dynamics (EFD) has been used to analyze the phenomena of fluid flows in various applications, experiments are increasingly expensive and do not offer the resolution as provided by computational studies. Over the last decade, access and advances in computing technology has enabled CFD to fast become an engineering design tool. In many engineering applications, it facilitates engineering analyses of flow. Recently, EFD increasingly serves as a benchmark dataset for CFD simulation. The use of CFD in design and development process was also substantially reduced the time needed to obtain flow solution and analysis.

This paper thus applies CFD simulation to objectively investigate the finned missile flow field at angle of attack and Mach number varying from 0 to 16 degrees and 0.6 to 5.5, respectively. The Steady Reynolds-Averaged Navier-Stokes (SRANS) equations with standard *k-ε* turbulence model were selected. The numerical results of aerodynamic coefficients are compared against semi-empirical data generated by Missile DatCom.

2 Mathematical Model

In engineering simulations of turbulent flow, the Navier-Stokes equations (NSEs) are modified. The modified equations often applied in engineering application are called the Reynolds-Averaged Navier-Stokes (RANS) equations. The Reynolds averaging process introduces the unknown term, the Reynolds Stress.

In order to simulate turbulent behaviors, the turbulent eddy motion is described by the turbulent viscosity term. Here, the standard *k-ε* turbulence model were selected. The turbulent model is briefly described.

For the standard *k-ε* Model, the turbulent viscosity is computed by a combination of the turbulent kinetic energy (TKE), *k*, and its rate of dissipation, *ε*, as follows,

$$\mu_t = \rho \cdot C_\mu \cdot \frac{k^2}{\varepsilon} \tag{1}$$

The quantities, *k* and *ε*, are obtained from the following equations,

$$\rho \cdot \frac{Dk}{Dt} = \frac{\partial}{\partial x_i} \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_k} \right) \cdot \frac{\partial k}{\partial x_j} \right] + \mathrm{P} - \rho \cdot \varepsilon \tag{2}$$

and

$$\rho \cdot \frac{D\varepsilon}{Dt} = \frac{\partial}{\partial x_j} \cdot \left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon} \right) \cdot \frac{\partial \varepsilon}{\partial x_i} \right] + C_{1\varepsilon} \cdot \frac{\varepsilon}{k} \cdot P - C_{2\varepsilon} \cdot \rho \cdot \frac{\varepsilon^2}{k} \tag{3}$$

P represents the production of turbulent kinetic energy. The terms, $\sigma_k, \sigma_\varepsilon, C_{1\varepsilon}$ and $C_{2\varepsilon}$ are constant values.

3 Detailed Formatting Instructions

As noted, the CFD simulation is conducted to predict the aerodynamics coefficients and flow fields of the missile. SRANS with standard *k-ε* turbulence model was selected. Fig. 1 demonstrates the missile configuration. The missile model is composed of fuze, warhead, cylindrical body, and tail sections. The aft part of the missile is tail section, mainly consisted of the fins.

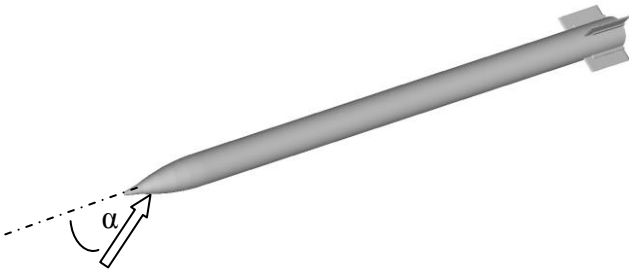


Fig. 1 The schematic of the compressible angle of attack flow over the missile

The CFD simulate the flow conditions by varying the incoming Mach numbers and angle of attacks (α) as sketched in Figure 1(b). In this study, the Mach numbers are varied from 0.6 to 5.5 at low angles of attack (4° and 8° degrees).

The grid (mesh) generation of the constructed missile geometry is shown in Fig. 2. The total numbers of computational cell generated is approximately 3 millions. The multiblock domain with viscous hexahedral

cells is applied to effectively capture the flow structure. Each grid identifies a position where numerical approaches are applied. For viscous mesh topology, the structured hexahedral cells are concentrated near missile surface. To resolve the boundary layer region, y^+ was less than 5 to ensure the resolution of the viscous sublayer.

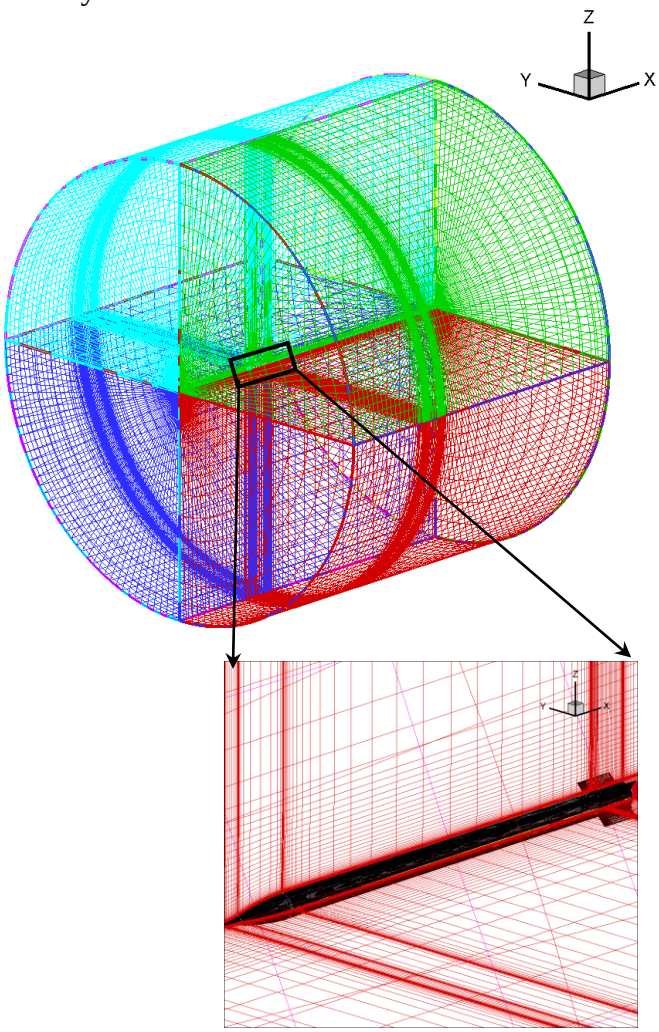


Fig. 2 The computational domain and mesh distribution of the compressible angle of attack flow over the missile

In the present study, we mainly used the commercial CFD Fluent code, the code use a time-independent compressible Navier-Stokes formulation and the standard *k-ε* turbulence model that are discretized using the finite volume method. Further, a second-order upwind and central differencing flow numerical schemes were applied for convective and diffusive terms, respectively. The discrete nonlinear equations were implemented implicitly. To evaluate the pressure field, the ideal gas law was used. The linearized equations were solved by multi-grid method.

The numerical boundaries, i.e. inlet, outlet, solid surface, are specified on the boundaries of the domain and surface. A streamwise far field condition was applied at the inlet, and lateral, and lateral boundaries. This specified constant Mach number. The missile surface is non-slip boundary condition.

4 Results and Discussion

4.1 Flow Structure

The steady flow structures are first presented in terms of contour plots. The Mach numbers computed by Standard *k-ε* SRANS are colored and shown in Figures 3 to 6 for subsonic, transonic, supersonic, and hypersonic incoming flows, respectively. Generally, we observe that the size of the wake region per Standard *k-ε* SRANS becomes smaller as the angle of attack increased for all compressible flow regimes.

For subsonic incoming flow (Fig. 3), the local shock pockets are developed in the vicinity of geometrical transition from head part to missile body and fin region. The shock pockets observed is asymmetric owning to the angle of attack flow. In addition, the higher angle of attack is, the stronger shock spot is developed.

For incoming Mach number 1.2 where the peak of axial force and pitching moment coefficients located, the bow shocks at nose region (in front of the missile) and tail region (in front of fins) are observed. The local shocks are developed at geometrical transition and fin area. At the higher Mach numbers (supersonic and hypersonic), both bow and local shocks downstream tends to attach the missile body as shown in Figs. 4 to 5. Note that the asymmetry flow structure is due to the effect of angle of attack.

Interestingly, we also observe the development of longitudinal shock (along the missile body) in the case of Mach number greater than 1.2 as a function of angle of attack

and Mach number. The increase of angle of attack results in the stronger longitudinal shock development. Referred to Figs. 5 (a) to (d), this behavior is represented by the expansion of shock flow from the missile surface. The longitudinal shock (compressibility) is first observed along the surface at 4 angles of attack as shown in Fig. 5 (a). The upstream and downstream Mach numbers are about 2.4 and 2.0, respectively. At Angles of attack 8°, 12°, and 16° degrees, the downstream Mach numbers are approximately increased to 2.3, 2.5, and 2.7 as shown in Figs. 5 (b) to (d). The increase in the downstream Mach number is associated with the expansion of the longitudinal shock. Based on the observation, we define this phenomenon as "shock lift-up". Finally, this phenomenon is attributed to the development of multi shock region as schematically shown in Fig. 5 (a).

Figs. 7 (a) to (d) display the development of the shock lift-up pattern as a function of Mach number at angle of attack 16. The similar phenomenon is observed. The initial development of shock loft-up is at Mach number 1.2. The strongest shock lift-up is observed at Mach number 2.5 for angle of attack 16, associating the minimum pitching moment coefficient. Beyond Mach number 2.5, the strength of shock lift-up is reduced and finally, it totally disappears at Mach number 5.5, as show in Fig. 6 (d).

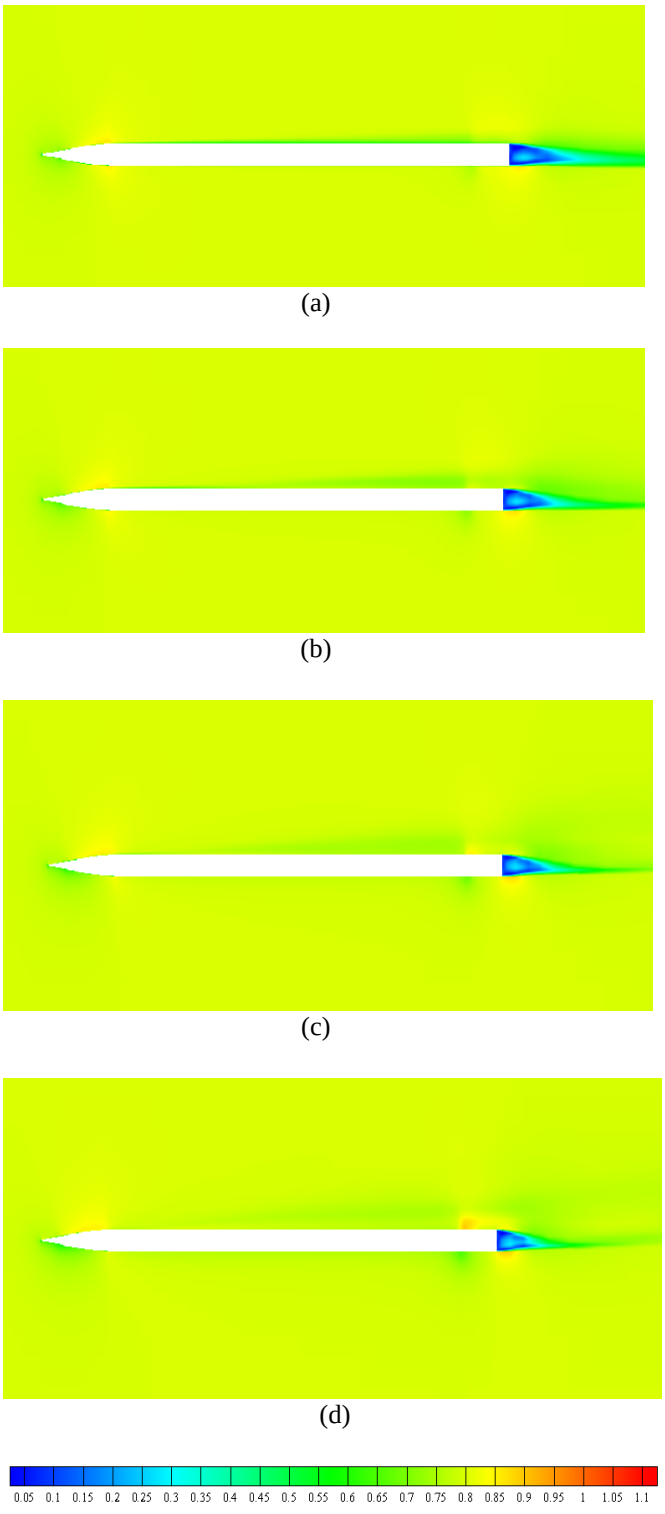
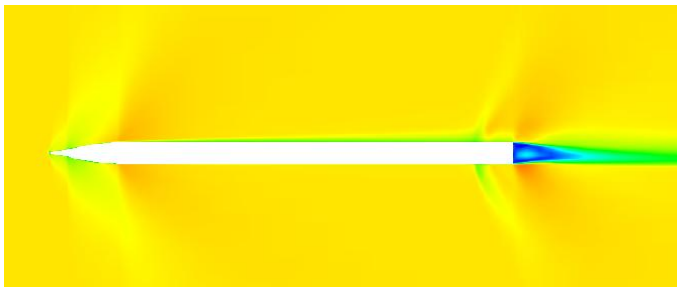
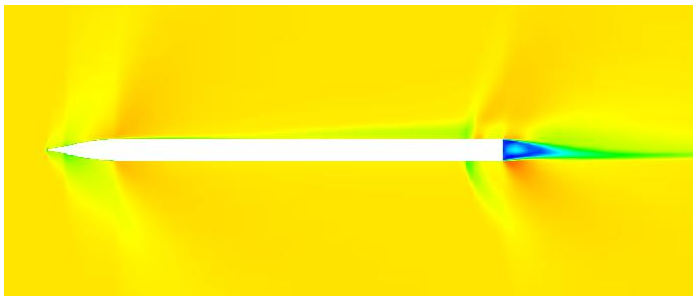


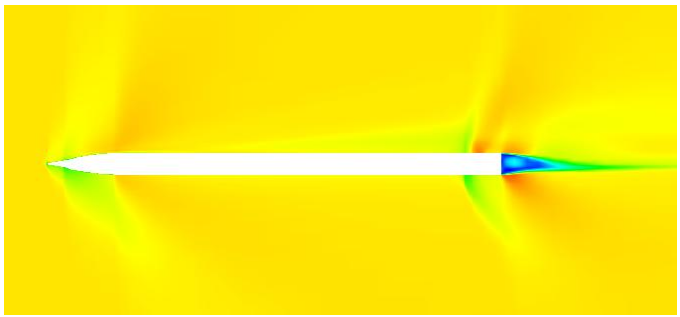
Fig. 3 The contour plot of subsonic flow (Mach number of 0.8) at angle of attack (a) 4°, (b) 8°, (c) 12°, and (d) 16° degrees



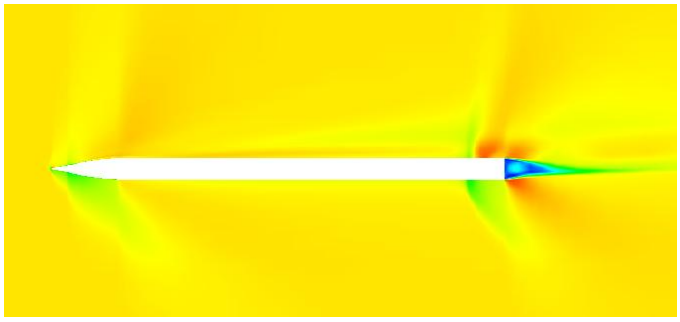
(a)



(b)



(c)



(d)

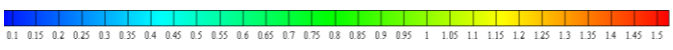
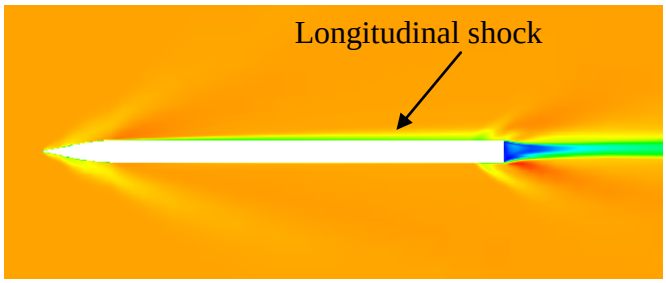
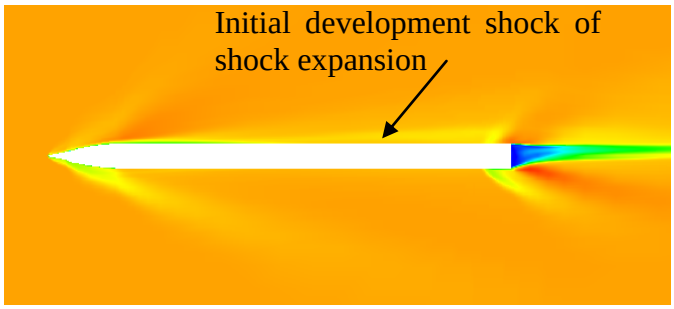


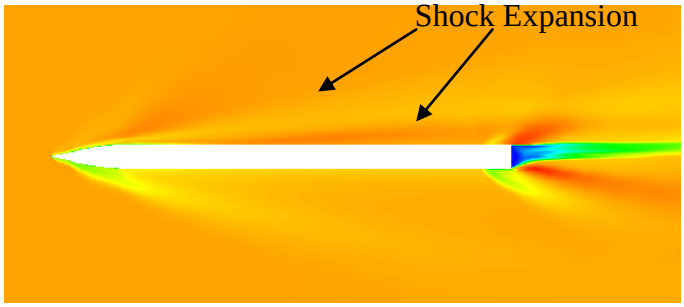
Fig. 4 The contour plot of subsonic flow (Mach number of 1.2) at angle of attack (a) 4°, (b) 8°, (c) 12°, and (d) 16° degrees



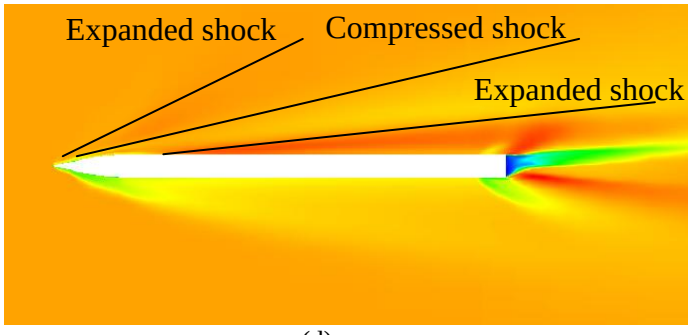
(a)



(b)



(c)



(d)

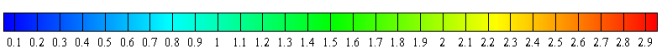
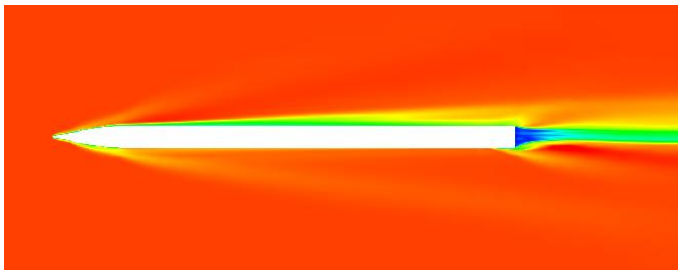
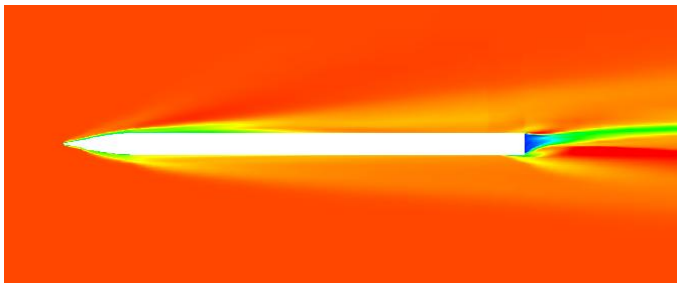


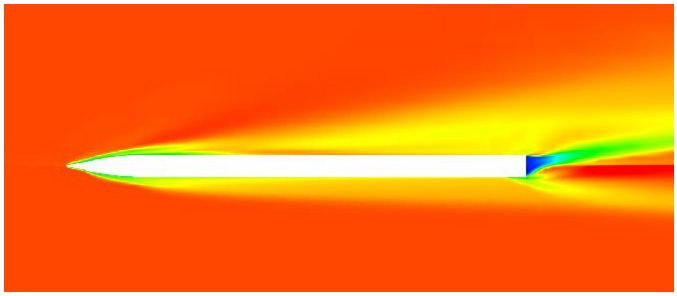
Fig. 5 The contour plot of supersonic flow (Mach number of 2.5) at angle of attack (a) 4°, (b) 8°, (c) 12°, and (d) 16° degrees



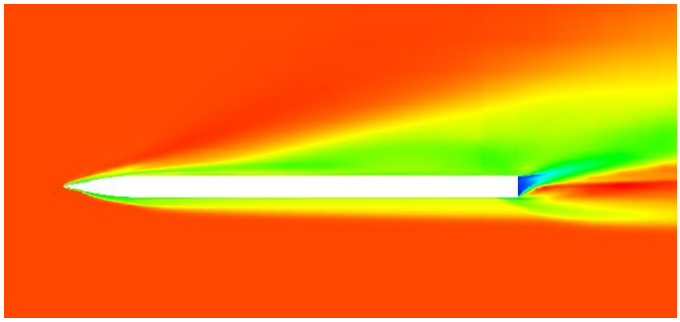
(a)



(b)



(c)



(d)

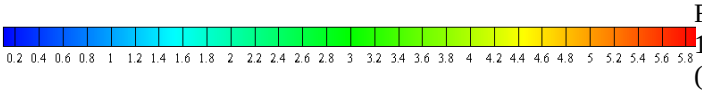
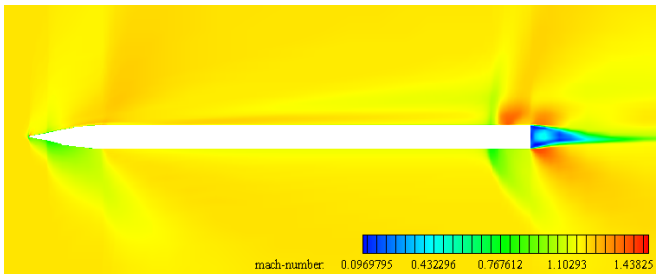
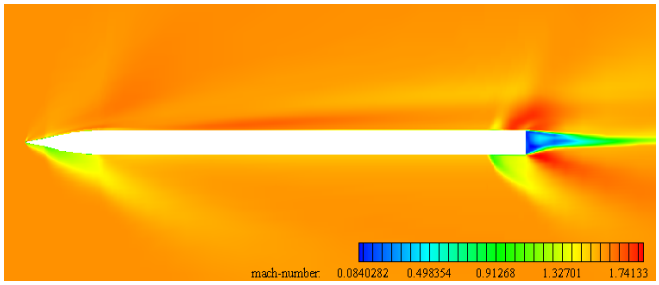


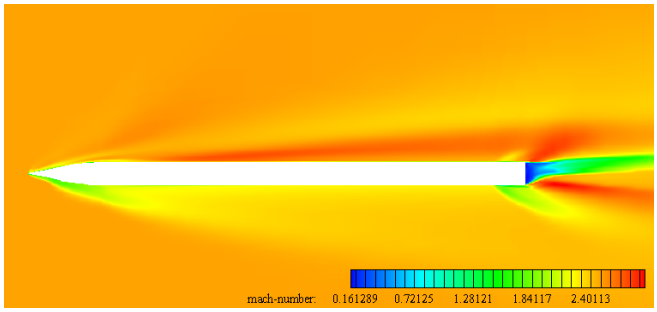
Fig. 6 The contour plot of hypersonic flow (Mach number of 5.5) at angle of attack (a) 4°, (b) 8°, (c) 12°, and (d) 16° degrees



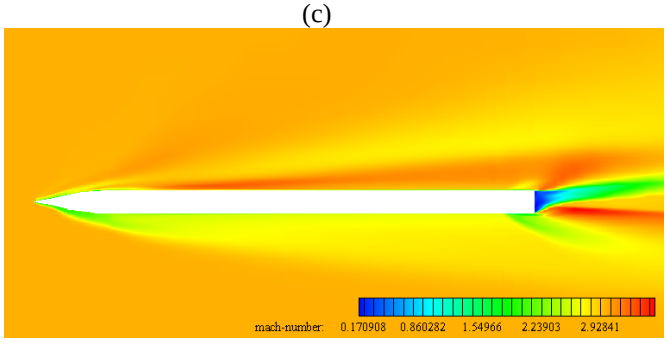
(a)



(b)



(c)



(d)

Fig. 7 The contour plot of Mach number at angle of attack 16° degrees for incoming Mach number of (a) 1.2, (b) 1.6, (c) 2.5, and (d) 3.0

The angle of attack introduces the pressure difference between upper and lower surfaces of the missile causing the flow of aero from the bottom to the top surfaces. The pressure driven flows are then roll up at the missile upper part and form as the streamwise vortices. The detailed vortices flow structures are quite complicated. Generally, on the upper of the missile, the two vertical flows were established. As the vortices advance downstream, both vortices gradually moves up, tending to move away from the missile surface. Ultimately, it reaches the fin region and confines between two adjacent fins. As they progress downstream, the pressure is recovered and their strengths are weaken, depicted in Figs. 8 to 10 and Figs. 11 to 13 for subsonic, supersonic, and hypersonic regimes at angles of attack 8° and 16° degrees, respectively.

As mentioned, the numerical results revealed the development of complicated vertical flow pattern and their interaction with fins at the rear part of the finned missile. For hypersonic flow at angle of attack 8, the vortices is firstly formed at x/L=0.16 as shown in Figure 10 (a), while demonstrated in Figures 8 and 9 (a) for subsonic and supersonic flows at location x/L 0.16, no vortices formation is observed. In addition, the development of secondary vortices is clearly observed and formed underneath the primary ones at x/L=0.5 and x/L=0.66 for Mach numbers 5.5 and 2.5, respectively. This indicates that the higher Mach number also promotes the earlier secondary vortices development. We note here another parameter promoting the development of secondary vortices is angle of attack. Shown Figs. 11, 12, and 13 (a) for angle of attack 16, the secondary vortices are observed at location x/L=0.33 for all flow regimes

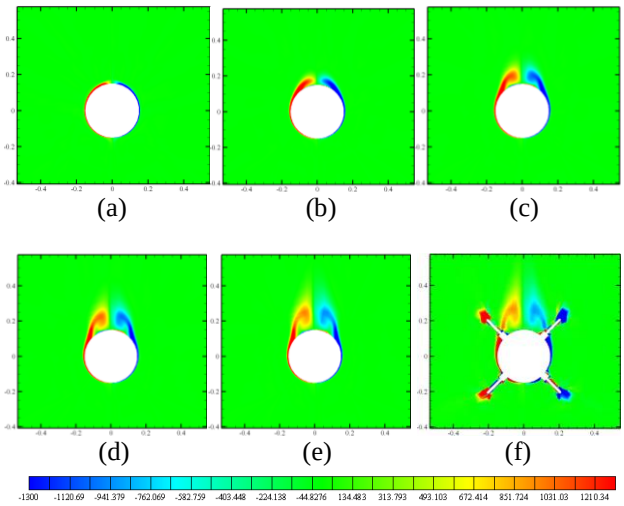


Fig. 8 The displays of x vorticity contour for angle of attack 8° degrees and Mach number 0.8 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99.

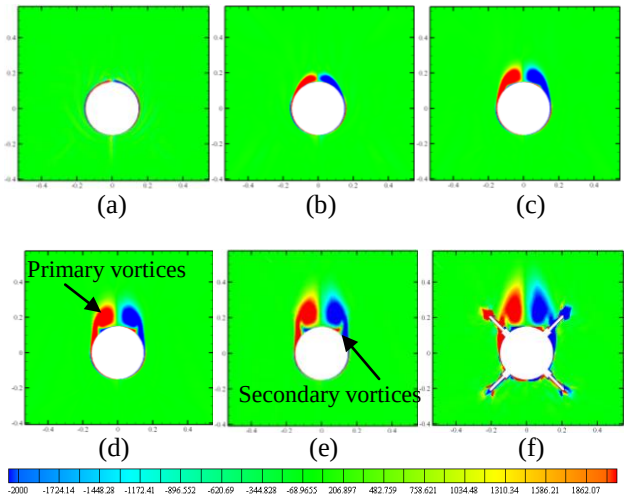


Fig. 9 The displays of x vorticity contour for angle of attack 8° degrees and Mach number 2.5 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99

Further observation for the vortical flow at 8° and 16° degrees of angle of attack, the visualization results revealed that, at 8° degrees of angle of attack, the vortices are symmetric, while asymmetric pattern of the primary vortices is clearly observed at 16° degrees of angle of attack in the fin region. In addition, the prolongation of the vortex flow is also observed as the vortices advance downstream. For subsonic and supersonic flows at angle of attack 8° degrees, the vortex stretching is clearly observed at the rear parts of missile (at x/L=0.83 and x/L=0.99) as shown in Figs. 8 and 9 (e) to

(f), while the prolongation and distortion of vortices patterns is observed in hypersonic regime (Fig. 10) . The complexity of vortices is increased at angle of attack 16° degrees, the strongly stretched and distorted vortical flows are found (Figs. 11 to 13), especially for hypersonic flow. The increasing complexity of vortex flow pattern results in the nonlinear behavior of aerodynamics coefficients.

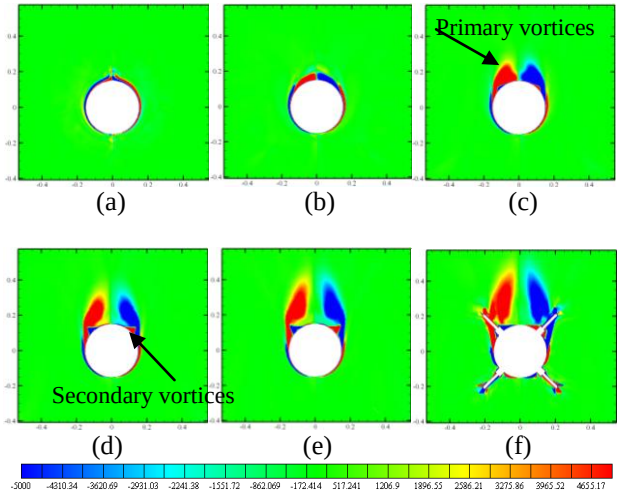


Fig. 10 The displays of x vorticity contour for angle of attack 8° degrees and Mach number 5.5 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99

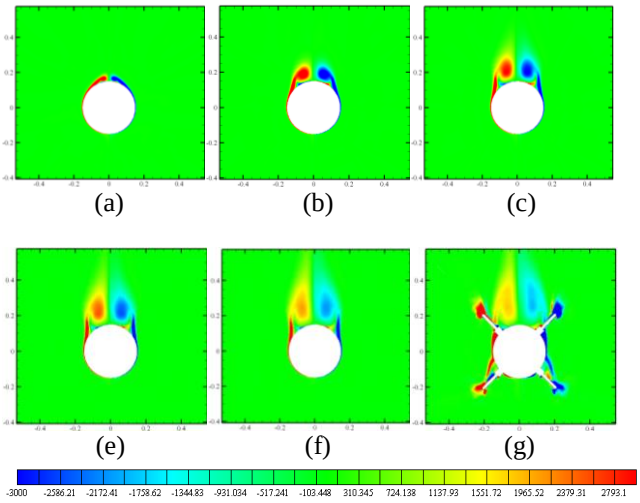


Fig. 11 The displays of x vorticity contour for angle of attack 16° degrees and Mach number 0.8 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99

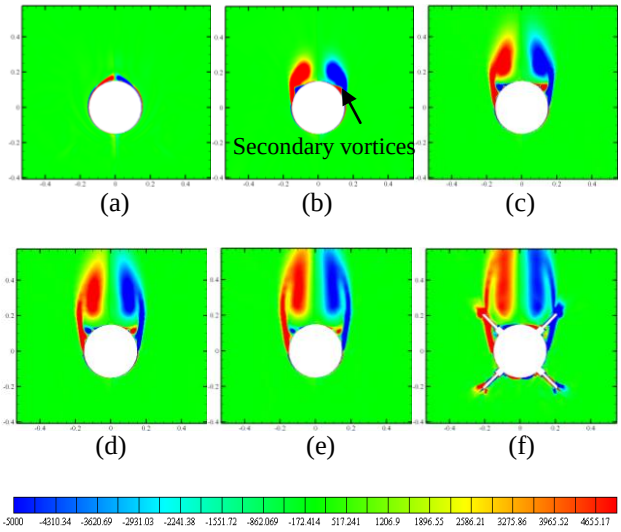


Fig. 12 The displays of x vorticity contour for angle of attack 16° degrees and Mach number 2.5 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99

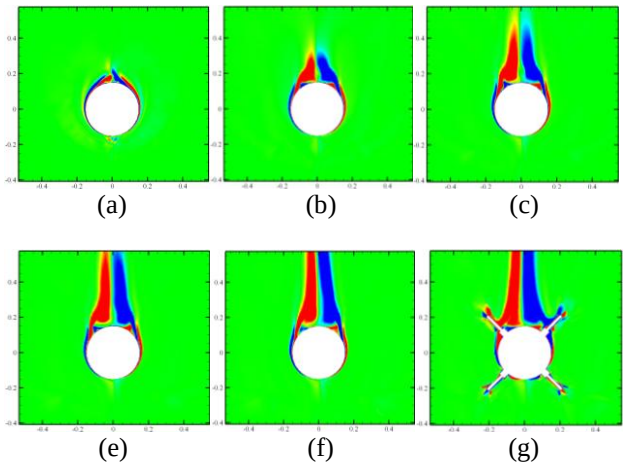


Fig. 13 The displays of x vorticity contour for angle of attack 16° degrees and Mach number 5.5 at x/L (a) 0.16, (b) x/L=0.33, (c) x/L=0.5, (c) x/L=0.66, (d) 0.83, and (e) 0.99

4.2 Static Aerodynamic Force and Moment Coefficients

With respect to semi empirical data generated by Missile DatCom of the static aerodynamic profiles including axial and normal force as well as pitching moment coefficients along the lines of Mach number and angle of attack, the predicted axial force and pitching moment coefficients by CFD are generally in marginal agreement; that is the reproducing trends similar to Missile DatCom as shown in

Figs. 14 and 16. However, the agreement between CFD simulation and Missile DatCom prediction is poorest especially for Mach number less than 3.5 in predicting the normal force coefficients (Fig. 16).

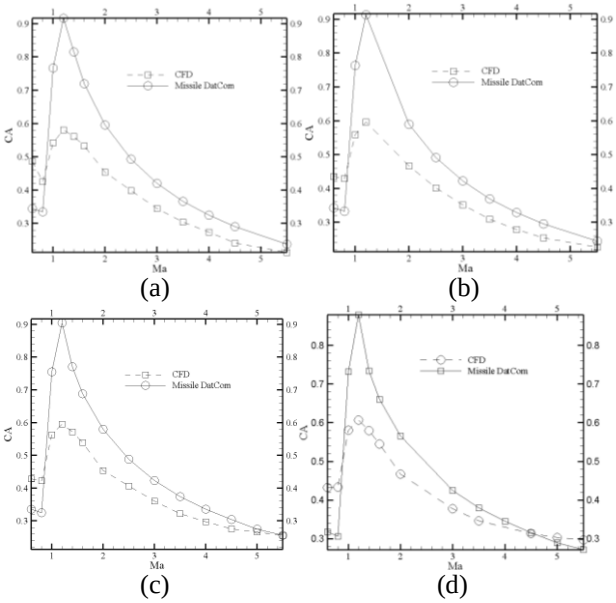


Fig. 14 The comparisons of aerodynamic axial force coefficient between CFD simulation and Missile DatCom prediction at AOA (a) 4, (b) 8, (c) 12, and (d) 16

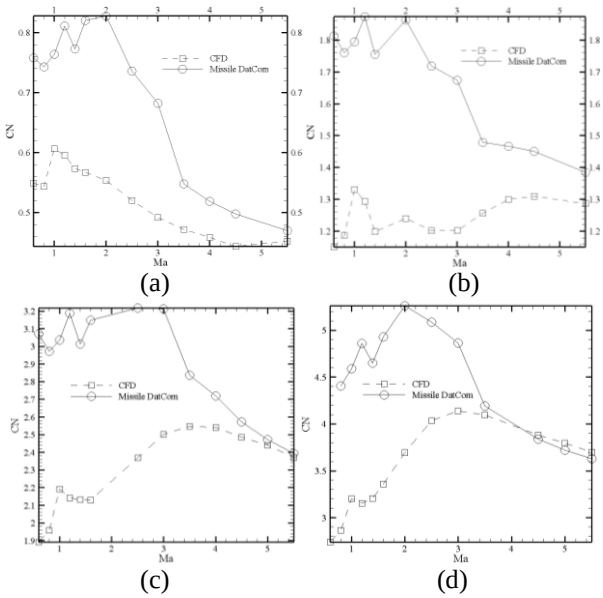


Fig. 15 The comparisons of aerodynamic normal force coefficient between CFD simulation and Missile DatCom prediction at AOA (a) 4, (b) 8, (c) 12, and (d) 16

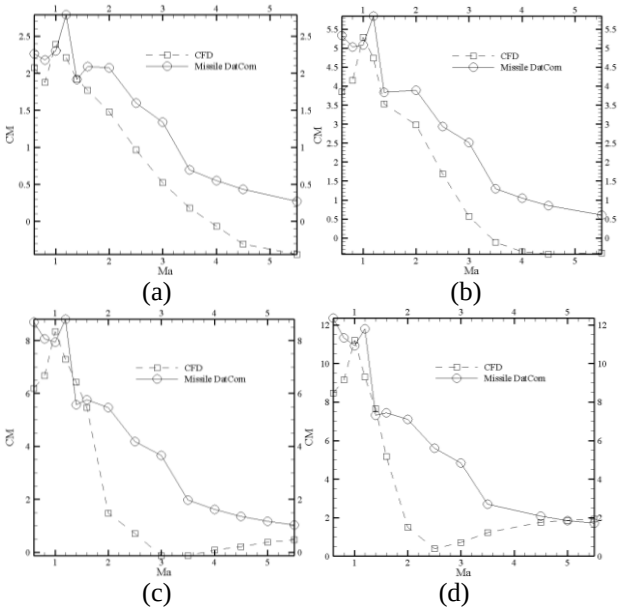


Fig. 16 The comparisons of aerodynamic normal force coefficient between CFD simulation and Missile DatCom prediction at AOA (a) 4, (b) 8, (c) 12, and (d) 16

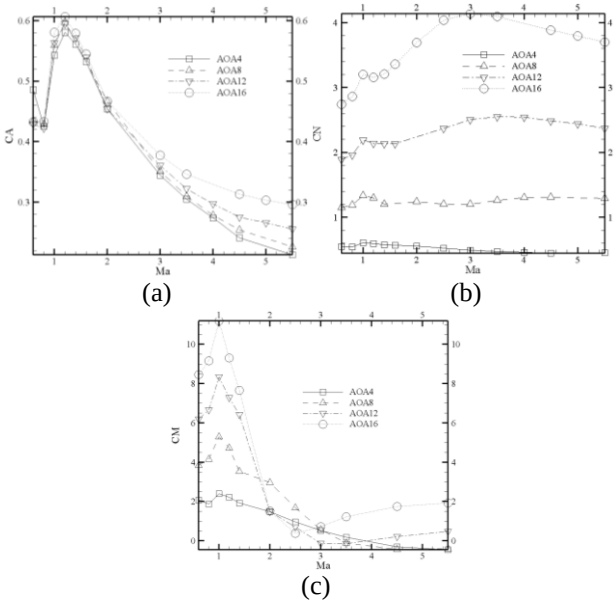


Fig. 17 The comparisons of (a) aerodynamic normal force coefficient, (b) aerodynamic axial force coefficient, and (a) aerodynamic pitching moment coefficient using CFD simulation at AOA 4, 8, 12, and 16

Figures 17 (a), (b), and (c) show the comparative plots of axial, normal, and pitching moment coefficients with respect to Mach number at angles of attack 4, 8, 12, and 16, respectively. The nonlinearity of the normal and pitching coefficients is dependent on angle of attack as shown in Figures 17 (b) and (c), while

plotted in Figure 17 (a), the trends of axial coefficient are similar for all angles of attack. We note that the axial, normal, and pitching coefficients are increased according to the increase of angle of attack. In addition for normal and pitching coefficients, the nonlinear behaviors are caused by the strong vortices at angle of attack 12 and 16. The minimum of pitching coefficients are located at Mach number 2.5 and 3.5 for angle of attack 16 and 12 degrees, respectively. In the case of Mach number 2.5, the minimal pitching moment coefficient is associated to the strongest surface shock expansion. Beyond the minimum location, the static margin is lost (the positive slope observed).

5 Conclusion

CFD simulations of the turbulent flow were conducted as part of an investigation to study flow through an finned missile configurations. Using FLUENT, we simulated compressible flow over the finned missile configurations at various angles of attack (4°-16°) and Mach numbers (0.6-5.5) using both Steady Reynolds Averaged Navier-Stokes (SRANS) with a first-order turbulence model, standard k-ε turbulence model.

Via standard k-ε SRANS, the development of shock flows and vortices were observed. The shock characteristics were represented by the abrupt change of flow properties. Here, we presented the shock flow pattern using the Mach number contour plot. The visualized results showed the development of bow shock at the missile nose and local shocks at geometrical transition region and finned area. Moreover, the so-called shock lift-up phenomena along the missile body was observed. This pattern exhibited by the expansion of shock flow from the missile surface. The vortices were characterized by vorticity contour. It was found that the strength and pattern of vortices are dependent on the compressible flow regime and the missile angle of attack.

Generally, the predicted axial and pitching moment coefficients using CFD were marginally agreed with the computed coefficients from Missile DatCom. The poorest agreement between CFD and missile DatCom was the normal force coefficient. The nonlinearity of normal and pitching coefficients was attributed to the effect of angle of attack.

References

[1] Bak K. M., "Experiment investigation and computational fluid dynamics analysis of missile with grid fin in subsonic flow", *International Journal of Engineering Science and technology*, Vol. 2, No. 11, 2010, pp. 6214-6220.

[2] Haroldsen D. J. and Sturek W. B., "Navier-stokes computations of finned missiles at subsonic speed," Army research laboratory, 2001.

[3] Champigny C., d'Espiney, P., Breduf M., Broussard H., Gillyboeuf J.-P., and Kergaravat Y., "Numerical Simulation of Vortex Flows around Missile Configurations," *the RTO AVT Symposium on Advanced Flow Management: Part A-Vortex Flow and high Angle for Military Vehicles*, Loen, Norway, 2000.

[4] Sahu J., Heavy K. and Dinavahi S., "Application of CFD to high angle of attack missile flow fields," Army research laboratory, 2000.

[5] Tuncer I. H., Platzer M. F. and VanDyken R.D., "Computational Investigations of Subsonic High Angle of Attack missile Flows," *the RTO AVT Symposium on Missile Aerodynamics*, Sorrento, Italy, 1998.

3rd CEAS Air&Space Conference
21st AIDAA Congress

LES Algorithm for Turbulent Reactive Flows Simulation

Ionut Porumbel, Cristian Carlanescu and Valentin Silivestru
National Research and Development Institute for Gas Turbines COMOTI, Romania

Keywords: *Numerical simulation, LES, reactive flow*

Abstract

The paper presents the development and implementation of a Large Eddy Simulation numerical algorithm for simulating turbulent reactive flows. The numerical algorithm is based on a 5 step modified Runge – Kutta numerical scheme with a dual time stepping numerical convergence acceleration technique, and employs a Localized Dynamic Turbulent Kinetic Energy turbulence model. The chemical reaction part of the algorithm is handled by a Linear Eddy Mixing combustion model. The paper described in brief the numerical algorithm, the turbulence model and the combustion model, highlighting the improvements they bring to the numerical solution in terms of accuracy and suitability to real life combustion problems. The algorithm validation is also performed, by comparison to experimental data and the literature and the superior algorithm accuracy is highlighted by comparison to numerical data resulting from previous numerical studies.

1 Introduction

1.1 Previous Work

With the recent advances in the computer technology, the numerical simulations of both reactive and nonreactive flow have become ever feasible and the use of numerical simulations for research studies is becoming ever more important. Reynolds Averaged Navier Stokes (RANS) simulations of premixed reactive flows were reported by many researchers in the field [1, 2, 3, 4, 5] with varying degrees of success. However, important discrepancies in the numerical results could not be avoided due to shortcomings stemming mostly from the RANS' inability to handle the smaller turbulent scales that are key to the combustion process.

The later development of the Large Eddy Simulation (LES) technique [6, 7, 8] allowed a significant improvement of numerical simulation accuracy in general. In LES, the three dimensional large scale motion is resolved, hence the geometry dependent flow features are captured accurately, and only the small scales that exhibit local isotropy are modeled.

Due to its ability to resolve a larger range of turbulent scales, the method is more appropriate to simulate turbulent flows. Numerical simulations of premixed reactive flows using this technique were reported in references [9-18].

1.2 Combustion and Turbulence

In most combustion related applications, the Reynolds number characteristic of the fluid flow in the flame region is sufficiently high such that the combustion process occurs in a turbulent flow field. The effects of the turbulence are generally advantageous for the efficiency of the combustion, since turbulence enhances the mixing of component chemical species and heat [19], but adverse effects upon combustion can also occur, if the turbulence level is sufficiently high to create flame extinction. In turn,

combustion may enhance the turbulence through dilatation and buoyancy effects caused by the heat release. The field of turbulent combustion is still an open research topic ("the most significant unresolved problem in classical physics" [19]) and significant research efforts are currently underway towards this end.

The most important theoretical issue that arises when considering turbulent combustion problems is the interdependency of the various time and length scales involved. Besides the diversity of scales brought into play by the turbulence, combustion occurs at molecular levels and involves a multitude of elementary chemical reactions, each with its own characteristic length scale.

If one is to consider simultaneously the entire range of scales involved, the problem becomes a lot more complex and also, when numerical simulation is considered, more computationally expensive. Therefore, a simplifying hypothesis was sought: the so – called hypothesis of "scale separation" [19]. The hypothesis assumes that in the inertial sub-range the scales of the combustion process are separated from the scales characteristic for the turbulence. However appealing through its simplicity, and popular for the variety of combustion models based on it, this hypothesis does not always hold true and accurate modeling of combustion processes occurring in a real application combustor requires more sophisticated combustion models that avoid this hypothesis. The combustion model used here, the Linear Eddy Mixing model is among the very few that maintain validity over the entire range of combustion regimes and, therefore, has proven superiority in terms of numerical accuracy when compared to classical combustion models, as employed by most commercial CFD software [20].

2 Mathematical Formulation

2.1 Governing Equations

The equations governing the motion of an unsteady, compressible, reacting, multiple-species fluid, are the Navier-Stokes equations.

The fully compressible Navier-Stokes equations describing the conservation of mass, momentum, total energy and conservation of N chemical species are:

$$\begin{cases} \frac{\partial \rho}{\partial t} + \frac{\partial \rho u_i}{\partial x_i} = 0 \\ \frac{\partial \rho u_i}{\partial t} + \frac{\partial}{\partial x_j}\left[\rho u_i u_j + p\delta_{ij} - \tau_{ji}\right] = 0 \\ \frac{\partial \rho E}{\partial t} + \frac{\partial}{\partial x_i}\left[\left(\rho E + p\right)u_i + q_i - u_j\tau_{ij}\right] = 0 \\ \frac{\partial \rho Y_m}{\partial t} + \frac{\partial}{\partial x_i}\left[\rho Y_m\left(u_i + V_{i,m}\right)\right] = \dot{w}_m \text{ ,where m = 1,N} \end{cases} \quad (1)$$

In the above equations, u_i is the *i*-th velocity component, ρ is the mass density, p is the pressure, Y_m is the species mass fraction of the *m*-th species, V_{im} is the diffusion velocity of the *m*-th species in the *i*-th direction,

$$E = e + \tfrac{1}{2}(u_k u_k) \quad (2)$$

is the total energy per unit mass, and τij is the viscous
stress tensor, defined as:

$$\tau_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{2}{3}\mu\frac{\partial u_k}{\partial x_k}\delta_{ij} \quad (3)$$

where δij is the Kronecker function.

Also, e is the internal energy per unit mass computed as:

$$e = \sum_{m=1}^{N} Y_m h_m - \frac{p}{\rho} \quad (4)$$

where h_m is the species enthalpy per unit mass given by:

$$h_m(T) = \Delta h^0_{f,m} + \int_{T_0}^{T} c_{P,m}(T')dT' \quad (5)$$

In the above, $\Delta h_{f,m}{}^0$ is the enthalpy of formation per unit mass of the *m*-th species at the reference temperature T_0, T is the temperature, and $c_{P,m}$ is the specific heat at constant pressure for the *m-th* species.

Returning to Eq. 1, ω_m is the mass reaction rate per unit volume of the *m-th* species.

The heat flux vector in Eq. 1 contains the thermal conduction (I), enthalpy diffusion (i.e. diffusion of heat due to species diffusion) (II), the Dufour heat flux and the radiation heat flux.

Dufour heat flux and radiation heat flux are neglected [20], therefore:

$$q_i = \underbrace{-\kappa \frac{\partial T}{\partial x_i}}_{(I)} + \underbrace{\rho \sum_{m=1}^{N} h_m Y_m V_{im}}_{(II)} \quad (6)$$

where the mixture averaged thermal conductivity is:

$$\kappa = \bar{c}_p \mu / Pr \quad)$$

the mixture averaged specific heat at constant pressure is:

$$\bar{c}_p = \sum_{m=1}^{N} Y_m c_{p_m} \quad (8)$$

and *Pr* is the mixture Prandtl number.

The pressure *p* is directly derived from the equation of state for perfect gas:

$$p = \rho RT = \rho \frac{R_u}{MW_{mix}} T = \rho R_u T \sum_{m=1}^{N} \frac{Y_m}{MW_m} \quad (9)$$

Fick's Law is used to determine the species diffusion velocity:

$$V_{im} = -\frac{D_m}{Y_m} \frac{\partial Y_m}{\partial xi} \quad (10)$$

where *Dm* is the *m-th* species molecular diffusion coefficient. Gradients of temperature and pressure can also produce species diffusion [20] (Soret and Dufour effects, respectively) but these two contributions are neglected hereafter.

2.2 LES Formulation

The earliest application of the Large Eddy Simulation (LES) methodology was performed by Smagorinsky [21] and important further developments of the method were introduced later [22, 23, 24].

LES resolves both the large, geometry dependent turbulent scales and a fraction of the smaller energy containing scales within the inertial range, up to a level dictated by the resolution of the numerical grid, and only the remaining scales are modeled. If the grid resolution is appropriately chosen, the unresolved scales, by Kolmogorov's hypothesis [25] are isotropic and, therefore, more amenable to modeling. This approach not only provides a lot more information but since the energy contained in the unresolved scales is much less than in the RANS approach, is also less sensitive to modeling hypotheses.

The separation between the large (resolved) and the small (modeled) scales is determined by the grid size Δ.Therefore, the Navier-Stokes equations have to be filtered with respect to the grid size in order to obtain the LES governing equations. A Favre spatial top-hat filter is employed to derive the LES equations. More details regarding the LES filtering and the different techniques are given by Ghosal [26] and Pope [27].

By applying the above mentioned filtering process to Eq. 1, the LES filtered Navier-Stokes equations can be written.

The LES filtered equations [20] contain unresolved terms representing the effects of the unresolved scales on the resolved motion, resulting from the filtering process and denoted by the superscript *sgs*.

Summarizing, the unclosed terms that need to be modeled are: the sub-grid kinetic energy, k^{sgs}, the shear stress tensor, τ_{ij}^{sgs} , the sub-grid enthalpy flux, H_i^{sgs}, the sub-grid viscous work, σ_i^{sgs} , the sub-grid convective mass flux, $\Phi_{j,m}^{sgs}$, the sub-grid diffusive mass flux, $\Theta_{j,m}^{sgs}$, the sub-grid heat flux, $q_{i,m}^{sgs}$, the sub-grid temperature - species correlation term, Y^{sgs} and the filtered reaction rate, $\widetilde{\omega}_m$

2.3 Turbulence model

The turbulence model, used to close the unresolved terms mentioned earlier, is a non - equilibrium model based on the sub-grid kinetic energy transport equation, initially developed by Schumann [28] and later improved by other researchers [29, 30, 31]. The model assumes isotropic turbulence at the sub-grid scales and tracks the sub-grid kinetic energy (k^{sgs}) using a formally derived transport equation, equation

solved along with the rest of the LES equations. Details on the model and the implementation can be found in reference [20]

The modeling coefficients are obtained dynamically as a part of the solution, using a method called the "Localized Dynamic Kinetic energy Model (LDKM) [32, 33].

2.4 Combustion model

To close the remaining terms mentioned earlier, a more comprehensive closure of the scalar mixing and combustion, based on the Linear Eddy Mixing (LEM) model proposed by Kerstein [34] and developed later into a sub-grid model [34], is used.

LEM is a stochastic approach aimed at simulating, rather than modeling the effects of turbulence on the chemistry, and it is not limited by the scale separation hypothesis, discussed earlier [19]. The parameters controlling the LEM turbulent mixing model require only the validity of the Reynolds number independence of free shear flows in the limit of large Reynolds numbers, which is a safe assumption for any flow of engineering interest [19]. Due to this extended validity range, the LEM model can be expected to perform well in any combustion regime, and to be able to accurately handle flames near to, or even outside, the flammability limits.

In LEM, the scalar equations are not filtered, and instead the large scale advection, turbulent mixing by eddies smaller than the grid size, molecular diffusion and chemical reaction are resolved at their appropriate length and time scales inside each LES cell. While the LES filtered conservation equations for mass, momentum and energy are numerically integrated on the LES grid, the evolution of the species fields is tracked using a two - scale, two - time numerical approach.

Molecular diffusion and chemical reaction contribution to the small scale transport are resolved on a one dimensional grid inside each LES cell at a resolution much finer than the LES resolution, and approaching the Kolmogorov scale. The one-dimensional computational domain is aligned in the direction of the flame normal inside each LES cell, ensuring an accurate representation of flame normal scalar gradients [34]. On this domain (denoted the LEM domain hereafter) molecular diffusion (term A below), chemical reactions (term B), diffusion of heat via species molecular diffusion (term C), heat diffusion (term D) and chemical reaction heat release (term E) are resolved, according to the equations below, written for the chemical species mass fractions (Eq. 11), respectively for the temperature (Eq. 12):

$$\rho^{LEM}\frac{\partial Y_m^{LEM}}{\partial t} + F_m^{stir} + \underbrace{\frac{\partial}{\partial s}\left(-\rho^{LEM}D_m\frac{\partial Y_m^{LEM}}{\partial s}\right)}_{A} = \underbrace{\dot{w}_m W_m}_{B} \quad (11)$$

$$\rho^{LEM}c_P\frac{\partial T^{LEM}}{\partial t} + F_T^{stir} \underbrace{-\sum_{k=1}^{N}\rho c_{p,m}D_m\left(\frac{\partial Y_m^{LEM}}{\partial s}\right)\left(\frac{\partial T^{LEM}}{\partial s}\right)}_{C} - \underbrace{\frac{\partial}{\partial s}\left(\overline{\kappa}\frac{\partial T^{LEM}}{\partial s}\right)}_{D} - \underbrace{-\sum_{m=1}^{N}h_m\dot{w}_m W_m}_{E} \quad (12)$$

Here, the superscript LEM indicates values at the subgrid LEM level, and s is the spatial coordinate along the LEM domain. F_m^{stir} and F_T^{stir} represent, respectively the effect of the sub-grid turbulence on the species m mass fraction field and on the temperature field.

The chemical reaction rates are given using a chemical reaction mechanism for the problem at hand. For the case of the numerical simulation presented herein, a 5-species, 1-step, reduced chemical mechanism for propane / air combustion was used. The chemical equation describing the mechanism is:

$$C_3H_8 + 5O_2 \Longleftrightarrow 3CO_2 + 4H_2O \quad (13)$$

with a reaction rate given by [35]:

$$\widetilde{\dot{\omega}_c} = \frac{\nu_m W_m}{\overline{\rho}}A_k exp\left(-\frac{E_a}{R_{u,cgs}\tilde{T}}\right)[C_3H_8]^{c_1}[O_2]^{c_2} \quad (14)$$

where ν_m is the stoechiometric coefficient, A_k is a pre-exponential factor equal to $8{,}6 \times 10^{11}$, E_a is the activation energy, equal to $3{,}0 \times 10^4$ calories/g, $R_{u,cgs}$ is the universal gas constant expressed in calories / gram Kelvin, $[X]$ represents the molar concentration of species X, in moles / cm^3, and c_1 and c_2 are two

coefficients, set to *0,1*, respectively *1,65* for this case [35].

The current implementation assumes a calorically perfect gas, and the sub-grid pressure, p^{LEM} is assumed constant over the LEM domain, and equal to the LES grid value, p, which is a valid assumption in the absence of strong pressure gradients [36].

Radiation effects are neglected. The small-scale turbulent stirring (F_m^{stir} and F_T^{stir}) is implemented explicitly on the same grid using stochastic rearrangement events that mimic the action of an eddy upon the scalar field using a method known as triplet mapping and designed to recover the 3D inertial range scaling laws [37]. Details on the triple mapping method are given elsewhere [20], and its effect on the scalar field is illustrated in Fig. 1.

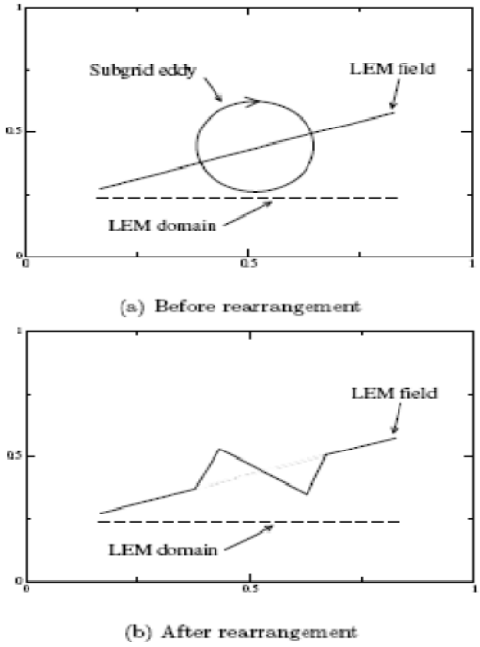


Fig. 1 - Schematic representation of the triplet mapping

The location of this stirring event is chosen from a uniform distribution. The frequency at which stirring events occur is given by [34]:

$$\lambda = \frac{54}{5}\frac{\nu Re_{\overline{\Delta}}}{C_\lambda \overline{\Delta}^3}\frac{[(\overline{\Delta}/\eta)^{5/3}-1]}{[1-(\eta/\overline{\Delta})^{4/3}]} \qquad (15)$$

where C_λ stands for the scalar turbulent diffusivity, set to *0.0675* [38].

The eddy size, l, ranges from the Kolmogorov scale, η, to the grid size, $\bar{\Delta}$ with a distribution given by [34]:

$$f(l) = \frac{(5/3)l^{-8/3}}{\eta^{-5/3} - \overline{\Delta}^{-5/3}} \qquad (16)$$

where the Kolmogorov scale is determined as:

$$\eta = N_\eta \overline{\Delta} Re_{\overline{\Delta}}^{-4/3} \qquad (17)$$

and N_η is an empirical constant that reduces the effective range of scales between the integral length scale and η but without altering the turbulent diffusivity [37]. The value used for this study is *5* [37].

The large scale processes are modeled using a Lagrangian transport of the scalar field across the LES cells that ensures exact mass conservation and called splicing [36]. Thus, once the LES computations are completed at a given time step, LEM domain cells (and / or cell fractions) are exchanged between the LES cells in a manner that accounts for the mass fluxes across the LES cell faces.

Thus, LEM cells are transferred between the LES volumes accounting for the mass fluxes through the LES cell faces. The order in which the cell transfer on each of the spatial directions is performed is dictated by the magnitude of the mass flux in the respective direction at the resolved level [39, 40]. Next, the number of LEM cells containing the mass flux to be transported to the adjacent cell is determined. If a fractional number is obtained, the LEM cell is split so that exact mass conservation is achieved.

The sufficiently small LES time step ensures that scalars are transported from one LES cell only to an adjacent LES cell, thus drastically reducing the complexity of the problem.

Since the pressure in the LEM domain is assumed constant and there is no pressure gradient term in Eq. 11 or Eq. 12, the volumetric expansion of the LEM cell needs to be modeled separately, after each diffusion step,

to account for the increase in volume through thermally generated pressure waves.

Both the thermal expansion and the splicing procedure cause the LEM linear grids to be neither uniform, nor have the same number of cells in different LES cells. However, the triplet mapping procedure and the discretization method used to integrate the reaction diffusion equation require an uniform grid. Hence, a re-gridding procedure aimed at producing uniform grids with equal numbers of cells is applied [40].

3 Validation Results and Discussion

3.1 Numerical setup

The validation numerical simulations presented here where performed in a geometry that reproduces a previous experimental study the Volvo experiment [41] and consists of a rectangular duct of size *1.0 m x 0.24 m x 0.12 m* with a triangular prism that extends between the two lateral walls of the combustor, as shown in Fig. 2. The side of the bluff body triangular base, *a*, measures *0.04* m.

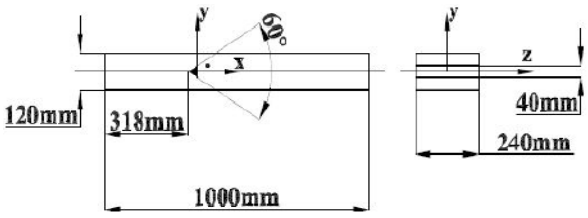


Fig. 2 - Schematic of the validation geometry

The computational domain was divided into 10 blocks, and each block was discretized by a body fitting grid with dimensions in Table 1.

The computational grid is stretched both axially and transversally and provides the maximum resolution in the two separated shear layers, immediately downstream of the bluff body, resolved by about 20 grid points. The grid stretching is maintained under 5 percent for reasons of numerical stability.

For the LEMLES, 12 LEM cells are used in each LES cell. Using the predicted k^{sgs} and the local $\bar{\Delta}$, the maximum local sub-grid Re_{Δ} is *130* and $\eta = 25 \times 10^{-6}$ m.

Block	i	j	k
1	45	52	90
2	35	52	90
3	278	52	90
4	90	52	90
5	278	81	90
6	90	81	90
7	45	52	90
8	35	52	90
9	278	52	90
10	90	52	90

Table 1 - Computational grid dimensions

The inflow velocity is *17.3* m/s with a *2* percent turbulence intensity under standard atmospheric conditions. The reference Reynolds number based on inflow velocity and bluff body height is *45,500*. The inflow consists of a propane - air mixture of *0.65* equivalence ratio. The simulations are carried out for five flow-through times before the flow statistics are collected, and the time averaged data presented herein are collected over a period equal to five flow-through times.

3.2 Time averaged results

This section will presents the time averaged results of the validation numerical simulation denoted LEMLES, compared to experimental data in the literature [41].

Results of an earlier numerical simulation of the same geometry using an Eddy Break - Up [42] combustion model are also included, for comparison, and denoted EBU.

Figure 3 presents the normalized axial velocity profile along the combustor centerline, behind the bluff body.

The numerical result matches closely the experimental data. Immediately downstream of the bluff body, the velocity is negative and reaches a negative maximum of about *0.75* of the inflow velocity at about twice the size of the bluff body, in agreement with the experimental data The length of the recirculation region is about *3.75 a*. After the end of the reverse flow zone, the mean axial velocity increases upstream to gradually approach an outflow

velocity about three times larger than the inflow velocity, due to the addition of chemical energy through combustion.

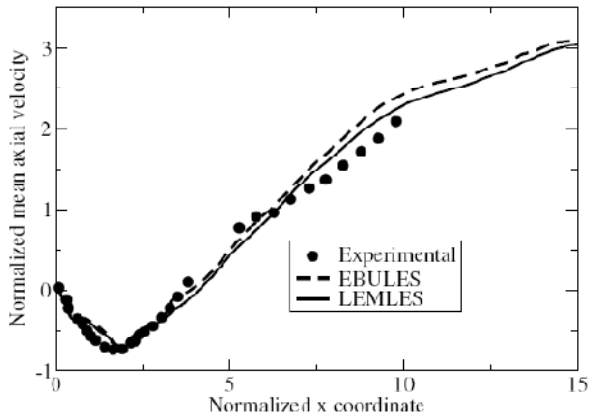


Fig. 3 - Centerline variation of the normalized time averaged axial velocity. The velocity is normalized by the inflow value, and the distance is normalized by the bluff body size, *a*.

Both numerical studies using two different combustion models yield equally accurate results, although the current study appears to show slightly better agreement, which is due to a better temperature prediction in the far field, as it will be shown later.

In the transverse direction, the accuracy of the current numerical simulation results is also acceptable, as seen in Fig. 4. In addition to the experimental and numerical results mentioned earlier, a third set of numerical data pertaining to another previous study [18], marked by letter G in Fig. 4 and using an Eddy Dissipation Model (EDM) is also shown for comparison, The numerical accuracy of the current study is not significantly improved in this case when compared to the earlier work, except that the current study predicts a more accurate axial velocity in the free stream, correlating well with the improved temperature predictions, to be shown later. Also, far downstream, the centerline velocity becomes over-predicted by the EBU models, indicating that the predicted acceleration rate is slightly off.

In the case of the transverse component of the time averaged velocity, the agreement with the experimental [41] is generally good, as shown in Fig. 5.

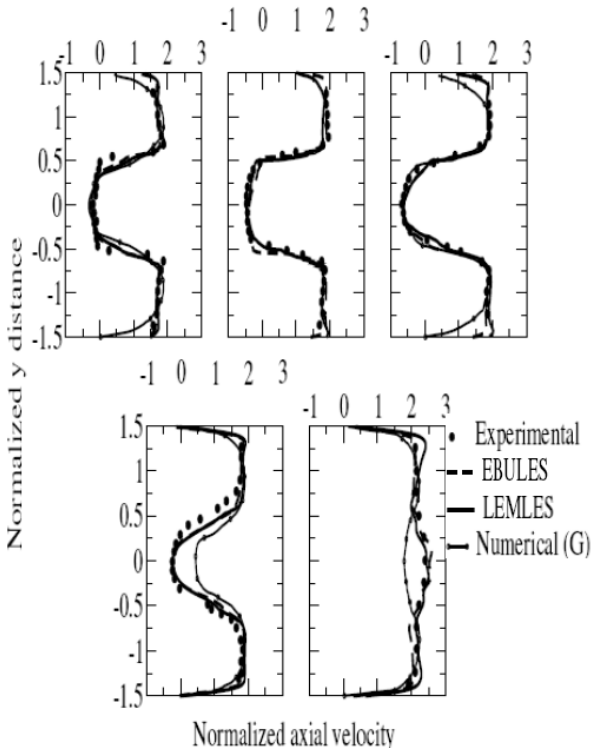


Fig. 4 – Transverse profiles of the normalized time averaged axial velocity, at the normalized axial locations, from left to right: *0.375 a, 0.95 a, 1.53 a, 3.75 a* and *9.4 a*. The velocity is normalized by the inflow value, and the distance is normalized by the bluff body size, *a*.

The large scale vortical structures, created by the presence of the bluff body, decrease in intensity with the distance from the obstruction, as the wake momentum deficit diminishes and the flow tends to recover its initial axial direction. The maximum transverse velocity is about *30%* of the inflow velocity, and is achieved at about *0.4 a*. A notable feature of the flow field is the sudden decrease of the transverse velocity at the flame front. This behavior correlates with the pair of stationary, counter-rotating vortices that form at the sharp edge of the bluff body in the case of a reacting flow around it. The effect disappears further downstream, where the intensity and the coherence of the vortices weaken.

Overall, even if from a validation standpoint the numerical results obtained by using the algorithm presented herein are excellent, these results appear to suggest only modest improvements when using the LEM combustion model. However, the improvements do appear

in regions where mixing between the products and the reactants is occurring and where the flame structure exists. This subtle fact becomes clearer when the time-averaged temperature profiles are compared in Fig. 6. As earlier, the numerical data reported by Giacomazzi et al. [18] is shown for comparison It can be seen that the earlier EBU study under-predicts the mean temperature in the centerline region. As a consequence, the flame thickness is overestimated by EBU as well [20] and the EBU flame is slower to respond to the turbulent fluctuations, and the intermittency effect is not captured accurately in this case. As a consequence, the turbulent flame brush is not captured accurately and a smaller spreading rate in the transverse direction is predicted by EBU, error that becomes more prominent further downstream.

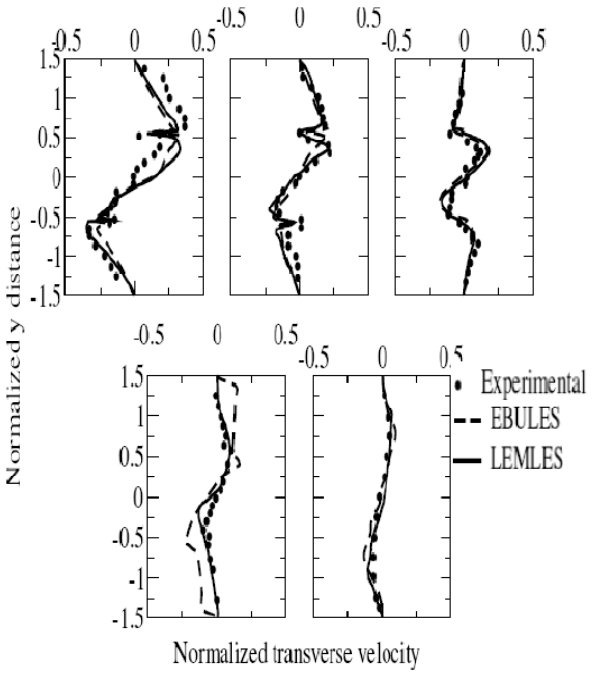


Fig. 5 – Transverse profiles of the normalized time averaged transverse velocity, at the normalized axial locations, from left to right: *0.375 a, 0.95 a, 1.53 a,* 3.*75 a* and *9.4 a*. The velocity is normalized by the inflow value, and the distance is normalized by the bluff body size, *a*.

Also, EBU tends to under-predict the centerline values by as much as 10%. It can be noted that the centerline region is also the region of low turbulent kinetic energy and EBU will predict here a reduced turbulent mixing rate. However, the experimental data show that the temperature maintains its high value over a large portion of the domain, so even with a reduced mixing rate the premixed reactants entrained in this region should burn at a high rate. The LEM model, on the other hand, avoids estimating the controlling rate and simulates the involved processes, thus allowing for a more accurate prediction of the temperature. The more accurate prediction of heat release results in more accurate spreading rates, which is reflected in the velocity field, as discussed earlier.

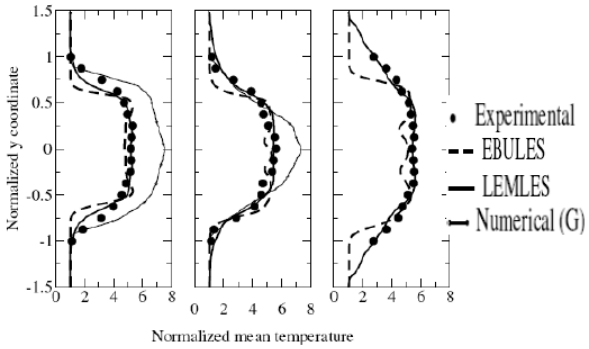


Fig. 6 – Transverse profiles of the normalized rms of the axial velocity fluctuation, at the normalized axial locations, from left to right: *0.375 a, 0.95 a, 1.53 a, 3.75 a* and *9.4 a*. The velocity is normalized by the inflow value, and the distance is normalized by the bluff body size, *a*.

These results demonstrate the subtle and global effects of using a more comprehensive combustion and mixing model as LEM. The significantly larger errors yielded by the EDM numerical study [18] are probably due to insufficient grid resolution.

4 Conclusion

The main goal of this work was to develop and implement an efficient, highly scalable computational algorithm able to accurately capture the subtle features of a highly turbulent reactive flow. The best suited CFD method for this is a compressible LES approach.

Although numerous such algorithms exist and have been successfully used in the past, this is, to the knowledge of the authors, the first such

complex implementation on Romania and will be used by Romanian researchers to tackle complex turbulent combustion problems with state-of-the-art methods and computational tools.

The LES algorithm was validated by simulating a reactive flow behind a bluff body and by comparing the results against earlier experimental [41] and numerical data [18, 42]. The simulation results have shown the superiority of the LEM closure. On theoretical grounds, this was to be expected, since LEMLES needs no assumption about the decoupling between the turbulent scales and the combustion scales in the inertial range (the so-called "scale separation hypothesis" and, therefore, maintains its validity over the entire range of combustion regimes. Previous numerical simulations of similar geometries [2, 9, 18, 42, 43] have all failed to capture the correct experimentally measured spreading rate. A common feature of these studies was the laminar chemistry assumption embedded in the combustion models (Eddy Dissipation Model [18, 43], or EBU [2, 9, 42]). As the simulated flame was proved to extend out of the validity range of the above mentioned combustion models, into the broken reaction zone regime [20], the flame thickness predicted by such models was overpredicted rendering the flame front less susceptible to turbulent fluctuations in the flow. Hence, flame surface wrinkling is significantly more pronounced in LEMLES, the flame structure is more complex, and the far field spreading of the wake is closer to the experimental observations. The time-averaged temperature in the low turbulence centerline region is underestimated by EBU, due to under-predicted turbulent mixing rates in the region. Consequently, the velocity fluctuations are also better predicted by the current study using LEM.

The current LEMLES implementation was proven to yield accurate results and is, therefore, suited for in-depth numerical studies of turbulent combustion. In combination with the LDKM model, the algorithm presented here has no adjustable ad-hoc parameters and is able to handle the entire range of combustion regimes and equivalence ratios.

References

[1] X. Bai, L. Fuchs, "Modeling of turbulent reactive flows past a bluff body: Assessment of accuracy and efficiency", *Computers and Fluids*, Vol. 23, No. 3, 1994, pp. 507-521.

[2] C. Fureby, S. I. Moller, "Large eddy simulations of chemically reactive flows applied to bluff body stabilized flames", *AIAA Journal*, Vol. 33, No. 12, 1995, pp. 2339-2347.

[3] M. Saghafian, P. K. Stansby, M. S. Saidi, D. D. Apsley, "Simulation of turbulent flows around a circular cylinder using nonlinear eddy viscosity modeling: Steady and oscillatory ambient flows", *Journal of Fluids and Structures*, Vol 17, No. 8, 2003, pp.1213-1236.

[4] F. Lien, E. Yee, "Numerical modeling of the turbulent flow developing within and over a 3-d building array. part i: A high-resolution Reynolds-averaged Navier-Stokes approach", *Boundary-Layer Meteorology*, Vol. 112, No. 3, 2004, pp. 427-466.

[5] A. Frendi, G. Skarath, A. Tosh, "Prediction of noise radiated by flow over a smooth square cylinder", *10th AIAA/CEAS Aeroacoustics Conference*. Manchester, UK, 2004.

[6] J. Smagorinsky, General circulation experiments with the primitive equations. I. The basic experiment", *Monthly Weather Review*, Vol. 91, No. 3, 1963, pp. 99-164.

[7] D. K. Lilly, "The representation of small-scale turbulence in numerical simulation experiments", *Proceedings of the IBM Scientific Symposium on Environmental Sciences*, IBM. Yorktown Heights, NY, 1967, pp. 195-210

[8] J. W. Deardorff, "Three-dimensional numerical study of the height and mean structure of a heated planetary boundary layer", *Boundary-Layer Meteorology*, Vol. 7, 1974, pp.81-106.

[9] C. Fureby, C. Lofstrom, "Large-eddy simulations of bluff - body stabilized flames", *Twenty- Fifth Symposium (International) on Combustion*, 1994, pp. 1257-1264.

[10] C. Fureby, "Large-eddy simulations of turbulent anisochoric flows", *AIAA Journal*, Vol. 33, No. 7, 1995, pp. 1263-1272.

[11] C. Fureby, "On sub grid scale modeling in large eddy simulations of compressible fluid flow", *Physics of Fluids*, Vol. 8, No. 5, 1996, pp. 1301-1311.

[12] S. Moller, E. Lungren, C. Fureby, "Large-eddy simulations of unsteady combustion", *Twenty - Fifth Symposium (International) on Combustion*, 1996, pp. 1257-1264.

[13] C. Fureby, "Large eddy simulation of

combustion instabilities in a jet engine afterburner model", *Combustion Science and Technology*, Vol. 161, 2000, pp. 213-244.
[14] C. Fureby, "A computational study of combustion instabilities due to vortex shedding", *Proceedings of the Combustion Institute*, Vol. 28, 2000, pp. 783-791.
[15] R. Ryden, L. Eriksson, S. Olovsson, "Large eddy simulation of bluff body stabilized turbulent premixed flames", *International Gas Turbine and Aero - Engine Congress and Exposition*, ASME - 93 - GT - 157, 1993.
[16] S. Inage, V. Perez, N. Kobayashi, "An evaluation of premixed combustion around a bluff body using a combustion model, *JSME International Journal*. Series B, Vol. 41, No. 3, 1998, pp. 650-656.
[17] V. K. Chakravarthy, "Stochastic Subgrid Modeling of Turbulent Premixed Flames", PhD Dissertation, Georgia Institute of Technology, Atlanta, Georgia, USA, 2000.
[18] E. Giacomazzi, V. Battaglia, C. Bruno, "The coupling of turbulence and chemistry in a premixed bluff-body flame as studied by LES", *Combustion and Flame*, Vol. 138, 2004, pp. 320-335.
[19] N. Peters, *Turbulent Combustion*, Cambridge University Press, Cambridge, U.K., 2000.
[20] I. Porumbel, *LES of Bluff Body Stabilized Premixed and Partially Premixed Combustion,* VDM Verlag Dr. Muller, Saarbrücken Germany, 2006
[21] M. Germano, P. Piomelli, U. Moin, W. H. Cabot, "A dynamic sub-grid scale eddy viscosity model", *Proceedings of the Summer Program*, 1990, pp 5-17, 1990.
[22] P. Moin, W. Squires, W. Cabot, S. Lee, "A dynamic sub-grid scale model for compressible turbulence and scalar transport, *Journal of Fluid Mechanics*, Vol. 3, No. 11, 1991, pp. 2746-2757.
[23] G. Erlebacher, M. Y. Hussaini, C. G. Speziale, T. A. Zang, "Toward the large eddy simulation of compressible turbulent flows", *Journal of Fluid Mechanics*, Vol. 238, 1992, pp. 155-185.
[24] S. Menon, "Active combustion control in a ramjet using large eddy simulations", *Combustion Science and Technology*, Vol. 84, 1992, pp. 51-79.
[25] A. N. Kolmogorov, "Local structure of turbulence in incompressible fluids with very high Reynolds number", *Dokl. Acad. Nauk*, Vol. 30, 1941, pp. 229.
[26] S. Ghosal, "On the large-eddy simulation of turbulent flows in complex geometries", Technical report, Center for Turbulence Research, 1993.
[27] S. B. Pope, *Turbulent Flows*, Cambridge University Press, 2000.
[28] U. Schumann, "Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli", *Journal of Computational Physics*, Vol. 18, 1975, pp. 376-404.
[29] W. W. Kim, S. Menon, "A new incompressible solver for large - eddy simulations", *International Journal of Numerical Methods in Fluids*, Vol. 31, 1999, pp. 983- 1017.
[30] W. W. Kim, S. Menon, H. C. Mongia, "Large eddy simulation of a gas turbine combustor flow", *Combustion Science and Technology*, Vol. 143, 1999, pp. 25-62.
[31] S. Menon, P. K. Yeung, W. W. Kim, "Effect of subgrid models on the computed interscale energy transfer in isotropic turbulence", *Computers and Fluids*, Vol. 25, No. 2, 1996, pp. 165-180.
[32] W.-W. Kim, S. Menon, "A new dynamic one equation subgrid-scale model for large-eddy simulations", *AIAA-95-0356*, 1995.
[33] C. C. Nelson, S. Menon, "Unsteady simulations of compressible spatial mixing layers", *AIAA-98-0786*, 1998.
[34] A. R. Kerstein, "Linear-eddy modeling of turbulent transport ii", *Combustion and Flame*, Vol. 75, 1989, pp. 397-413.
[35] C. K. Westbrook, F. L. Dryer, "Simplified reaction mechanisms for the oxidation of hydrocarbon fuels in flames", *Combustion Science and Technology*, Vol. 27, 1981, pp. 31-43.
[36] V. Sankaran, I. Porumbel, S. Menon, "Large-eddy simulation of a single-cup gas turbine combustor", *AIAA-2003-5083*, 2003.
[37] T. M. Smith, S. Menon, "One-dimensional simulations of freely propagating turbulent premixed flames", *Combustion Science and Technology*, Vol. 128, 1996, pp. :99-130.
[38] V. K. Chakravarthy, S. Menon, "Large - eddy simulations of turbulent premixed flames in the flamelet regime", *Combustion Science and Technology*, Vol. 162, pp. 1-48.
[39] V. Chakravarthy, S. Menon, Subgrid modeling of premixed flames in the flamelet regime, *Flow, Turbulence and Combustion*, 2001.
[40] V. Sankaran, "Subgrid Combustion Modelling for Compressible Two - Phase Reacting Flows, PhD Dissertation, Georgia Institute of Technology, Atlanta, Georgia, USA, 2002.
[41] A. Sjunesson, S. Olovsson, B. Sjoblom, "Validation rig - a tool for flame studies", Technical Report 9370-3088, 1991.
[42] I. Porumbel, S. Menon, "Large-eddy simulation of bluff body stabilized flames", *AIAA-2006-0152*, 2006.
[43] A. Sjunesson, R. Henriksson, C. Lofstrom, "CARS measurements and visualization of reacting flows in bluff body stabilized flame", AIAA - 92 - 3650, 1992.

CEAS 2011

All Electric Aircraft

RAPID200-FC - Flight Test of First All Electric Propulsion Aeroplane powered by Fuel Cells

G. Romeo , F. Borello , E. Cestino and G. Correa
Politecnico di Torino, Dept. of Aerospace Engineering, giulio.romeo@polito.it

Keywords: *Hydrogen, Fuel Cell, Green Aircraft, Flight Tests*

1 Abstract

The paper contains a description of testing activities related to experimental flights of an all-electric general aviation aircraft fuelled by hydrogen. A general description of the components of the highly complex power system are first presented. Great importance has been given to the testing phase of the prototype and examples of each testing stage are shown ranging from the single components to the final test flights. During the 6 experimental flights a rotation speed of 84 km/h was obtained in 184 m of taxi at power of 35 kW. Level flight was attained at 135 km/h by mean of only a fuel cell power setting. A new speed world record of 135 km/h and endurance of 39 min. were established during several flights conducted for the FAI Sporting Code Category C (airplane).

2 General Introduction

The hydrogen and fuel cell power based technologies that are rapidly emerging can now be exploited to initiate a new era of propulsion systems for light aircraft and small commuter aircraft. These technologies can also be developed for the future replacement of on-board electrical systems in larger 'more-electric' or 'all-electric' aircraft. The main objective of the ENFICA-FC project (**ENvironmentally Friendly Inter City Aircraft powered by Fuel Cells** - European Commission funded project coordinated by Prof. Giulio Romeo) was to develop and validate the use of a fuel cell based power system for the propulsion of more/all electric aircraft. The fuel-cell system was installed in a light sport aircraft RAPID 200 which was flight and performance tested as proof of the functionality and future applicability of this system for inter city aircraft.

The ENFICA-FC consortium consists of 9 partners representing the whole chain of Aircraft manufacturers (IAI, Evektor and Jihlavan Airplanes), a Fuel cells Power system producer (Intelligent Energy), a Hydrogen distributer (Air Product), Research Institutes (the Politecnico di Torino, the Université Libre de Bruxelles and the University of Pisa) as well as an SME in the field of administrative management (Metec). Within the course of the ENFICA-FC project, which was launched in October 2006, two key objectives were attained:

1) A feasibility study was carried out to provide a preliminary definition of new forms of Inter-City aircraft power systems that could be provided by fuel cell technologies (APU, Primary electrical generation supply, Emergency electrical power supply, Landing gear, De-icing system, etc); the safety, certification and maintenance concepts were also defined. Parametric sizing of different aircraft categories was performed; these ranged from two-seater aircraft to small 32 passenger commuters. Very interesting results were obtained from the preliminary parametric sizing and analysis of a more-electric 32 passenger regional jet aircraft fuelled by liquid hydrogen. The study has led to a better understanding of the practical meaning of transition from kerosene to hydrogen in transportation airplanes.

2) A two-seater electric-motor-driven airplane, powered by fuel cells, was assembled and

tested. The high efficiency, two-seater aircraft Rapid 200, manufactured by Jihlavan Aircraft (now Sky Leader), was selected for conversion from over more than one hundred light sport aircraft through a multi-criteria analysis performed by the Politecnico di Torino. After the selection of the aircraft for conversion [1] an extensive conversion design activity was undertaken. The first step was to define the demonstrative mission; a complete, but limited, mission profile was selected by the consortium since the demonstration goal was to show the feasibility of a new concept propulsion system. The requested mission performances were based on a parametric study and system architecture design which was conducted by the authors and is reported in [1]; mission properties were chosen because they could guarantee that the mission could be flown while keeping the total weight at around 550 kg, i.e the maximum total weight at which original RAPID200 was tested. This result was confirmed at the end of the conversion activity. A better understanding of the aerodynamic behavior of the aircraft was needed to define the power requested for the mission phases; a CFD analysis was performed, for this purpose, by the mean of the VSAERO commercial code; the analysis concerned not only the overall aircraft, but also the critical components that had to be designed; as an example, the CFD results were vital for the design of the engine cowl [2] which must guarantee a proper cooling of the different systems installed in the engine bay (see section 3), but also be a passive safety system for the prevention of hydrogen accumulation. Moreover the CFD results were important to design the complete hydrogen venting system and to predict the best locations for the static pressure ports.

3 Power System Description

A particular architecture was adopted for the power system in order to obtain an aircraft that could fly the prescribed mission. Relying solely on fuel cells for the entire mission, including take-off, leads to an excessive weight due to the required large fuel cell system at high power (40 kW); for this reason a hybrid battery/fuel cell system was chosen (see the following sections for a complete description of the sub-systems).

It was decided to limit the power supply from the battery during normal cruising operations as much as possible so that the battery would only be used during the most power-demanding phases of the mission (take-off and climbing); the fuel cells always work up to 20 kW, which is their maximum power output (approximately 50% of the power requested during take-off) for all the flight phases (take-off, climbing, cruising and descending). Moreover, having two completely separate power sources has an important impact on flight safety, which was the main driver of all the decision taken during the design phase; the battery was designed to supply 20 kW for 18 minutes so that it could work as an emergency power source in the case of the improbable failure of fuel cell, in order to allow the pilot to land safely.

The introduction of the second power source required a more complex electronic control system; the fuel cell was automatically selected as the main power supplier to minimize the usage of the battery, which is “activated” only when power that exceeds the fuel cell maximum is requested; at the same time, the controller needs to be able to instantly draw power from the battery to replace the fuel cell power in the case of a fuel cell malfunction. This controlling function was achieved by integrating the inverter that runs the electric motor with two innovatively boosters (designed and manufactured by Mavel srl, Italy) which modify the voltage of a power source (inverter-side) in order to “activate” or “deactivate” them as needed . The innovative design ensured a very low weight for the entire booster-inverter hardware set. The weights of subsystems are reported in Table 1 and a schematic of the power system is reported in Fig. 1.

The engine is a brushless electric motor produced by Phase Motion Control, it relies on air cooling and this has led to a saving in the weight as a water cooling system is no longer required. The motor-case was linked directly to the electronic boards (DC/AC inverter and DC/DC chopper). This can be considered an excellent solution, in terms of layout integration and cooling, because the air flow that runs along the motor wing tabs goes directly to the external surface of the converter case, where it continues

to carry out its cooling action. The fuel cell system, which is able to provide 20 kW of net unregulated power, consists of:

- A Fuel Cell Stack & Electrochemical System
- A Heat Exchanger System
- An Air Delivery and Water recovery system
- A Water Management Subsystem
- An Electrical and Electronic Support System and a Control & Internal Battery Subsystem.

An overview of the complete FC system is shown in Fig. 2.

Table 1

Component	Mass [kg]
Empty Aircraft [EA]*	221
Electro-Chemical Sun-System [ECSS]	103
Water Management Sub-System [WMSS]	
Heat Exchanger [HEXX]	
Control Sub-System [CtrlSS]	
Pressurized Hydrogen Sub System [H2SS]	51
Electric Motor [EM]	38
AC/DC+DC/DC Sub-System [ADDD]	14
Battery Pack 1 [BP1]	26
Battery Pack 2 [BP2]	26
Pilot [Pil]**	75
TOTAL	**554**

*Aircraft operative empty weight minus engine weight. The estimated weight includes a modified engine mount (3 kg) and the new propeller (4 kg).

**The ENFICA-FC converted Rapid200 is designed for a single pilot.

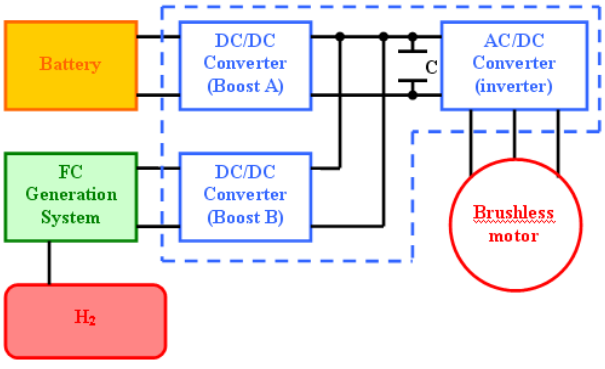


Fig 1. Power system schematic.

The Electro-Chemical Sub System (ECSS) consists of two separate fuel cell units. In order to provide a safe mounting system, the fuel cell stack was enclosed in a lightweight structure that also provides safe ventilation of any hydrogen leak and electrical isolation. The stacks were designed for a maximum current of 110A. The Air Compressors Sub System (ACSS) was designed as two-stage centrifugal compressors in series. This system brings fresh air from the engine cowling inlet and feeds fuel cell stack with compressed air.

The heat exchanger assembly (HEXSS) was placed in the front part of the engine bay under the electric motor. The fresh air flows through the engine cowling inlet and heat exchanger matrix, cools the waste water-air mixture from the fuel cell stacks and leaves the engine bay through the outlet opening at the front gear housing.



Fig. 2. Fuel cell system assembled in the mock-up fuselage

The cooled waste water-air mixture arrives at the cyclone where water is separated from air and directed back to the water tank to be reused. The Water Management Sub System (WMSS) consists of a water tank assembly situated in the right central wing leading edge part (originally occupied by fuel tanks) and a water pump, filter and flow meter situated in the engine bay. Finally, the Control Sub System (CtrlSS) comprises an FCS central communication and control module and an Internal Battery Sub System (a battery that is used to start-up the fuel cell).

The hydrogen storage and distribution system was one of the most important systems, from the conversion point of view; its volumes, weights and important impact on safety made it the starting point of each configuration designed during the project. The system consists of two Dynatek L026 tanks with accessories and it is shown in

its final configuration in Fig. 3. These tanks have a capacity of 26 litres each and they were manufactured for a working pressure of 350 bar (leading to a total H2 mass capacity of 1.2 kg). The whole assembly was installed in the baggage compartment behind the pilot, and it was separated from the cockpit by an aluminum wall. The tank compartment was sealed off from the cockpit to avoid any H2 leakage into the cabin. The tanks were secured with brackets mounted onto a lightweight construction which was directly attached to the load bearing structure. This solution ensures that all the operational loads and also the crash loads are properly absorbed by the aircraft structure.

Access to the luggage compartment, for the refilling and inspection of the tanks, is through the side door, which can be opened from outside. A refill valve was placed directly behind the door. Pressure regulators were installed in the former baggage compartment in order to ensure that the hydrogen flowing from the tanks to the fuel cell crosses the cabin at almost atmospheric pressure. The tank compartment was equipped with a passive venting system, which is aerodynamically activated during normal operations and with emergency activated relief valves (over-pressure and over-temperature).



Fig. 3. Hydrogen storage system.

Two Li-Po battery packs supply the additional energy that is necessary for take-off and climbing; packs (manufactured by Air Energy) are able to deliver 20 kW for 18 min. They are stored in two carbon fibre containers (and glass/fibre covers) which are secured with rails to the cabin floor on the co-pilot's side. The rails were necessary to remove the batteries easily for safe recharging operations; moreover, the batteries can be placed in different positions in the cabin and this allows a center of gravity shift, when necessary.

As stated before the presence of two different power sources and the will to fly relying only on fuel cells as much as possible, without compromising safety, led to the necessity of properly managing interaction between the fuel cell and the battery pack. As shown in Fig. 1, the power system includes a power electronics management unit that consists of two DC/DC converters (boosters) and an AC/DC inverter, whose function is to control the brushless motor. The boosters raise the input voltage so that only the selected power source becomes effective; without boosters, the higher voltage source would be the only one power is drawn from, and this doesn't allow correct selection. The purpose of the inverter is to properly modulate the direct current bus coming from the boosters in order to provide a sine current to the motor phases. The frequency and amplitude are closely connected to the rotational speed and torque of the motor, and hence to the aircraft performances. The system, designed and produced by Maver srl for the Enfica-fc project is very compact, very efficient (>97%) and very light (14 kg) (Fig. 4).

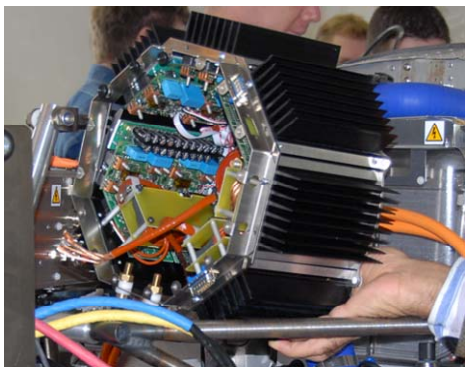


Fig. 4. Integrated vehicle controller.

4 Testing the System

Several experimental testing activities were performed at different levels of integration. Initially, the manufacturers or suppliers of each sub-system provided test results on their own systems; then a intermediate test campaign was carried on a semi-integrated fuselage mock-up in order to allow easier modifications or replacements of the single components to be made; finally, the complete system was tested on the real aircraft. This section presents a brief

summary of the most significant results obtained during the experimental activities.

4.1 Individual Sub-System Testing

Since the fuel cell system operates during the entire mission and represents the main power source, it was carefully tested for endurance at its maximum power output. The system was continuously tested by IE for more than 6 hours and no degradation of the performances were registered during the experiment. Several 6 hour long tests were performed to prove the reliability of the FC system [4]. The battery system is technologically more developed and so more reliable than the fuel cell system; the testing hence mainly regarded the safety of the system during charging and discharging. Attention was paid in particular to the behavior of the cell temperatures and the minimum single cell voltage during discharge; this latter aspect is very important because, for safety reasons, the battery system is not provided with an automatic cut-off (which has the purpose of protecting the battery from any damage that could occur because of a too low voltage level). Even though the flight mission was programmed to allow for safe gliding emergency landing from any point of the flight path, without automatic cut-off, the pilot is able to draw on all the energy accumulated in the battery (i.e. the most reliable energy source onboard), possibly damaging the cells, in order to land during a fuel cell failure if the gliding range is not sufficient enough to reach the airport. Some of the cells could experience a faster voltage drop after some charge/discharge cycles or after an excessive discharge below safe limits (Fig. 5) and a substitution of the cells that would performing less was needed for this application.

The motor, power electronics and vehicle controller were tested simultaneously. The main concerns about these systems pertain to the temperatures that can be reached during a full duty cycle. The most stressing conditions are [2]:

- At the very beginning of the take-off phase, because of the maximum power in conjunction with the low speed and there is therefore limited cooling for a short period;
- When climbing, because of the maximum power necessary for a relatively long time, although there is substantial cooling.

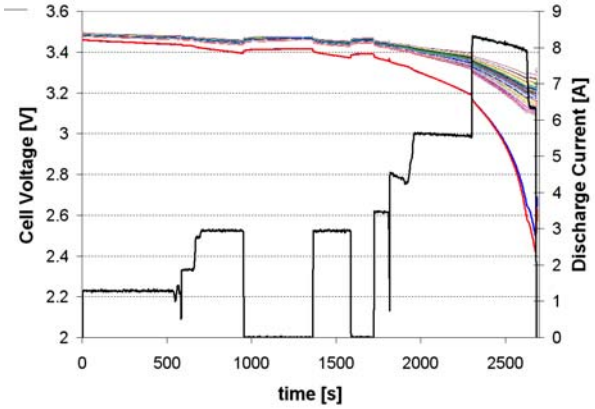


Fig. 5. Battery packs anomalous discharge.

Moreover, starting with the early experiments, the vehicle controller was tested for its capability to be able to switch immediately from the main to the second power source and back without any interruption or unexpected change in motor operations. These systems were bench tested with a DC power supplier simulating the two different onboard power sources and an air-blower simulating the air flow due to aircraft speed. A typical power profile adopted during the tests is shown in Fig.6: after a first part, representing an extended flight duty cycle, a "power blending" test was carried out; the total time was approximately twice the real flight duty cycle. The behavior of the temperature of the critical components during the same test is reported in Fig. 7. The maximum temperature reached in the inverter was 78° C (the maximum allowable temperature is 120°C), while the maximum temperature reached in the motor was 80 °C (the maximum allowable temperature is 180°C). The hydrogen storage system was tested by the tank manufacturer and by the supplier of the entire system for the maximum working pressure and burst pressure. The tests were conducted according to the test specification identified in ECE draft Regulation Annex 7 B9. The final maximum allowable pressure is 438 bar (350 bar is the normal working pressure of this application), while the burst pressure (representing the ultimate load of the tank) is 984 bar.

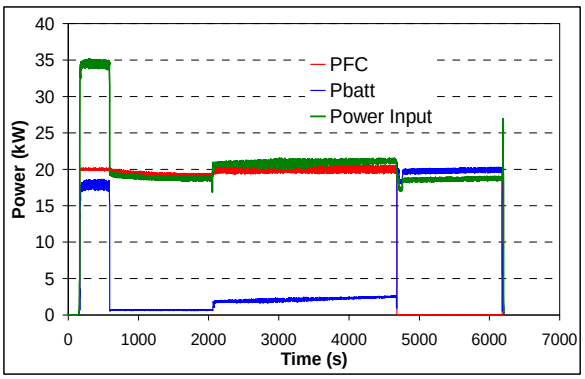


Fig. 6. Power profiles during an extended duty cycle power electronics test

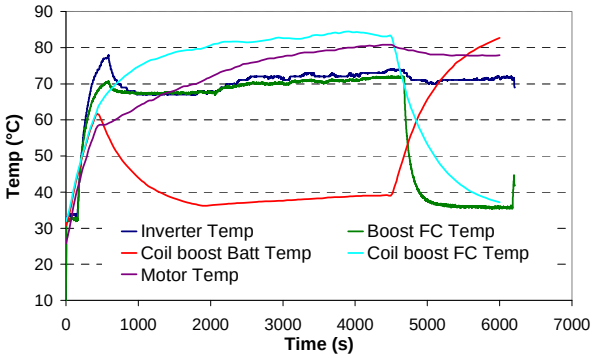


Fig. 7. Temperature profiles during an extended duty cycle power electronics test

4.2 Testing of the Semi-integrated System

Extensive testing of the semi-integrated system was carried out by POLITO, IE, APL and UNIPI at the University of Pisa laboratories. The whole fuel cell system in its final configuration, was completely installed on a fuselage mock-up (Fig. 2) together with the telemetry system; the motor/power electronic block was linked to a bench brake; hydrogen was first supplied from hydrogen bottles located in a bunker, for safety reasons, until the system proved to be reliable; other tests were then carried out with the actual hydrogen system with de-rated pressure (200 bar). Each system was provided with an air blower that simulated the theoretical airflow expected for that particular system. The batteries were replaced by an external generator for most of the tests to prevent deterioration of the cells due to excessive charge/discharge cycles. The main goal of this testing stage was to investigate and tune the communication between the systems, above all the vehicle controller and fuel cell. Moreover, as the fuel cell system is extremely complex and opportune strategies needed to be defined to pilot it during the normal and abnormal operations that may occur during flight operations, extensive testing was devoted to software related issues and tuning. From the hardware point of view, attention was paid to the same aspects reported in previous section, the temperatures being the most critical issue. Some typical duty cycle test results are shown in Fig. 8 and Fig. 9 for one of the final tests, where the system was basically ready for final installation.

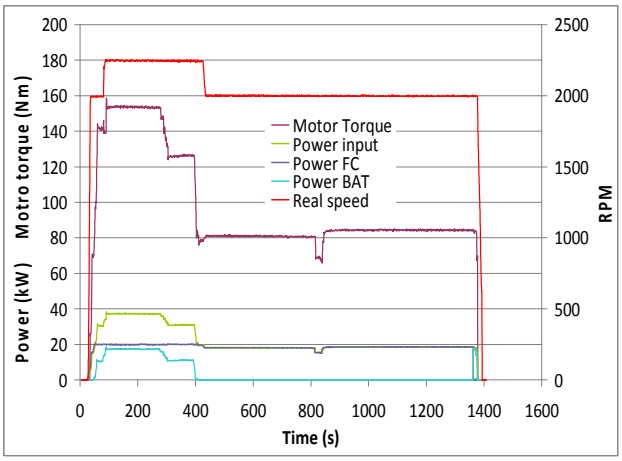


Fig. 8. Power, Torque and RPM profiles during a Typical Duty Cycle.

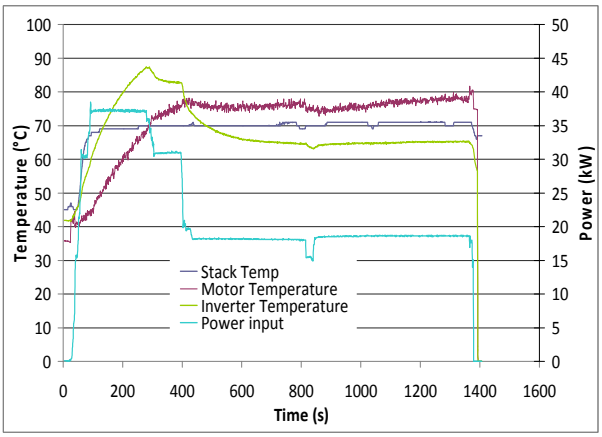


Fig. 9. Temperature profiles during a typical duty cycle.

4.3 Testing of the Semi-integrated System

The final and most extensive test campaign was devoted to the complete aircraft; the ground tests and flight tests were performed at the Reggio Emilia airport with the goal of validating the design and installation of the complete converted aircraft. This stage mainly involved the investigation of the behavior of the output power when connected to the real load (i.e. the

propeller), the behavior of the propeller, the handling of partial failures of the system, the temperatures with the real cooling system (i.e. the cooling system exposed to aircraft speed) and finally of the aircraft performances during take-off and cruising. Again great attention was paid to correct the handling of the two onboard power sources; a simulation of fuel cell failure is reported as an example in Fig. 10.

The real speed (purple line) has to be considered as a reference performance of the motor and hence of the propeller, while the power input (green line) is the power requested by the throttle. It can be seen that the system selects the fuel cell (red line) as the main source until 20 kW are required and when this threshold is exceeded, the controller starts drawing power from the battery (blue line). If, for any reason, the fuel cell cannot provide the requested power, the system immediately demands power from the battery, but the performance of the motor does not change. Moreover, the system tries to recover the fuel cell from its inoperative state and, if successful, to re-establish the fuel cell priority.

Several tests of this kind were performed, simulating different failures, and completely satisfactory behaviour was observed. The effect of the real load was investigated in terms of developed thrust, rpm coupling and absorbed power. As mentioned above, having the complete system installed allowed to be checked for the first time the real efficiency of the cooling systems (fuel cell, motor and power electronics). In order to investigate this aspect, the temperatures were observed during high speed roll-outs, which were performed to test the theoretical data pertaining to the take-off distances and speeds (Fig. 11)

As shown in Fig. 12, the cooling systems showed very satisfactory behavior, and the temperatures were kept below the admissible limits (the external temperature was 19°C).

In order to investigate the potential performances of the system for future developments, it was decided to check take-off without battery support. The aircraft was accelerated up to rotation speed (80 km/h) and, for safety reasons, the take-off was aborted before the climbing phase started (Fig. 13). It was possible to reach the rotation speed in 350 m (180 m is the usual distance when 35kW of power is supplied by both fuel cells and battery), but further testing should be carried out for the climbing phase in order to be sure that the aircraft can effectively run entirely on fuel cell power, and careful considerations have to be made on reliability before completely removing the batteries from the system.

4.4 Flight Tests

After the extensive test campaign, the aircraft was finally flown at Reggio Emilia airport. Six flights were performed, first with a 2 minute maiden flight and ending up with the world speed record for electric aircraft powered by fuel cells, according to a draft FAI sporting code, being broken (Fig. 14). The 350 bar APL S100 hydrogen refuelling system was installed at the Reggio Emilia airfield for these tests.

Six flights were performed and the telemetry-recorded flight data for mission 3 are reported in Fig. 15 as an example of a significant flight in terms of endurance. The flight area was chosen so that the aircraft was always able to land at the airport or at another airfield close by, while gliding with no available power. The final endurance obtained during this particular flight was 40 minutes (Fig.15), the limiting factor not being the hydrogen, as expected, but the water consumption; the capacity of the water tank (8 l) is under sized and this will be optimized in a future development. As shown, the temperatures were kept under their respective limits and the cooling system showed better performances than the ones recorded during the ground tests.
An FAI certified GPS data logger (LX Navigation Colibrì) was also installed on board to record the ground speed, flight altitude and flight path during the tests. According to the FAI draft rule for electrically powered flights, the speed was measured during two continuous 3km long runs and with an altitude variation of less than 100m between the start and the finish points.

The main results obtained during flights were:

- A maximum endurance of 39 minutes was recorded.
- A maximum speed of 135 km/h was recorded during runs 6-7.

- A greater maximum speed than 158 km/h was reached during a free flight.
- A minimum pressure of 70 bar was measured in the hydrogen tank at the end of the flights. 5.9 bar/min were approximately consumed during the flight; thus about 10 minutes more could be possible and this would increase the flight endurance to 49 minutes.
- A minimum value of the water level of about 15% was reached.
- The total GPS Horizontal Path Length (Taxi + roll out + climb + horizontal flight + landing) was 76,5 km

5 Conclusions

The extensive experimental campaign carried out during the ENFICA-FC project, as well as the theoretical estimations, have proved that fuel cell technologies represent a promising future innovation in aeronautics as a key-enabling technology for all-electric, zero emission, low noise aircraft. A new world speed record of 135 km/h and endurance of 39 min. was established during several flights for Category C (airplane) of the FAI Sporting Code. The previous record was established by the Boeing Research & Technology Centre (Madrid) in their first hydrogen flight (120 km/h for 20 minutes, but for a motor-glider, Class D, FAI Sporting Code); DLR also flew in 2009 but with a motor-glider (Class D, FAI Sporting Code) powered by fuel cells.

Higher flight speed values were measured during the free flight with altitude variations of 200m. Higher speeds than 155 km/h were measured several times, with a top speed of 180 km/h, which was measured during several diving and pull-up manoeuvre tests.

The positive handling qualities and satisfactory engine performances of these six flight tests led the team to consider these successful flights as a good starting point for further long endurance high speed flights. 2.8 hours of block time and 2 hours of effective flight were obtained during these 6 tests for a total path of 237 km.

The results obtained during the flights can be considered as a further step in the European and World Aeronautics Science field towards introducing completely clean energy (ZERO EMISSION).

At the moment, for general aviation aircraft, fuel cells and the related technologies seem to need improvement from the gravimetric efficiency point of view; for example the hydrogen storage system, weights 52 kg and contains 1.2 kg of hydrogen. The actual gravimetric efficiency does not allow the same performances to be achieved as the original aircraft, both as far as flights (speed, endurance) and for the payload capability are concerned (it was impossible to carry a second pilot/passenger in the converted aircraft); a mid-range development would be sufficient to obtain performances that could be compared to those of to a modern general aviation aircraft.

The real strength of the “all-electric aircraft” concept does not lie in an improvement in the performances, but in the environmentally friendly use of the aircraft itself; such an aircraft could be used in airports surrounded by urban centres, during the night and in an environments that are restricted because of excessive pollution risk.

Acknowledgments

The authors acknowledge the important contribution made by the European Commission funding the ENFICA-FC, EC 6th FP programme – Contract No. AST5-CT-2006-030779 (2006-2010) and of the Consortium partners.

6 References

[1] Romeo G, and Borello F. “Design and Realization of a 2-Seater Aircraft Powered by Fuel Cell Electric Propulsion”. *The Aeronautical Journal*, Vol. 114, n. 1155, May 2010.

[2] Romeo G, Cestino E, Borello F and Correa G. “Design and Realization of a 2-Seater Aircraft Powered by Fuel Cell Electric Propulsion. An Engineering Method For Air-Cooling Design Of 2-Seat Propeller Driven Aircraft Powered By Fuel Cells”. *Journal of Aerospace Engineering*, Vol 21, n.1, 2011.

[3] Romeo G, Novarese C, Pistek A, Hlinka J, Kouril M, Sosovjcka R and Danek V. “Identification of requirements for an aircraft in selected cathegory”. Technical Report D6.2 ENFICA-FC. 2007.

[4] Braid S and Orsillo A. “Technical Demonstrator Report”. Tech. Rep. D7.3-2 ENFICA-FC. 2008.

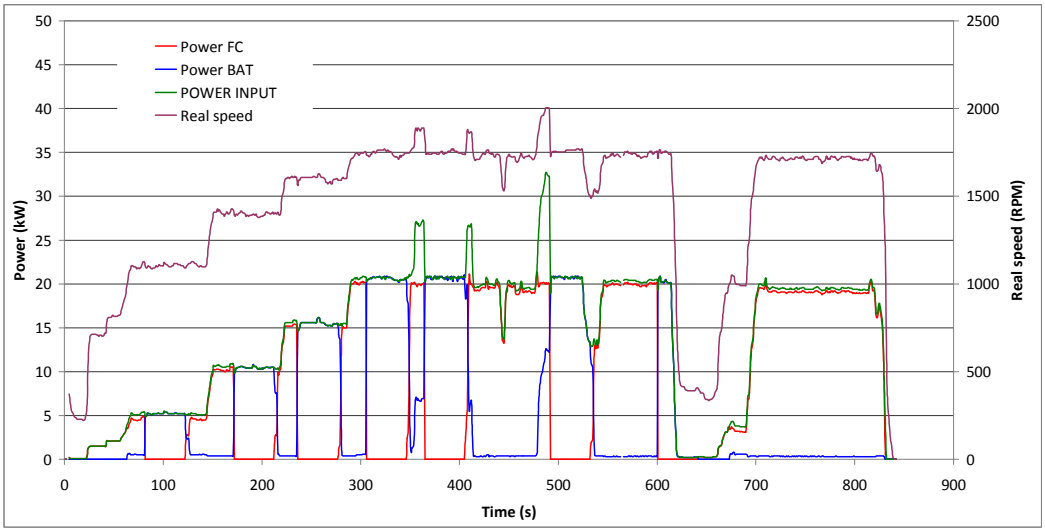


Fig. 10. "Power blending" ground test

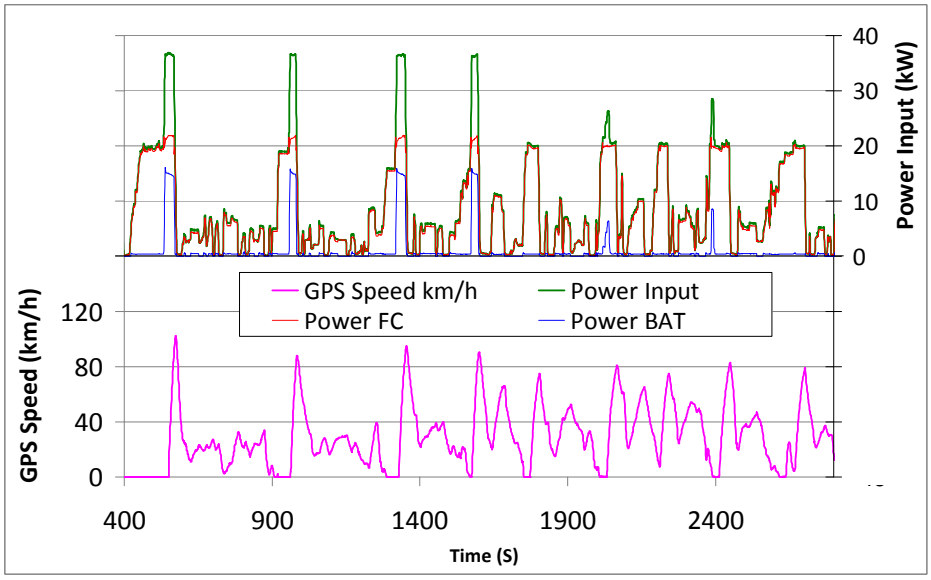


Fig. 11. Summary of the high speed roll-out powers and gps ground speed.

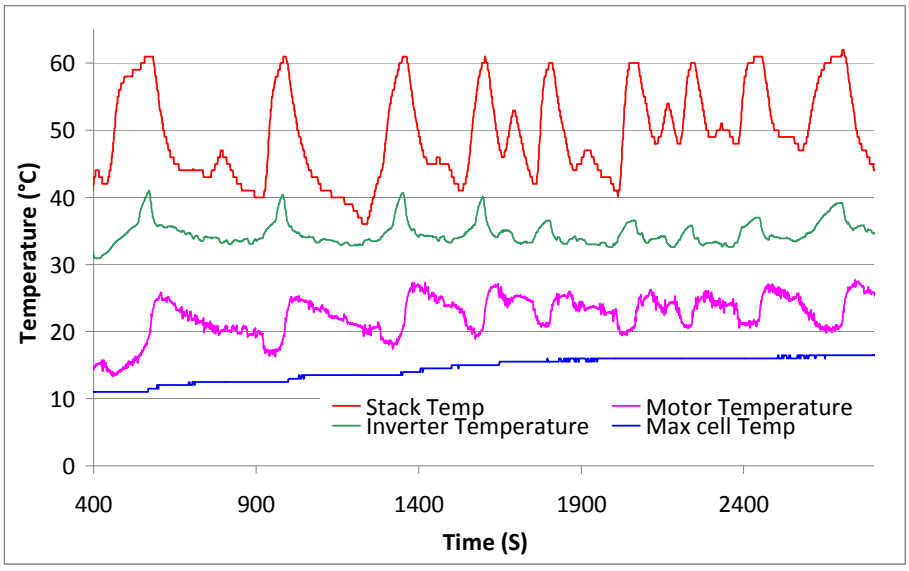


Fig. 12. Summary of the high speed roll-out temperatures

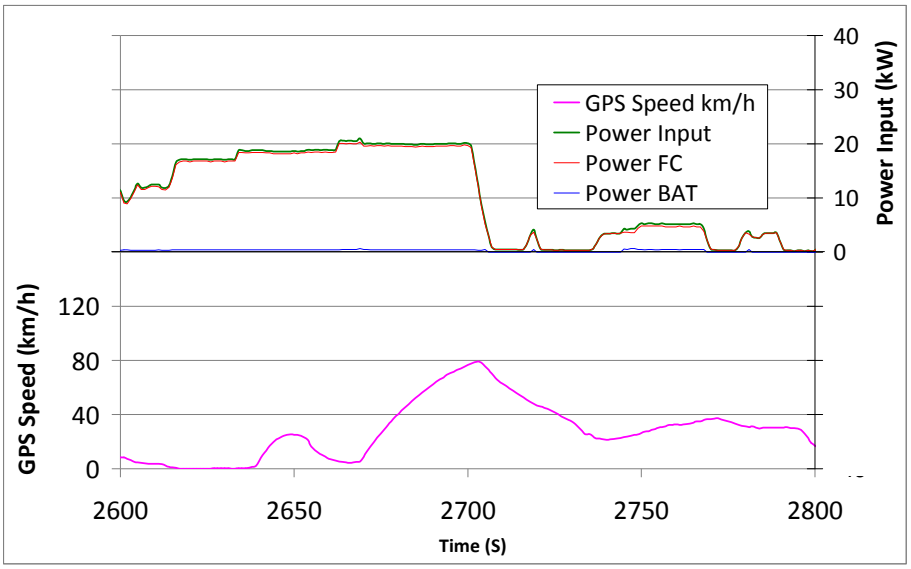


Fig. 13. Fuel cell take-off power and ground speed.



Fig. 14. RAPID-200 Fuel Cell in flight & the APL-S100 Fuelling Station
http://www.enfica-fc.polito.it/

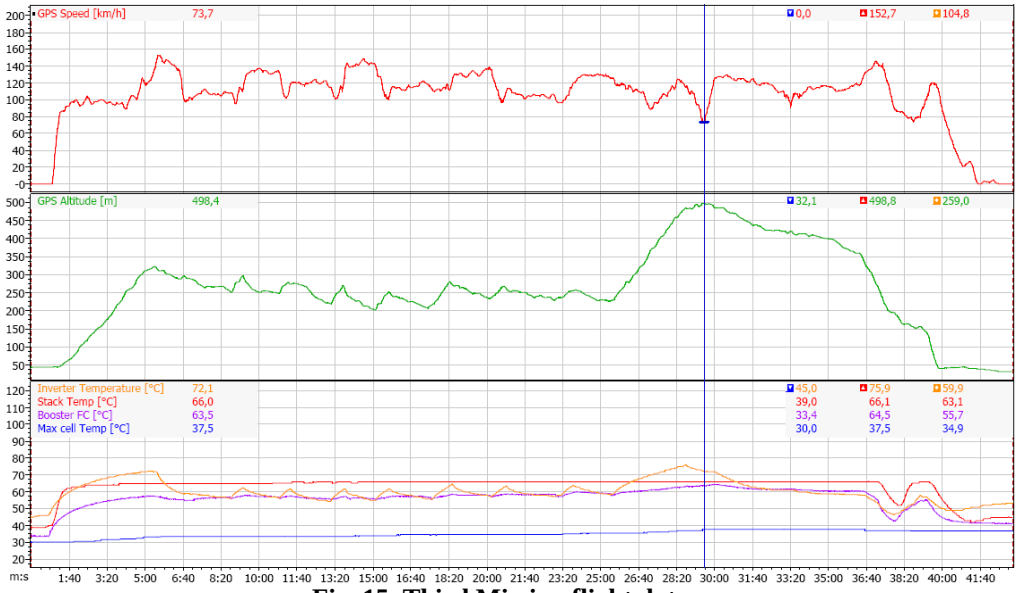


Fig. 15. Third Mission flight data.

© Copyright 2011by G. Romeo. Published by AIDAA/CEAS2011 with permission. All informations are Propriety of Prof. G. Romeo, and cannot be reproduced without his authorization. All right reserved.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Characterisation of the Extension/Retraction Dynamics of Electrically Powered Landing Gears

E. Denti, G. Di Rito, R. Galatolo and F. Schettini
Dipartimento di Ingegneria Aerospaziale – Università di Pisa
Via G. Caruso 8, 56122 – Pisa (Italy)

Keywords: *landing gear, electro-mechanical actuator, power absorption, simulation*

Abstract

The paper describes a simulation model of the electro-mechanical extension and retraction system of the landing gear of regional aircrafts. The main scope of the work is the evaluation of the electrical power absorption of the landing gear system, taking into account the effects of both aerodynamic and inertial loads. The simulation results reported in the paper highlight the relative importance, in terms of power absorption, of the different phenomena considered in the model (weight, inertia and aerodynamics). The mechanical, electrical and control sections of the landing gear system have been modelled in different simulation environments (Modelica-Dymola for the multi-body mechanics, Matlab-Simulink and AMESim for the electric/electronic parts) and then integrated into a Simulink platform for the simulation studies.

1 Introduction

The design of modern airborne systems and equipments constantly tends towards both technology innovation and integration of functions. This trend has been recently emphasized by research programs oriented to the conversion of aircraft systems to the "all-electric" solution [1][2][3][4], which point out important issues on power management for systems [5][6][7]. Actually, in the current aircraft technological frame, different types of power are derived from the engine for supplying the on-board systems: pneumatic, hydraulic, mechanical and electrical. The All-Electric Aircraft (AEA) concept is based on the total replacement of hydraulic and pneumatic power with the electrical one, and the rather suggestive acronym PbW (Power by Wire) has been coined for referring to a solution that satisfies each system power request by cables. This challenging objective can be only achieved by appropriately monitoring and managing the power requests (e.g. by temporarily reducing the power supplied to some systems during those flight phases in which the total request of electrical power could overcome the maximum available). As for any technological innovation, in particular when applied to safety-critical systems as the landing gear, new issues related to performances and reliability are expected, as well as specific concerns about systems interactions in terms of electrical power management. Actually, the design of the Energy Management System (EMS) logics requires the characterisation of the power absorption of each on-board system as a function of mission phase and aircraft operating point. In this context, a

strong effort is required for system engineers to develop models that are capable of predicting the aircraft systems performance, with particular attention to power absorption. The object-oriented modelling provides a convenient approach to this activity, since the model prototyping and the performance analysis of complex multi-physical systems can be carried out very efficiently.

The paper is focused on the development of an object-oriented simulation platform for the study of the Landing Gear System (LGS), with reference to a tricycle architecture for a typical regional aircraft. The extension and the retraction of each LG is obtained by an Electro-Mechanical Actuator (EMA). The simulation platform is composed of kinematic models of each landing gear, developed in the Modelica-Dymola simulation environment, while, for the EMAs simulation, two models have been used: a detailed (*behavioural*) EMA model, developed in the AMESim environment and oriented to the studies of the electrical power quality and the electrical network stability; a simplified (*functional*) EMA model, developed in Matlab-Simulink environment and oriented to low-frequency studies. The Matlab-Simulink environment has also been used as integration platform to interface the Modelica-Dymola and the AMESim models.

2 Simulation Platform

Figure 1 shows the Matlab-Simulink scheme for the nose LG. Similar architectures have been used for the left and right main LGs.

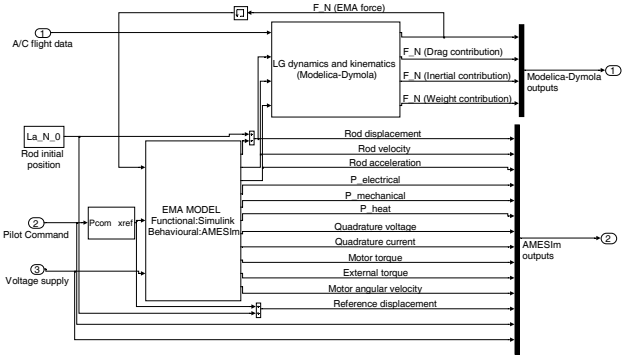


Fig. 1 Simulation platform for nose LG (Simulink)

The AMESim and the Modelica-Dymola blocks are included in the Simulink model in the form of S-Functions. The whole model is solved by the Simulink solver.

3 Landing Gear System Data

The model has been developed with reference to a LGS of a typical turboprop regional aircraft with one nose LG and two main LGs.

The LGS parameters have been defined starting from the basic design parameters reported in Table 1.

Tab. 1 Basic requirements for the landing gear system

Nose LG mass	40 kg
Main LG mass (right+left)	400 kg
Extension (Retraction) time	8 sec
EMA nose LG working stroke	148 mm
EMA main LG working stroke	310 mm
Position error steady-state	< 1%

4 General Features of the Model

The LGS model simulates the power absorption of LG extension and retraction actuators, as a function of LG legs position, between the up-lock and the down-lock boundary condition, environmental conditions, flight conditions and the electrical power made available by the EMS.

The model takes into account the kinematic link between the EMA rod and the LG leg during extension/retraction. The load transferred to each EMA, depending on the LG extension/retraction and on A/C manoeuvres is obtained by considering the LG weight (including the effects of A/C Euler's angles), the aerodynamic forces on the LG (including the effects of A/C angles of attack and sideslip) and the acceleration effects due to A/C manoeuvres and LG extension / retraction. The LG mass is assumed to be concentrated in a single point, as well as the aerodynamic drag force.

The flexibilities of A/C structure and LG components, as well as the gyroscopic loads, resulting from the wheels rotation, have not been taken into account.

The EMA dynamics has been developed by means specific models, with particular attention to its power absorption. The consumption of the

control units have not been taken into account, as well as the consumptions of up-lock and down-unlock actuators.

5 Model of the Landing Gear Structure

The same simplified architecture and the same extension/retraction kinematics are considered for both the nose LG and the main LGs (fig. 2 and 3). Each LG is represented by a generic rigid structure, which includes the fully extended shock absorber, bound to the A/C in the point *C*. The extension / retraction is obtained thanks to a rotation about an axis whose direction is assigned within the input parameters of the landing gear. The LG rotation is actuated by an EMA acting on the LG structure, and bound to the LG and to the A/C structure by spherical hinges (points *2* and *3* in Fig. 3).

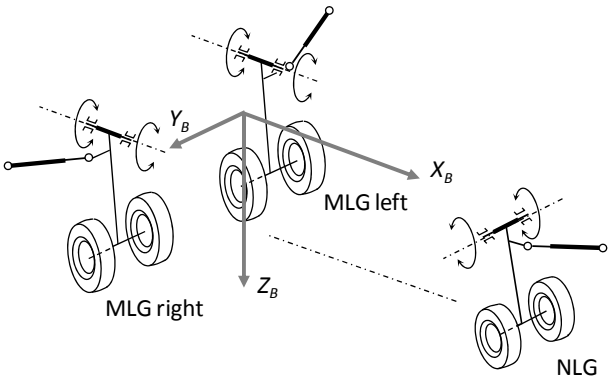


Fig. 2 Landing gear system layout and body-fixed A/C reference frame

The LG mass and the aerodynamic force are assumed to be concentrated in the point *H*. The latter is modelled as a drag force, with a constant drag coefficient (not dependent on the LG stroke). For each LG (nose, left, right), the points *H* and *3* are assigned through their coordinates on a LG-fixed reference frame (C; X_{LG}, Y_{LG}, Z_{LG} – Fig. 4), while the origin *C* of such a frame and the point *2* are assigned through their coordinates on the A/C body-fixed reference frame (O; X_B, Y_B, Z_B). The X_{LG} axis of the LG reference frame coincides with the axis of rotation of the landing gear during extension and retraction. The axis Z_{LG} has been chosen so that the X_{LG} -Z_{LG} plane contains the point *H*.

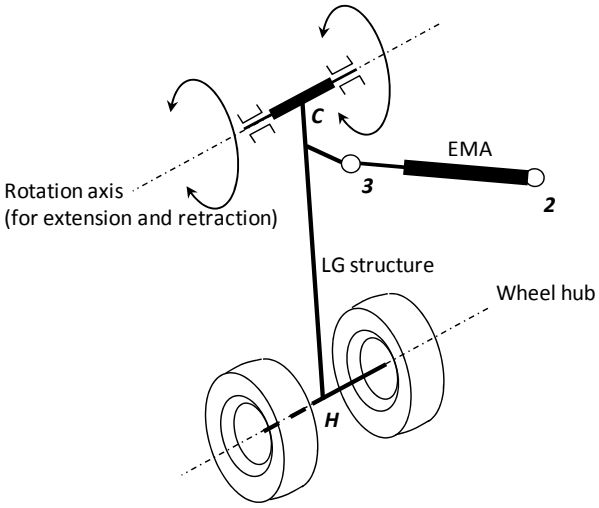


Fig. 3 Extension / retraction kinematics (scheme relevant to the nose LG)

The orientation of the LG reference frame is defined by the Euler's angles with respect to the A/C reference frame (Fig. 4).

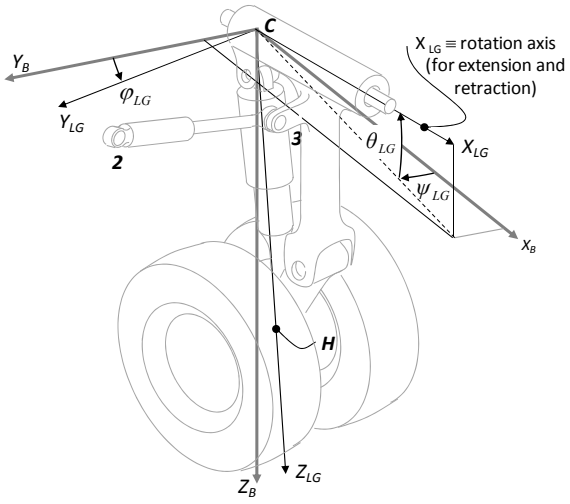


Fig. 4 Euler's angles of landing gear reference frame w.r.t. body reference frame

5.1 Modelica-Dymola Implementation

Figure 5 illustrates how the kinematic model of the nose landing gear has been implemented in the Modelica-Dymola simulation environment (similar schemes have been used for left and right main LG).

The interface *frame_a* is the A/C body-fixed reference frame (O; X_B,Y_B,Z_B) and the positions of the points *2* and *C* are defined with respect to such a frame through two elements, *b2_LG_B* and *bc_LG_B*, representing two massless rigid connections. The actuator is represented by the

element named *actuatedPrismatic*. This element receives the position signal *xa_LG* from the EMA model and two cylindrical hinges (*hinge_2a* and *hinge_2b*) allow the actuator to rotate about point *2*. Two elements, named *actuatedRevolved,* define two successive rotations for the orientation of the X_{LG} axis with respect to the reference frame *frame_a*. The cylindrical hinge *hinge_C* allows the rotation of the nose LG leg around the X_{LG} axis. To this element is linked the massless rigid connection *L3_LG_LG* which defines the position of the point *3* with respect to the point *C*. The spherical hinge *hinge_3* closes the kinematics loop by connecting the *L3_LG_LG* and the *actuatedPrismatic* elements. The LG mass, attached to the point *H*, is positioned with respect to the point *C* through the massless rigid connection *Lh_LG_LG*. This last element also identifies the frame *frame_b* which is used to apply the aerodynamic load at point H.

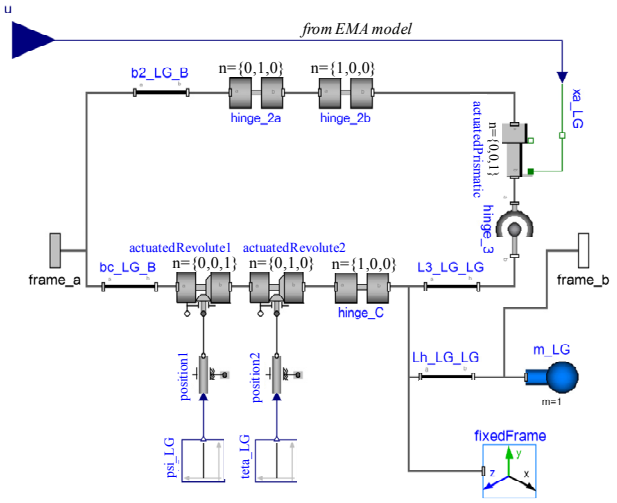


Fig. 5 Nose landing gear structure model (Modelica-Dymola simulation environment)

6 EMA Models

The LGS model is designed to be integrated within a Shared Simulation Environment (SSE) able to test and validate advanced technologies relevant to the implementation of the AEA concept. In particular, the SSE has to be able to validate the energy management logics used for the optimal sharing of the total on-board electrical power. Since the aircraft electrical network is a very complex system, the models to be integrated within the SSE have to be developed with particular attention to CPU time consumption, in order to permit the simulation of long-time mission phases.

In this research two EMA models have been developed: functional and behavioural. The first one is representative of the basic system dynamics and it is much less onerous in terms of computational time. The second one permits studies on electrical power quality and electrical network stability.

The detailed description of the two EMA models is reported in [8]. The inputs are: the pilot command, the actual voltage supplied by the EMS, and the reaction force applied by the LG to the EMA rod. The behavioural model, developed using the AMESim simulation environment, is organised into four main parts:

- electric motor (3-phase synchronous brushless type)
- mechanical transmission (gearbox, low-pitch screw jack and surface kinematics)
- power electronics (270 V_{dc} PWM drive of the motor coils via MOSFET switches)
- control electronics (closed-loops on actuator position, motor shaft speed and motor currents, including command limiting and current saturation).

The functional model, developed in the Matlab-Simulink environment, is simplified by replacing the AC network with an equivalent DC network. In addition, the absorption of the power electronics is neglected.

The results presented in this paper are obtained using the functional EMA model. This may lead to an underestimate of the power absorption, mostly when the EMA works in continuous duty cycle condition [8], that is not the case of a LGS.

6.1 Power Absorption

The EMA can work in motor or brake mode. In the first case, the mechanical power (P_{mech}) is an output and it is assumed to be positive as so the electrical power (P_{elect}) absorbed by the system. In the second case, the mechanical power is an input, representing the braking power, and it is negative as so the electrical

power. The overall power balance can be written as follows:

$$P_{mech} = P_{elect} - P_{trans} - P_{diss} \qquad (1)$$

where P_{diss} represents the power dissipated as heat and P_{trans} represents the power accumulated by the system during the dynamic transients (terms due to the circuit inductances and the EMA system inertia). The negative electrical power could be regenerated thanks to electrical equipments. This is not allowed by the current aerospace regulations, even if this restriction may be relaxed in the future. As a consequence, the braking power must be dissipated through proper circuits. The negative electrical power in the results presented in the next section indicates the braking energy to be dissipated and provide indications for the design of such circuits.

7 Simulation Test

The results presented in this work are referred to the command time histories illustrated in Fig. 6, characterized by an extension phase and a retraction phase. They have been obtained in levelled flight conditions, for an A/C speed of 70 m/sec. During the extension phase the nose LG EMA retracts, unlike the EMAs of the right and left main LGs which extend during such a phase. The data related to the aerodynamic drag forces have been derived from [9] [10].

7.1 Results

Figures 7-8 report the time histories of P_{elect}, P_{mech} and P_{diss} for nose LG and left main LG during the extension and retraction phases. The time history of the actual EMAs stroke is shown in Fig. 6, together with the commanded ones. For the considered LGS configuration, the nose LG EMA works in brake mode during all the extension phase and in motor mode during all the retraction phase. The behaviour of the main LG is slightly different because the EMA works in motor mode during the last part of the extension phase and in brake mode during the first part of the retraction phase.

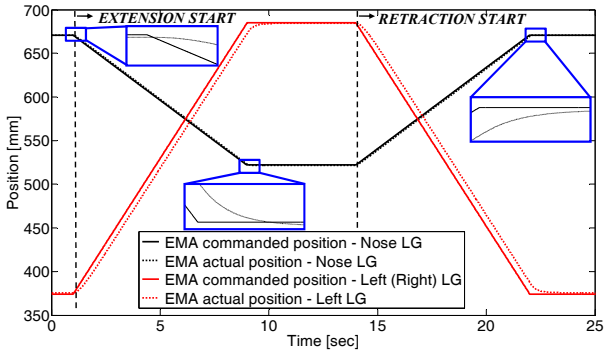


Fig. 6 EMA command positions for nose LG and main LGs (left and right)

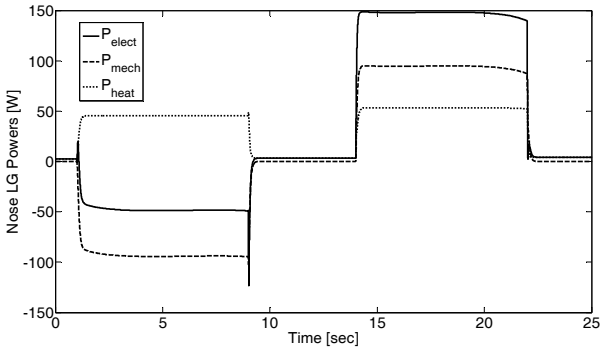


Fig. 7 Power flows of the nose LG

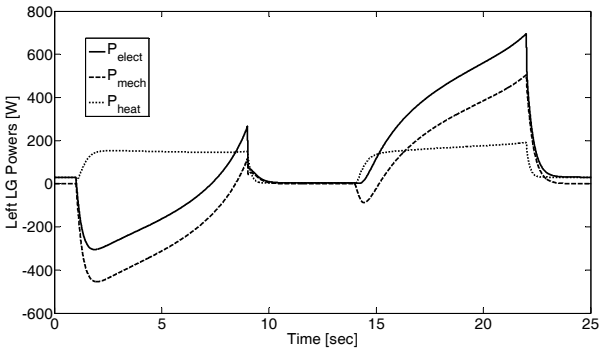


Fig. 8 Power flows of the left main LG (same results for right main LG)

To better understand the power flows, is useful to examine Fig. 9-10 where the time histories of the forces applied by each EMA are split in the three different contributions: drag (*FD*), weight (*FW*) and inertial (*FI*). Concerning the nose LG, the contribution to the force applied by the EMA and due to the LG weight (FW_{nose}) is maximum at the extraction start and it decreases during the rotation of the nose leg because of the reduction of torque due to the LG weight. In the meantime, the torque due to the drag force increases because the nose LG opens downwind. The consequence is that the total

force is nearly constant. Similar situations (about constant F_{nose}) occur during the retraction phase with an exchanged roles between the weight and drag contributions.

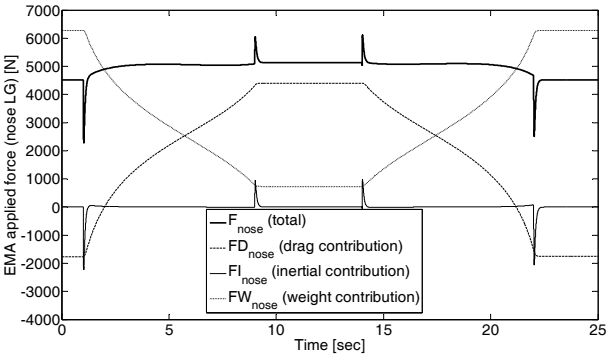


Fig. 9 Force applied by the nose LG EMA

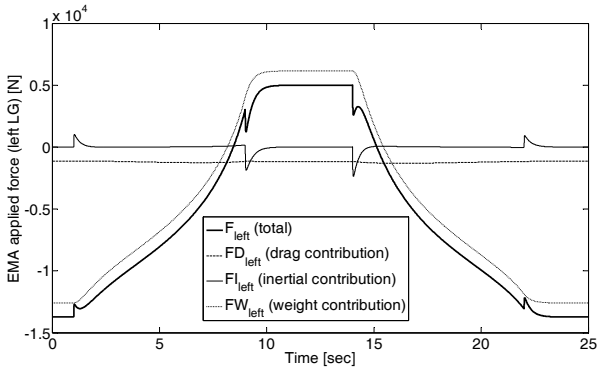


Fig. 10 Force applied by the left main LG EMA

Concerning the main LG, the weight contribution is predominant in both the phases, while the drag contribution is very little.

Finally, it is worth noting that the inertial contributions are relevant for both the nose LG and the main LGs.

8 Conclusion

The electrical power absorption assessment for a tricycle landing gear system of a typical regional aircraft has been assessed via an object-oriented modelling.

The model developed also evaluates the amount of braking power to be dissipated when the electro-mechanical actuators work in the brake mode.

Simulation results point out that the power flows for the nose LG depend on weight and drag forces in equal measure, while for the left and the right main LG the contribution due to the weight force is predominant. Obviously, the contribution of the LG mass increases with the A/C load factor. Concerning the inertial effects due to the relative motion between the LG and the A/C, the results show that they cannot be neglected in evaluating the total force acting on the electro-mechanical actuator. Nevertheless, such effects could be controlled by limiting the acceleration of the actuator rod at the beginning and at the end of the extension/retraction phases, through a proper modulation of the command time history.

It is interesting to note that the modulation of the command time history can be also performed in order to eliminate the peaks of electrical power, obtaining a lower, quite constant, level of power absorption during LG operation.

Acknowledgment

This research was partially conducted in the frame of the Clean Sky European Project (Green Regional Aircraft ITD).

References

[1] Emadi A. and Ehsani M.. “Aircraft power systems: technology, state of the art, and future trends”, *IEEE AES Systems Magazine*, January 2000, pp. 28-32.

[2] Cutts S. J.. “A collaborative approach to the more electric aircraft”, *IEE Conference on Power Electronics, Machines and Drives,* Bath (UK), 2002, pp. 223-228.

[3] Howse M.. “All electric aircraft”, *IEE Power Engineering Journal*, Vol. 17, No. 4, 2003, pp. 35–37,.

[4] Rosero J. A., Ortega J. A., Aldabas E. and Romeral L. “Moving towards a more electric aircraft”, *IEEE A&E Systems Magazine*, March 2007, pp. 3-9.

[5] Maldonado M. A. and Korba G. J.. “Power management and distribution system for a more electric aircraft (MADMEL) - Program status”, *IEEE AES Systems Magazine*, December 1999, pp. 3-8.

[6] Schallert C., Pfeiffer A. and Bals J.. “Generator power optimisation for a more electric aircraft by use of a virtual iron bird”, *25th International Congress of Aeronautical Sciences (ICAS)*, Hamburg (Germany), 2006.

[7] Anastasio V. and Di Donna L.. “Power by wire meets environmental friendly requirements”, *XIX*

AIDAA National Congress, Forlì (Italy), 2007.

[8] Denti E., Di Rito G., Galatolo R. and Schettini F.. "Power absorption characterisation of Electro-Mechanical Flight control actuators via detailed system modelling", 3rd *The International Conference of the European Aerospace Societies*, Venice (Italy), 2011.

[9] Roskam J.. "Methods for Estimating Drag Polars of Subsonic Airplanes", *Boards & Wraps*, Baltimore, MD, U.S.A., pp. 69-70,1973.

[10] Hoerner S. F.. "Fluid-Dynamic Drag", published by the author, pp. 251-252, 1958.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

New trends of power electronics converters architectures applied to more electric aircrafts

M. M. Prats, C. García, A.L. Rodriguez, L. Parrilla

Electronics Engineering Department. Escuela Técnica Superior de Ingeniería.
Avda/ Camino de los descubrimientos s/n
41092-Sevilla. Spain
Telf: 34 954486177
Fax: 34 954487373

Keywords: *power electronics, more electric aircraft, electrical generation, distribution system*

Abstract

In this work, a study of the different subsystems included in the electrical generation and distribution system of a commercial aircraft is explained. The most commonly used AC power generators in aircraft is described specially focused in the Variable Speed Constant Frequency (VSCF) technologies, because they are more efficiency, compact and reliable than Constant Speed Drive (CSD). That is the commonly solution in most of the currently commercial aircraft.
A complete review of the current and future electrical architectures trends are discussed in this work, where an important point developed is the Power Electronics section.

1 Introduction

The aircraft industry is developing the More Electric Aircraft (MEA) concept with the main goal of distributing only electrical power across the airframe. The replacement of existing systems with electric equivalents has, and will continue to, significantly increase the electrical power requirement. As consequence of that, it has created a need for the enhancement of generation capacity and changes to distribution systems. The higher powers will push distribution voltages higher in order to limit conduction losses and reduce cable size, and hence weight. A power electronic interface may be required to regulate generator output into the distributed power form, and elaborate more complex control strategies to manage the electrical power across the aircraft.
It is expected that significant cost and performance improvements can be achieved by consolidating power distribution into a single all electric system. A reduction in complexity and an improvement in maintainability are the anticipated benefits. The electrical system is the only system that has the potential to perform the tasks of all the other power systems [1][2][3].

Most current civil aircraft electric power systems use a combination of 115V 400Hz AC for large loads and 28V DC for avionics and battery-driven vital services.
Variable frequency AC systems are now being considered by the industry, in order to replace the Integrated Drive Generators (IDGs) on current engines (which need complex hydraulic systems to generate constant frequency AC voltages from a variable speed engine shaft power inputs) by simpler and lighter generators. Given that the More Electric Engine would generate electricity at a range of frequencies (because the various engine shafts to which the motor would/generators are connected will all rotate at different speed), it would make sense to use power electronics to convert all the motor/generator outputs to a single high voltage DC output (the voltage level of which would be possibly determined by the capabilities of DC switchgear) for transmission around the airframe.
The use of programmable solid-state devices and switching power devices in place of traditional electromechanical circuit breaker technology provides benefits to the aircraft in terms of load

management, fault isolation, diagnostic health monitoring, and improved flexibility to accommodate modifications and system upgrades.

2. Description of the electrical generation and distribution system

In conventional aircraft, a wound-field synchronous machine has been used to generate AC electrical power with a constant frequency of 400Hz. This machine/drive system is known as Constant Speed Drive (CSD) system.

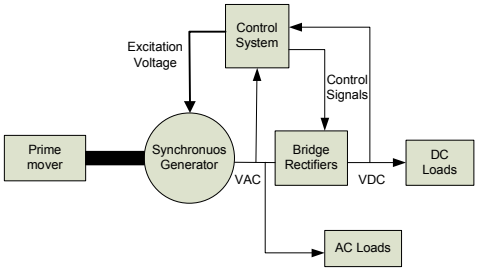


Figure 1. Typical Constant Speed Drive System

Figure 1 shows a typical constant speed drive system, where the synchronous generator supplies AC constant frequency voltage to the AC loads in the aircraft. Then, AC/DC rectifiers are used to convert the AC voltage with fixed frequency at the main AC bus to the multi-level DC voltages at the secondary buses which supply electrical power to the DC loads. Excitation voltage of the synchronous generator and firing angles of the bridge rectifiers are controlled via the control system of the CSD.

Recent advancements in power electronics, control electronics, electric motor drives, and electric machines have introduced a new technology: Variable Speed Constant Frequency (VSCF) system. The main advantage of this system is to provide better starter/generator systems. Other advantages are: higher reliability, lower recurring costs, and shorter mission cycle times.

Figure 2 shows the block diagram of a typical variable speed constant frequency starter/generator system. In the generating mode, the aircraft engine, which has variable speed, provides mechanical input power to the electric generator.

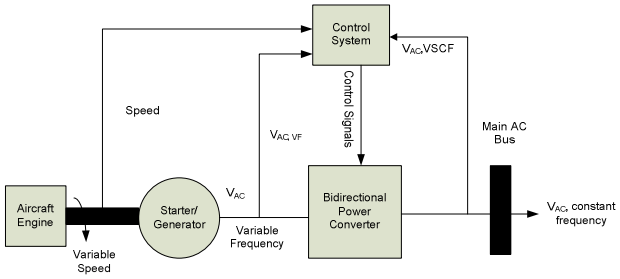


Figure 2. Typical Variable Speed Constant-Frequency Starter/Generator System

Therefore, the electric generator supplies variable frequency AC power to the bi-directional power converter which provides AC constant frequency voltage to the main bus. In the motoring mode, the constant frequency AC system, via the bi-directional power converter, provides input electric power to the electric machine which is a starter to the aircraft engine.

Synchronous (permanent magnet), induction, and switched reluctance machines are three candidates for VSCF starter/generator systems.

2.1. Induction Machine Generator

This section has focused on two VSCF technologies that have the potential in aerospace applications. The first technology is based on either the singly-fed induction machine or the doubly-fed machines (typified by the brushless reluctance machine). The second technology uses the switched-reluctance machine (SRM). The practical advantages of the simple construction and lower cost of induction machines compared to synchronous machines, have resulted in a widespread use of the squirrel-cage induction machine as a generator in the utility power system which uses wind or hydro power [3].

2.2. Squirrel Induction Machine

The operation of a squirrel cage induction machine as an isolated starter/generator unit, for aerospace applications, has been studied in [4]. The use of power electronics permits, in a more advanced system, to eliminate many of those problems associated with induction generators used in connection with utility power systems. For example, one advantage of the system was that no capacitors were required for excitation, and the system

operated at unity power factor. The system was based on field-oriented-controlled squirrel-cage induction machines, interconnected via a 20 kHz high frequency (HF) resonant ac link. The research demonstrated the theoretical and experimental results of a 10 hp system.

The initial power conversion system used a single phase high frequency (HF) resonant link proposed in A low frequency, variable amplitude and frequency three-phase system is synthesized from a zero-voltage switching Pulse Density Modulated (PDM) Converter which offers low distortion. The study focused on the operation and characteristics of the interface converters, outlined the principle of PDM, and demonstrated the attractiveness of HF link for isolation and power density-sensitive aerospace applications. Power system requirements for aircraft secondary power system is among some of the examples. A block diagram of the HF link is shown in The study, also, investigated interfacing the HF link to an induction motor using a volt-sec or flux (V/f) controller.

The use of a HF link reduces the converter losses and improves efficiency, but imposes a limit on the control system and introduces EMI problems.

2.3. Doubly-fed brushless Machine

Brushless doubly-fed machines have been investigated as variable as variable speed constant frequency generators for wind power generation in which the speed of the prime mover (horizontal- or vertical-axis wind turbine) is allowed to vary, but the output electrical power is maintained at a constant frequency (such as 60 Hz for the utility grid and 400 Hz for the aircraft system). These generators include systems based on induction and reluctance machine technologies. Although the investigation of these machines has shown promising results for wind power generation, the application can be extended to aerospace power generation. However, one distinction that should be made is that the wind power generation is obtained at low speeds (up to 1,000 rpm), whereas the aerospace starter/generators are operated at high speed, typically above 20,000 rpm types have been used as doubly-fed machines for wind power generation.

Related to the configuration of a variable speed constant frequency VSCF doubly-fed wound-rotor induction machine system, the machine is excited at both the stator and the rotor terminals. The stator terminals are connected to the ac power supply directly, while the rotor windings are connected through a bidirectional variable frequency converter to handle the slip power in both directions. Among the three power flow ports, namely, the stator terminals, the rotor terminals and the rotor shaft, the rotor terminals act as the energy regulating port, balancing the power flow of the entire system. An important feature of slip recovery power is that the frequency of the induced electric power can be kept constant, even if the rotating speed of the induction machine changes widely. However, since the rotor is controlled by a power electronic converter, harmonic currents are fed to the rotor windings and transmitted to the power supply, thereby changing the frequency of the stator windings. One advantage of the system is its controllability by a small capacity converter. This promotes cost reduction for the doubly-fed machine in VSCF operation.

A brushless doubly-fed squirrel cage induction machine has typical structure. The fundamental feature of the system is that two sets of windings, electrically isolated and displaced by a phase angle, with different pole numbers, are equipped in the stator of the machine. One set of the three-phase stator windings, labeled as “power windings,” is excited directly from the utility supply. The other set of three-phase windings, “control winding,” acts as an equivalent set of rotor windings, and is used to regulate the energy flow through a bidirectional frequency converter. The rotor is of the cage-or-reluctance-type.

In conclusion, the potential of doubly-fed generators for VSCF wind power applications, analyzed and experimentally evaluated in a large amount of researches. However, the potential use of the doubly-fed machines as starter/generators in aircraft systems has not been studied in detail.

There is a significant difference between the speed range in wind power applications and that of a starter/generator in an aircraft system.

2.4. Permanent Magnet Generator

A PM generator satisfies a large number of the criteria for a MEA generator. Achievable power densities are high, and maximum torque is available from zero to a wide rage of speed. In comparison with a wound field machine the rotor is mechanically simple, however difficulties are associated with magnet retention during high speed operation. High generator operating speed is

anticipated in More Electric Engines (MEEs) so a high strength magnet retention method will have to be utilized. The machine is fully excited at all speeds so in the wind milling condition sufficient power can be generated to power the emergency bus in the critical fault condition, this is one of the alternatives to replace the currently air driven generator (ADG) emergency power system. This is in contrast to the wound field machine where the speed of the rotor determines the power provided to the main exciter stage and thus the main rotor field strength.

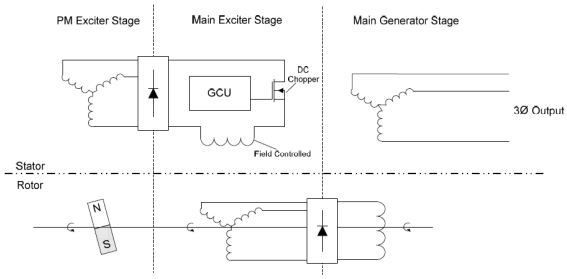


Figure 1. Typical three stage wound field permanent magnet aircraft generator.

In Figure 1, a three-stage wound field PM aircraft generator is shown. A PM exciter stage generates sufficient energy to induce a magnetic field on the main portion of the rotor. Field control of the main exciter stage is used to regulate main rotor field strength so that a constant voltage output is produced.

Operating temperature significantly affects the field strength of PMs; Samarium Cobalt (SmCo) based compounds exhibit the highest tolerances to heat with grades available with a temperature coefficient of -0.03%/ºC that can be used in environments up to 350ºC. This is the current operating environment ceiling to which PM machines can be employed.

Voltage regulation of PM generators has to be achieved with power electronics, which have considerable mass and volume. Advantageously the use of a power electronic converter stage allows ac or dc distribution bus requirements to be achieved.

Being permanently excited, the field in a PM machine cannot be disabled as with a wound field machine. This can result in undesirable torque ripple and elevated temperatures due to excessive current flow during fault operation. Fault tolerant PM machines have been developed with an increased number of phases and separated phase windings to mitigate these issues.

2.5. Switched Reluctance Generator

The switched reluctance machine (SRM) presents considerable interest for use as an integral starter/generator on future aircraft engines [10]. In this application, the starter/generator will be the primary aircraft electric power, and will also provide engine starting. It is meant that the rotor of the generator will be mounted on the shaft of the aircraft itself. By this reason, the electrical loads future aircraft will be much greater than on present aircraft as electrical actuator and electric motor drive pump become more usual. There are a few reasons for choosing the SRM for this application, all of them due to the categories of reliability and fault tolerance. With the rotor of the generator mounted on the shaft of the aircraft engine, the fact that the switched reluctance generator does not employ permanent magnets (in fact, there is only iron on the rotor) which will generate into a fault becomes a major advantage. This fault could be internal or external to the machine. The switched reluctance generator does not have a field winding which could fail. This maximizes the reliability of the machines rotor which will be internal to the aircraft engine and thus difficult to maintain. Other advantages of the switched reluctance generator are the independence of its phase windings, both electrically and magnetically, and the fact that machines windings do not physically overlap like they do in other electric machine types.

The mechanical integrity of the rotor permits high speed, high power density operation. The ability to operate in high temperature environments, in the region of 400ºC, and high speed operation, allow the possibility of direct-drive and, hence, the elimination of gear box and hydraulic system accessories on the aircraft.

However, the salient shape of the rotor causes wind age losses which become significant at high speed. The gaps maybe filled with a non magnetically permeable material to prevent this loss, but a mass penalty is incurred. Achievable power densities in a SR machine are lower than those possible with PM designs.

Related to the typical switched reluctance starter/generator system used in aircraft, in the motoring mode, the dc bus (270 V in aircraft) delivers power to the reluctance machine through the converter, and starts the aircraft engine. In the generating mode, the dc bus delivers the excitation to the machine and the gas turbine acts as a prime

mover. The power delivered by the machine exceeds the excitation power delivered by the source, and can be delivered to the dc source or pumps and other loads in the aircraft. Based on the machine phase currents and rotor position, the controller sends the appropriate signals to turn on and off the converter phases.

2.6. Comparison of different generators for aircraft applications

The comparison of the practical properties for above machines is given in Table 1 by qualitative analysis. PM machine has higher efficiency than induction machine and the SRM. The price paid for the simple structure and control of the SRM is its high torque ripple and acoustic noise. The efficiency and capability of the PM machine are compromised during constant voltage operation, especially in the light load and high speed range, while the SRM can run very fast and even be used as aerospace ISA at speed up to 60,000 rpm or higher. There are many on-shelf drives and technologies for induction machine control and the resolution requirement to its speed sensor is relatively low. Except for induction machine, all other machines require high resolution position sensor although the sensorless technology is under development.

Electric Machine type	Induction Machine	Permanent Magnet Machine	Switched Reluctance Machine
Efficiency and Compactness		√	
Low torque ripple and noise	√	√	
Easy loop control			√
Wide Speed Range		√	√
High power application	√	√	√

Table 1. Comparison between different generator for aerospace applications

3. Electrical Distribution Architectures

MEA would replace three individually optimize but collectively sub-optimized secondary power systems (electric, hydraulic and pneumatic) with one globally-optimized electrical system. However, in order to fully optimize the whole aircraft, the electric system architecture would need to be chosen carefully [6][11][12].

Recently studies demonstrated that the use of such technologies could, for a typical 200-seater aircraft, result in a 10% reduction in aircraft empty weight, a 13% in required engine thrust and a 9% reduction in aircraft fuel burn. These are significant improvements, both economically and environmentally.

Most current civil aircraft electric power systems use a combination of 115V 400Hz AC for large loads and 28V DC for avionics and battery-driven vital services.

Variable frequency AC systems are now being considered by the industry, in order to replace the Integrated Drive Generators (IDGs) on current engines (which need complex hydraulic systems to generate constant frequency AC voltages from a variable speed engine shaft power inputs) with simpler, lighter generators. Given that the More Electric Engine would generate electricity at a range of frequencies (because the various engine shafts to which the motor would/generators are connected will all rotate at different speed), it would make sense to use power electronics to convert all the motor/generator outputs to a single high-voltage DC output (the voltage level of which would be possibly determined by the capabilities of DC switchgear) for export to and transmission around the airframe. [9][10][11]

The rationale for such a choice can be summarized as follows:

- In order to reduce weight, power should be transmitted at high voltage (resulting in low current).
- Higher voltages also reduce the power losses in the transmission of the electric power because, for a given cable, the losses are proportional to the square of the current.
- Using the same voltages for generation and for transmission would avoid the need of voltage converters.

- AC systems require lager cables than DC systems, because the reactive power flows and, at high frequencies, the skin effect.
- Historically, AC systems were preferred because changes in voltage level could be achieved easily through transformers. Today, power electronics can now be used to make these changes in voltage level and also in frequency an in efficiency and safety way.

4. Power Electronics in MEA

Nowadays, novel ways of generating, distributing, and using power onboard are examined at the aircraft level.

Hybrid or bleed-less air conditioning systems, "More Electric Engines" (MEEs), fuel cells, variable frequency generators, complex embedded digital systems and distributed system architectures are just a few of the technologies vying for space on forthcoming aircraft; the concept is known as "More Electric Aircraft" (MEA) as presented in Figure 24. Recently, worldwide research into the future development of commercial aircraft has given rise to more advanced approaches to on-board energy power management and drive systems (Figure 55). These are now being carefully considered, and it is believed that electrical and power electronics systems have far more potential for future improvement than conventional ones regarding energy efficiency.

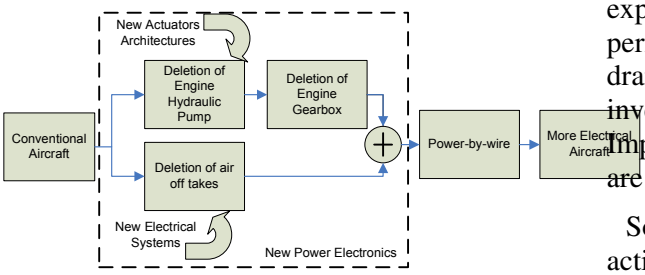


Figure 2. Current trends toward the MEA

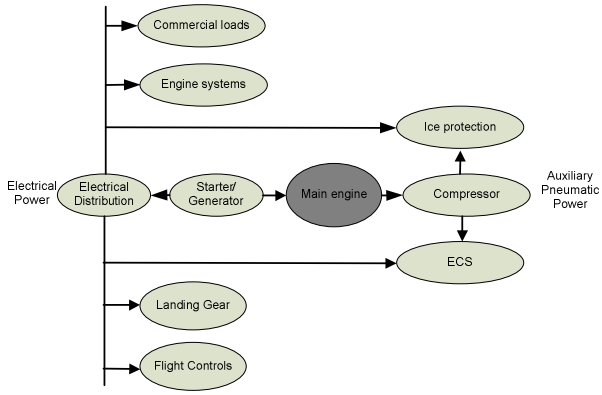


Figure 5. Schematic of a MEA power distribution

The first factor to take into account is the large amount of power electronics for power conversions and power users that MEA will involve: at least 1.6 MW for a next-generation 300 pax aircraft. The development of efficient and secure power electronics technologies is a great challenge. However, not only are power electronics necessary, but also efficient control of the electronics must be developed.

One major evolutionary technological advance that has contributed greatly to the feasibility of an electric aircraft non-propulsive power system has been the development of reliable, solid-state, high power-density, power-related electronics. Generator power control units, inverters, converters, and motor controllers consist of state-of-the-art silicon-based power semiconductor switching devices that include integrated gate bipolar transistors (IGBTs). It is expected that advanced composition, high-performance multi-layer ceramic capacitors will dramatically improve the power density of future inverters, converters, and motor controllers. Improved, high-efficiency electric circuit topologies are also the subject of on-going research.

Some of the higher power level equipment is actively cooled through the use of oil circulation or forced air convection. The extent of the use of active, fluid-based cooling systems is extremely application-specific and is yet to be determined. Lightweight, simplified, passive (non-pumped fluid-based) thermal management techniques are also a focus of research and will be used, wherever feasible, to maintain high reliability.

Power Distribution and Management Systems (PDMS) provide fully automatic monitoring,

control, protection, and switching of aircraft electrical loads under normal and emergency conditions with load management, including automatic load shedding and restoration, to make best use of available power. These systems comprise the Primary Power Buses, located close to the generators, with high power contactors and circuit breakers, and the Secondary Power Distribution Buses, located in the avionics bay, which provide the monitoring and control of the system, and contain some same circuit breakers and remote power switches.

The use of programmable solid-state devices and switching power devices in place of traditional electromechanical circuit breaker technology provides benefits to the aircraft in terms of load management, fault isolation, diagnostic health monitoring, and improved flexibility to accommodate modifications and system upgrades.

These power electronics systems have to be the interface between the generator and the distribution bus. The interface is required to facilitate a multitude of tasks:

• Regulation of the distribution bus from a variable voltage and frequency source.

• Efficient operation over a wide speed and power range i.e. from full-throttle to emergency power provision.

• Operation at very high fundamental electrical frequencies e.g. over 1 kHz when the engine is at full-throttle.

• Bi-directional power flow to allow engine start.

• Combine power from multiple generators operating at different frequency and voltage levels.

• Be physically distant from the generator i.e. a long cable length is present between the generator and the power electronic interface with possible EMC implications.

• Stable and rapid response to load changes

5. Conclusion

In this work, a description of the different parts included in the electrical generation and distribution system of a commercial aircraft is explained.

The most commonly used AC power generators in aircraft is described specially focused in the Variable Speed Constant Frequency (VSCF) technologies, because are more efficiency, compact and reliable than Constant Speed Drive (CSD), that is the commonly solution in most of the currently commercial aircraft.

In addition, a complete review of the current and future electrical architectures trends is discussed and Power Electronics systems are explained, where AC/DC inverters, DC/DC converters, battery chargers, Low Regulated Units are few examples of the possibilities of the energy conversion.

With these technologies, it will be feasible to use high power-density electrical power components to drive the majority of aircraft subsystems. These will become easier to maintain, more durable, lower in cost, and higher in performance.

References

[1] Beyond the More Electrical Aircraft, American Institute of Aeronautics and Aerospace, 2005.

[2] Elbuluk, M. E. and Kankam, M.D., "Potential Starter/Generator Technologies for Future Aerospace Applications", *IEEE Aerospace and Electronics Systems Magazine*, vol. 11, no.10, 1996, pp.17-24.

[3] A. Emadi and M. Ehsani , "Aircraft Power Systems: Technology, State of the Art and Future Trends", , *IEEE AES Systems Magazine*, 2000.

[4] J.G. Vaidya, “Electrical Machines Technology for Aerospace Power Generators,” *Proceedings of IECEC* 1991, Vol. 1, pp. 7-12.

[5] P.K. Sood and T.A. Lipo, “Power Conversion Distribution System Using a Resonant High Frequency AC Link,” *IEEE Trans. On Industry Applications*, Vol. 24, No. 2, March/April 1988, pp. 586-596.

[6] P.K. Sood, T.A. Lip0 and I.G. Hansen, “A Versatile Power Conversion for High Frequency Link,” *IEEE Trans. On Power Electronics*, Vol. 3, No. 4, 1988, pp. 383-390.

[7] S.K. Sui and TA. Lipo, “Field-Oriented Control of an Induction Machine in a High Frequency Link Power System,” *IEEE Trans. On Power Electronics*, Vol. 5, No. 4, 1990.

[8] S.K. Sul and T.A. Lipo, “Design and Performance of a High Frequency Link Induction Motor Drive Operating at Unity Power Factor,” *Proceedings of IEEE-*

IAS Annual Meeting, 1998, pp. 308-313.

[9] E. Alan and T.A. Lipo Control of a Poly-Phase Induction Generator Induction Motor Conversion System Completely Isolated from the Utility," *IEEE Trans. On Industry Applications*, Vol. 30, No.3, 1994.

[10] S.K. MacMinn and W.D. Jones, "A Very High Speed Switched-Reluctance Starter Generator for Aircraft Engine Applications," *Proceedings of NAECON '89*, Dayton, Ohio, 1989, pp. 1758-1764.

[11] C.R. Avery, S.G. Burrow, P.H. Mellor, "Electrical Generation and Distribution for the More Electric Aircraft," *Universities Power Engineering Conference, 42nd International, Brighton (UK)*, 2007, pp. 1007-1012.

[12] Elbuluk, M.E. and Kankam, M.D., "Potential Starter/Generator Technologies for Future Aerospace Applications", *IEEE Aerospace and Electronics Systems Magazine*, 1996.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Knowledge-based strategies for electrical aircraft wiring

B. Figlar and B. Mohr
Institute of Aircraft Design
Technische Universität München
Boltzmannstrasse 15, 85748 Garching, Germany
Email:mohr@tum.de

Keywords: *system modeling, aircraft wiring, knowledge based engineering, automated aircraft cable route generation*

Abstract

The trend towards "more or all electric aircraft" in combination with ever increasing requirements towards aircraft systems leads to a steadily growing complexity of electric wiring in commercial aircraft. Consequently, suitable methods for rapid and regular design of cable runs are required. This study aims at finding solutions for (semi-) automatic support of cable routing processes according to the principles of knowledge-based engineering. The developed approaches are implemented and tested in digital mock-ups. Additionally, industrial implementation feasibility is assessed.

The approach encompasses digitalization of the design and installation directives, the development of metrics for quality assessment of cable runs and the implementation of a parametric standard parts catalogue. Different solutions of communication between the use case CAD system and the knowledge database are presented. As a first step the feasibility of concurrent support in the design process and of the industrial implementation is demonstrated. In a second approach the complete automation of the definition of a cable run is demonstrated in order to find optimal routings. A pathfinding algorithm enables to communicate with the CAD system. The capability to find an optimal route is demonstrated with a simplified mock-up. Industrial implementation, however, is challenging in the short-term due to complex extensions of the mock-up and time-consuming computation.

1 Introduction

Knowledge-Based Engineering (KBE) has significantly gained relevance as a tool to speed up product development processes, especially in large scale projects such as the development of new aircraft and respective systems. The translation of engineering knowledge from domain experts into computer code that solves complex problems has demonstrated significant savings in manpower and time for routine design problems [1].

The base for KBE is the storage of design guidelines and rules in a database that can later be interpreted by computer software. All implemented design guidelines and rules represent cumulated expert knowledge leading to accelerated decision making. The ultimate goal of KBE is to rapidly solve dedicated problems and to build a knowledge-base for solving similar development problems in the long-term [2].

This paper presents different methods developed and assessed for the process of

routing electrical cables through an airframe segment by the means of KBE. The goal was the demonstration of a (semi-) automatic process to support virtual cable routing and the associated post-processing such as the qualitative assessment of cable routings.

2 Motivation

The trend towards "more or all electric aircraft" in combination with ever increasing requirements towards aircraft systems leads to a steadily growing complexity of electric wiring in commercial aircraft. Today the cabin and the respective cabin systems are the main differentiation criteria for airlines. As a result, a broad variety of aircraft interior systems exists and product complexity is hardly controllable with conventional methods. Furthermore, needs and attitude towards cabin interior evolve throughout the aircraft service life and cabin systems need to be designed for maximum flexibility with regard to possible upgrades. Each cable in each cabin version must be routed efficiently and according to the rules and design directives. The routes varying due to airline-specific requirements are called "customised routes". The increasing degree of electrification and customisation requires suitable methods for rapid design and inspection of cable runs. In this study different approaches to support and automate the cable run design process have been developed.

The current design process of customised routes is very iterative and time consuming. The routes are defined by subsequently setting attachment points into a reduced digital mock-up of a certain section of the aircraft's fuselage. At this stage the engineer can perform a visual check of the route and has functionalities available to measure the distance to surrounding parts or systems. By the time the route is virtually defined it can be transferred into another more complex digital mock-up (DMU) containing all systems of the fuselage. At this point the route can be inspected regarding clashes or conflicts. If, however, conflicts are detected the route has again to be changed in a different environment without supporting functionalities. Thus, nonconformities are identified very late and an immediate check of performance criteria during the routing process could accelerate the process. A support for finding the "optimum" route would also be desirable.

This study aimed at finding solutions for (semi-) automatic support of cable routing with support for optimal routing quality and demonstrates results with DMUs implemented in the CAD/ CAM software CATIA V5 by Dassault Systemes. It is additionally of major interest if the developed methods could be industrially implemented.

3 Proceeding

A necessary prerequisite for the study was the development of a DMU implemented in CATIA V5. For this purpose structural elements as well as system parts like air-conditioning ducts or pipe runs were modeled in CATIA. This DMU was the test bed for the methods in the study.

To thoroughly understand the complex system of aircraft cable routes it was necessary to analyze and structure all relevant system elements and their interfaces to the system environment. Official regulations and system specifications had to be implemented into the knowledge-base. It consists of three elements being a digital version of rules and directives that can be interpreted by CATIA, metrics to assess the quality of entire cable runs and a standard parts catalogue including, for instance, brackets and plugs.

Different approaches are then assessed regarding their possibilities to support the routing process. Algorithms for compliance checks with the rules database are implemented either in Visual Basic (VB) communicating with the CAD system via its VB-interface or directly implemented by means of macros or scripts. As CATIA V5 has a modular build-up of different design environments with coinciding functionalities it had to be assessed which module is best suited for automated cable routing applications.

Additionally, the specifications and constraints for an automatic cable routing procedure in CATIA V5 have been set up.

Different methods have been developed, each contributing to the goal of demonstrating the possibility of automation the generation of custom cable routings in commercial aircraft.

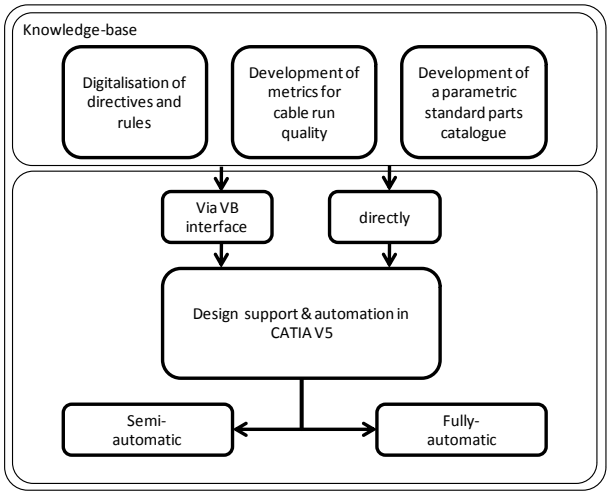


Fig. 1: Strategic approach of the study

The strategic approach chosen for the project is shown in Fig. 1 above.

3.1 Establishment of knowledge-base

For the definition of the knowledge-base the authors were able to analyse an original set of electrical design requirements and installation specifications of an aircraft OEM. For preliminary simplification of the problem only a subset of key rules and directives was chosen to be implemented in the knowledge-base.

Digitalisation of rules and directives

Three main types of key design rules were identified. They encompass distance requirements, minimum bending radii and compatibilities of routes. The different types of rules are shown in Table 1 below. The subset of chosen rules was structured and deposited in spreadsheet format to be accessible for CATIA or any other software. Other possibilities for storage of data in larger scale applications could be relational databases or object-oriented databases.

Table 1: Rules implemented in Knowledge-base

Distances	
	Minimum distances between routes
	Minimum distances between routes and structure
	Minimum distances between routes and pipes
	Minimum distances between routes and moving systems
	Maximum distances between fasteners
Bending radii	
	Single wires
	Bundles/harnesses
Compatibilities	
	Compatibility of routes in bundles/harnesses
	Compatibility of routes in connectors
Route Segregation	
	Spatial segregation of routes for redundancy reasons

It should be noted that this subset does not claim to include all important rules. Other directives, for instance, concern excess lengths for repairing purposes, the correct design of drip loops and grounding requirements.

A major issue that arose by implementing the knowledge-base was that rules exist that allow the designer to choose between solutions of different quality. For example, there are minimum distances between routes and pipes containing oxygen or liquids, which are defined as optimal solutions. However, the distances can be further reduced if precautionary measures are applied. Then the solution is no longer optimal but still acceptable. Since the electrical harness routing in commercial aircraft is a very complex problem within a very limited space envelope, it is very unlikely to find many optimal routing solutions. For this reason metrics are needed to assess and evaluate a run regarding its quality.

Metrics for cable run quality

Automation of aircraft cable routings requires quality indices in order to measure and compare different routings and routing alternatives. One possibility of implementation is to virtually discretize the cable routing and to analyze it at each discrete observation point. We have chosen a constant interval length between the single discrete points; a variable interval length can be chosen for instance in critical areas where a high resolution is desired. To obtain best possible results the interval length should be minimized, however it must be mentioned that computing time may limit this issue.

The cable routing is analyzed at every observation point with regards to all design rules implemented in the knowledge-base e.g.: minimum distances. As a result, the system engineer can determine the quality of complete cable runs as well as of single observation points regarding different quality criteria.

For the establishment of quality indices the rule base has to be investigated in detail. Two different kinds of rule categories have been identified:

- Category 1: Rules can be described binary (e.g. required minimum distance is maintained or not maintained)

- Category 2: Rules that cannot be described binary (e.g. there is an optimal minimum distance and if that distance needs to be further reduced there are precautionary measures that need to be applied)

A performance index ε is established which represents the degree of performance for a certain criterion from the installation guidelines. Category 1 decisions are binary, i.e. behind a certain limit no alternative exists. Category 2 decisions might have several solutions which offer different performance levels. Fig. 2 shows how performance functions can be implemented in the knowledge-base.

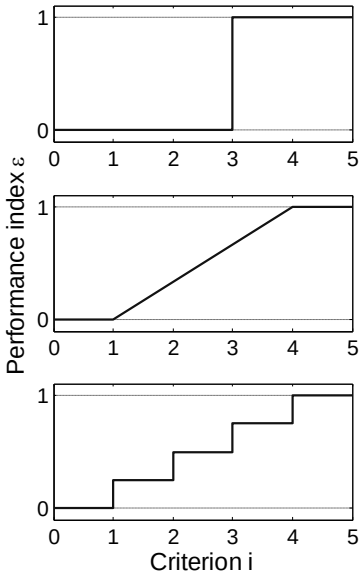


Fig. 2: Different functions for performance indices

To express the importance of a performance criterion a weighting factor δ_i is introduced (in our case the different criteria are: determination of importance of route segregation, bending radii, compatibility and minimum distances). A certain percentage value is assigned to each performance criterion; the sum of all values has to add up to 100% [3]. The overall quality of an observation point is expressed by a quality index C which is considering the performance indices $\varepsilon_{i,k}$ multiplied by the respective weighting factor δ_i for k surrounding systems:

$$C = \frac{\sum_{k=1}^{n}\sum_{i=1}^{n}\delta_i \varepsilon_{i,k}}{\sum_{k=1}^{n}\sum_{i=1}^{n}\delta_i (\varepsilon_{i,k})_{\max}} \qquad (1)$$

In order to keep complexity low and computing resources manageable an observation volume is introduced. For every observation point a spherical volume is defined in which the point's performance is evaluated. Thus, spatial separation of systems is considered and no unnecessary data is being produced. It is in the responsibility of the systems engineer to determine the observation volume's dimensions to ensure correctness of results.

3.2 Parametric standard parts catalogue

A parametric standard parts catalogue constitutes the third element of the knowledge-base. For the purpose of this study a selection of standard brackets, fasteners and plugs was categorised and parameterised. The motivation to parameterise the parts results from the requirement to precisely describe each part according to its application and geometry. The use of a search screen then allows displaying the parts that are applicable. The standard parts catalogue is directly implemented in CATIA V5.

4 Semi-automatic support of cable routing processes

Through the automation of repetitive engineering tasks the development cycle of a product can be accelerated [4]. The objective of the study is to assess possibilities to speed up the design process of customised cable runs in aircraft fuselages. A first semi-automatic approach intends to support the design process, but leave the definition of the route to the systems engineer. The objectives are to support the cable run generation process, to automate parameterisation of the generated run, to automate compliance checks with the rules database and to assess the quality of the generated route.

For this purpose two approaches have been followed that are different in the way CATIA is controlled and how it communicates with the knowledge-base. In a first approach it is attempted to implement algorithms and the knowledge-base as far as possible directly into CATIA, while in a second approach CATIA's Visual Basic interface is used to communicate with Visual Basic algorithms and the knowledge-base.

4.1 Route generation

CATIA V5 incorporates several modules capable of routing cable runs and harnesses in a three-dimensional space. The functionalities of these modules have been assessed regarding their potential to support the route design process as typically required for aircraft wiring applications. The typical requirements identified, mainly result from the special task of attaching customised routes to already defined main- or standard-routes. It is required that for segments where several routes are attached the geometry is automatically adapted and the parameters of the segment are changed such that compliance checks are possible for the whole bundle. On the other hand it is required that a "physical" connection pertains between each start and end point in the model. This is necessary to being able to assess cable lengths and weights. Ideally, the virtual cable should have the same physical characteristics as the real cable, in terms of bending behaviour and slack. Other desirable issues would be to minimise the construction effort, to minimise the number of parts in the model and hence to minimise memory allocation and computing time.

Different functionalities are available within the CATIA construction modules which are based on different construction philosophies. The main functionalities provided are called:

- "Flexible Routing"
- "Route a Run"
- "Branchable"

All functionalities have been thoroughly tested to demonstrate their strengths and weaknesses and the "Branchable" method has been considered for further investigation being the most promising functionality regarding the requirements described before. Several methods of application of the "Branchable"-method were tested and assessed. A first possibility of application is to model harnesses containing single wires or cables. The advantage is that the parameters of every single cable are accessible and that weight and cable length can easily be computed. The disadvantage is that the design effort rises significantly and the high number of parts allocates large memory resources. Another method of application is based on so-called "multi-branchable" elements. If a single cable is

coupled with an existing harness, the harness will inherit all relevant parameters of the single wire for the common segment length. This decreases parts count and memory allocation significantly. However, the single wire is split up in several parts during the operation and a determination of cable length and weight becomes difficult.

Due to the deficits mentioned, a new functionality for the creation of cable bundles in CATIA V5 has been created on the basis of the main functions of the "Branchable" method for the approach of implementing algorithms directly in CATIA. The generation of route geometry was then partly automated by the use of templates and scripts.

For the approach of controlling CATIA externally a Visual Basic GUI was developed aiding the engineer through the geometry generation process. Instead of using the predefined "branchable" functionalities, standard geometry elements had to be used, since "branchable" functions are not reachable via the VB interface.

4.2 Parameterisation

The parameterisation of the cable runs was implemented as queries at the time of definition of the route. The assigned parameters are then automatically combined for harness segments were different routes are attached. Despite the user interface the method is identical for the CATIA external and internal approach.

4.3 Compliance checks

For the purpose of automation of compliance checks according to the rules database, all parts in the DMU must be identified and categorised. This is done by assigning each part to so-called lists, e.g. parts may be assigned to structural elements or air-conditioning elements, etc. Then the generated route can be inspected regarding conflicts with the differently classified elements in the DMU.

The compliance checks are implemented in the form of macros for the internal approach and for the external approach VB algorithms perform the checks or call the respective CATIA functions.

Finally, reports are created and conflicts are visualised in CATIA.

4.4 Automation of route quality assessment

Finally, a method for the quality assessment of a generated route was developed. To allow the use of the quality metrics described before the generated route is discretized by the use of CATIA point elements. These points are set in a distance determined by the systems engineer along the run's centre line. At each point the distances regarding all other categorised elements in the DMU are measured and assessed with the quality metrics. After the assessment of all points a report about the quality of the cable run is generated and provided to the engineer.

4.5 Conclusion

All solutions interactively guide the system engineer through the routing process. Relevant data (like route category) are requested as input from the user. These data are further processed and other parameters like minimum bending radii are calculated by using the rule dataset. Finally, route quality is inspected and assessed. The engineer is informed about inspection results through reports and visual feedback in CATIA V5.

It can be said that the proposed procedure for the interactive (semi-automatic) cable run creation in combination with an immediate conformity check helps to unburden the systems engineer and to decrease error ratios in DMUs. However, a fully automatic cable run creation is not possible with the current tools and requires a completely different approach. One possibility for fully automated cable routing is presented in the following paragraph.

5 Automation of cable run generation

The second approach to support the harness routing process intends to completely automate the search for regular route geometry between

two predetermined points. The requirements towards automation can be expressed as:

(1) *Find the shortest possible route between two points*
(2) *Do not violate any rules in the Knowledge-base like minimum required distances*
(3) *Allow reduction of minimum distances by use of additional precautionary measures only if path length can be significantly reduced (a trade-off must be defined between length and route quality)*

A possible solution to such a complex task is provided by so-called pathfinding algorithms. Pathfinding is a branch of artificial intelligence and widely implemented in computer games and car navigation systems. A pathfinding algorithm is a sequence of commands to find a connection between two preselected points in a discretized search space, called the graph. The typical problem in pathfinding is obstacle avoidance [5]. There are well-established pathfinding algorithms that do not only allow to find a connection between two points, but also to find an optimal path. The optimisation objective could basically be minimisation of the path's length or of another cost factor. The cost of a path is determined depending on the weighting of the different nodes of the discrete search space. The exact functionality of pathfinding algorithms is not subject to this paper, but is for instance described in [6] and [7].

5.1 Approach

Different pathfinding algorithms have been assessed in this study. The most promising algorithm for harness routing purposes is the so-called A*-algorithm (pronounced "A-star"). The A*-algorithm allows to find the optimal path between two points [8]. The use of pathfinding algorithms for the purpose of finding optimal route geometry requires a discrete search space. This means that the DMU of structure and other systems must be extended by a grid of nodes available for routing purposes. Contrarily to the semi-automatic approach described before, the compliance check with the knowledge-base is performed prior to the route definition process. For this purpose each node of the search space is evaluated regarding its goodness to be crossed by a specific cable run and then weighted according to its goodness. The fitness of a route can then be determined as a trade-off between the length of the cable run and its quality.

The procedural method of the approach is shown in Fig. 3. Necessary prerequisite for the method is that a discrete search space must be provided to the pathfinding algorithm and that all systems and structural elements in the DMU are classified and can be identified by the algorithm. The algorithm must know if an element in the DMU is, for instance, part of the structure or a pipe containing oxygen. The designer's task is to define the parameters of the harness or cable to be routed and to determine the search space available for this cable run. According to the classification of the different elements in the DMU and the parameters of the route, each discrete node of the search space is then weighted by the algorithm, according to the quality metrics described before. After all available nodes are weighted the actual pathfinding algorithm starts to search for the optimal route between the two predetermined points.

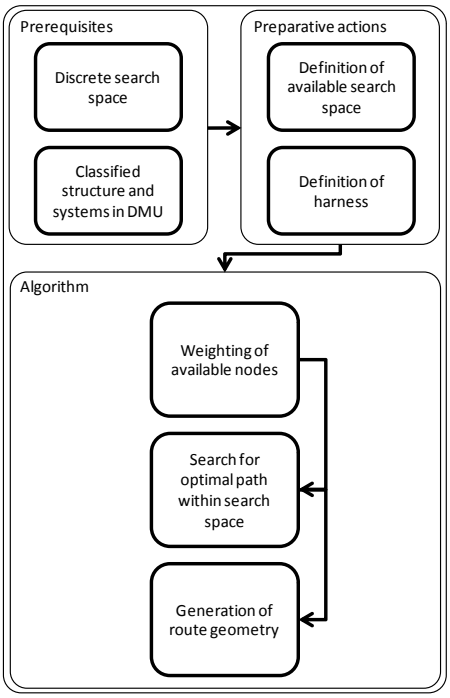


Fig. 3: Procedural method of automation of the route geometry generation

5.2 Implementation of the proposed methodology

A strongly simplified schematic model of an aircraft fuselage section was created in CATIA V5 to assess the practical applicability of the proposed method. This DMU only consists of structural elements, being frames and floor beams as shown in the Fig. 4.

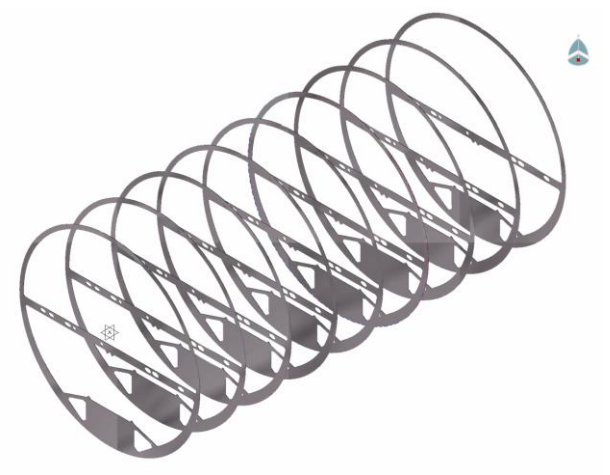


Fig. 4: Schematic structural mock up of the use-case aircraft fuselage

The model was discretized by applying a three dimensional grid composed of CATIA point elements. The search space was reduced to areas that would be preferably used for routing purposes, meaning that the cargo and passenger spaces were excluded from the search space. A graphical user interface (GUI) was developed communicating with CATIA via its Visual Basic interface. The GUI can be used to reduce the overall discrete search space to certain areas and to define the start and ending point of the route. Then a query requests all necessary parameters describing the harness or cable, like type of route, diameter, etc. The algorithms for weighting of nodes and pathfinding were implemented in Visual Basic and can subsequently be started via the GUI. When the optimal path is found the route geometry is automatically generated and saved in the CATIA mock up. In this preliminary simplified model the optimisation only accounts for the length of the route and distances to the structure. A simplified weighting method was developed for this purpose to describe the quality of the route depending on the distance to the structure.

5.3 Result and conclusions

The use of pathfinding algorithms is an innovative approach to support the process of route geometry definition. In the simplified model an optimal path could be automatically found. However, there are some issues that require further improvement and investigation. The weighting of all nodes in the available search space turned out to be the most time-consuming process of pathfinding. One means of avoiding unnecessarily long computation times could be to reduce the available search space as much as possible. Another issue is the definition and check of bending radii, which would require a more sophisticated adaption of the pathfinding algorithm, since the methodology of finding the path would need to be fundamentally changed. A third issue that has not been solved yet is that only one optimal route can be found by the path finding algorithm, meaning the designer cannot be provided with alternative optimal or “near-optimal” route geometries. This however could be helpful if knowledge of the designer about the system is at a higher level than stored in the DMU. Alternative routes are also necessary for the purpose of route segregation. This issue, however, can be solved by providing the algorithm with different available search spaces to receive alternative routes.

6 Conclusions and outlook

Based on the design and installation directives, a systematic knowledge-base was developed that can be interpreted by CATIA V5. Different methods have been developed to support the routing process, based on the idea of knowledge-based engineering to use a centralised database of rules, standard parts and quality assessment routines.

The first approach intends to support the engineer while defining the cable run geometry in the DMU. Rules like minimum distances or minimum bending radii are automatically checked during the routing process.

Additionally, a quality protocol can be generated for the entire cable run to make it comparable to alternative runs. The methods were implemented via the CATIA internal programming possibilities such as macros or Visual Basic communicating with CATIA via the respective interface.

In a second approach the possibility of completely automated routing was demonstrated. Here the engineer only needs to define start and end point of the desired run and the type of cable or harness connecting the two points. The algorithm based on the A*-pathfinding method then automatically finds the best possible run between the two points, respecting different quality criteria and without violating rules.

The first approach seems to be easily adaptable to industrial needs and its implementation seems to be realistic. The second approach is very innovative and it was demonstrated within a simplified mock-up that it is possible to automatically generate an optimal cable run. Industrial application however requires further investigations on how to manage the complexity of real-life cable routing in a pathfinding problem.

7 Acknowledgements

The authors would like to thank T. Richter and H. Mechler for making this project possible as well as L. Drees, F. Fisch, C. Fuchs, R. Gusmini and C. Kurz for their significant contribution to the achievement of the project goals.

8 References

[1] Bermell- García, P. et al.: “Effective abstraction of engineering knowledge for KBE implementation”, *International Conference on Engineering Design*, ICED. Glasgow, UK, 2001

[2] González, A.; Dankel, D.: *The engineering of knowledge-based systems*. Theory and practice. Englewood Cliffs, NJ: Prentice Hall, USA 1993

[3] Breiing, Alois; Knosala, Ryszard: *Bewerten technischer Systeme. Theoretische und methodische Grundlagen bewertungstechnischer Entscheidungshilfen*, Springer, Berlin, 1997

[4] Shea, K.: “Computer Aided Product Development”, Lecture Script, Technische Universität München, Garching, Germany, 2011

[5] Stout, B.: Smart Moves: Intelligent Pathfinding. *Game Developer Magazine*, Jg. 1996, H. 10/11.

[6] Patel, A.: The A* Algorithm. Online available at: http://theory.stanford.edu/~amitp/GameProgramming/AStarComparison.html.

[7] Lester, P.: A* Pathfinding for Beginners. Online available at: http://www.policyalmanac.org/games/aStarTutorial.htm, last update 2005.

[8] Richter, T.: „Simulationsmethodik zur Effizienz- und Komfortbewertung von Menschenfluss-prozessen in Verkehrsflugzeugen“. Dissertation Dr. Hut Verlag, München, Germany, 2008

CEAS 2011

Space Guidance, Navigation and Control

3rd CEAS Air&Space Conference
21st AIDAA Congress

Laser Altimetry at Ganymede: Mission and Performance Analysis

Aniello Memoli[1], M.Rosaria Santovito[2], MarcoD'Errico[1]
[1] *Seconda Università degli Studi di Napoli, Viale Beneduce n.10 – 81100 Caserta (Italy)*

[2] *CO.RI.S.T.A. (Consortium of Research on Advanced Remote Sensing Systems), Viale J.F. Kennedy n. 5 – 80125 Napoli (Italy)*

Keywords: *Ganymede, laser altimeter, performance, EJSM mission*

Abstract

This paper presents a study of performance of a laser altimeter for topographic analysis of Ganymede surface. The scenario is the Europa Jupiter System Mission (EJSM), a National Aeronautics and Space Administration (NASA)-European Space Agency (ESA) unmanned joint space mission slated to launch around 2020. EJSM will offer the chance to study Jupiter System and its origin, focusing on moons Europa and Ganymede. The use of laser altimeters is particularly appropriate whenever the topography of planetary surface is the objective of a space mission.

In this paper, after orbital propagation realised with a specific tool, Laser Altimeter performance analysis is carried out simulating Ganymede topographic characteristics and the potential targets on the planet with analytical models and computer codes.

1 Introduction

A laser altimeter has been considered for the joint NASA-ESA EJSM mission to study Jupiter System.

The main targets of this mission are the Jovian moons Europa and Ganymede. To this end, two spacecrafts should be sent: NASA Jupiter Europa Orbiter (JEO) that will investigate Europa in detail and ESA Jupiter Ganymede Orbiter (JGO) focused on Ganymede.

For Europa and Ganymede, the main scientific objectives are [1]:

- characterize and determine the extent of sub-surface oceans and their relations to the deeper interior;
- characterize the ice shells and any subsurface water, including the heterogeneity of the ice, and the nature of surface-ice-ocean exchange;
- characterize the deep internal structure, differentiation history, and (for Ganymede) the intrinsic magnetic field;
- compare the exospheres, plasma environments, and magnetospheric interactions;
- determine global surface compositions and chemistry, especially as related to habitability;
- understand the formation of surface features, including sites of recent or current activity, and identify and characterize candidate sites for future *in situ* exploration;

- understand the Jovian satellite system, especially as a context for Europa and Ganymede;
- evaluate the structure and dynamics of the Jovian atmosphere;
- determine the interactions occurring in the Jovian system.

In the last period, there has been the withdrawal of the NASA JEO contribution. This could bring to a re-formulation of the mission but in this paper will be presented the situation before the NASA withdrawal (at the moment the new scenario is still under investigation).

The EJSM mission will be launched in 2020 and it will reach Jupiter System after 6 years of flight. The circular orbit around Ganymede will start in 2028.

The basis of the laser altimeter measurements are Time-of-Flight (TOF) and backscattering laser pulse shape. TOF, that is the evaluation of elapsed time of short pulses for round-trip propagation at the speed-of-light between the space probe and the surface to be measured, permits to evaluate the distance between the spacecraft and the surface of Ganymede, allowing to obtain an altimetric map of the Galilean satellite. The analysis of backscattering laser pulse shape permits to determine Ganymede's superficial characteristics in terms of slope, roughness, and albedo. This technique provides accurate information on the surface profile allowing, within the characteristics of spacecraft orbit, global coverage of the planet surface.

Within the ambit of EJSM, laser altimeter on board of JGO spacecraft will contribute to the geodetic and geophysical characterization of Ganymede and particularly of its surface, but the accurate range data, combined with additional information, will be essential also for studies of the spacecraft orbit in the gravity field of the moon. It will be possible to reconstruct topographic profiles along the ground track of the spacecraft and also to characterize tidal forces effect by measuring the temporal changes in the global shape of Ganymede due to them.

In particular, scientific objectives that make so important the use of this instrument are:

- derive geo-referenced topographic profiles and 3D models of geological features and different terrain types (regionally and locally);
- derive an accurate Ganymede's global shape model and reference figure;
- derive a geodetic network based on accurately measured positions of prominent topographic features;
- determine tidal deformation (tidal amplitudes as a function of Ganymede's orbital longitude);
- measure topography to interpret the gravity signal (possible mass anomalies);
- assist in studies of geodetic parameters (coordinate system, rotation, librations);
- assist in orbit determination and gravity data modeling;
- measure surface roughness, local slopes and albedo variations.

The laser design is on the heritage of BepiColombo Laser Altimeter (*BELA)* in the environment of Jovian system. Laser emitter of Laser Altimeter for *JGO* would have, preferably, an adjustable pulse output energy for allowing operations at Callisto, which has an albedo (0,19) much less than Ganymede's (0,43), and for giving much options for power efficient laser operations at Ganymede. The pointing shall be accurate to within the size of the laser footprint and both night-time and daytime observations are equally possible (in the second case it is fundamental to overcome the solar background noise). In scenario in which *JGO* will operate, laser altimeter will not find limitations due to the presence of an atmosphere (like in terrestrial applications), which attenuates backscattered power for absorption. Moreover, because of the highly radiative environment which *JGO* will encounter, a critical issue for the instrument is surely its limit to radiation tolerance and in particular the knowledge of least tolerant components. Elements like receiver diode, laser pump diodes, laser rod, Q-switch, optical coatings for example are rad-hard only up to 60 − 100 krad (according to the value of radiation dose foreseen for the mission).

In this work, after a review of the main topographic features of Ganymede and the potential targets of laser altimeter on the planet, the orbital propagator and the performance model are presented. Then the results of the

model will be presented taking into account the reference scenario of the mission.

2 Ganymede

Ganymede is a moon of Jupiter, the largest among the whole Solar System's satellites. It is planet-sized world and belongs with Io, Europa and Callisto to the group of the Galilean moons of Jupiter.

Ganymede's geology, geophysics, surface composition and evolution are rather complex to analyze. The satellite's density implies a bulk composition that is about 60% rock and 40% water, which is mainly in the form of ice. Because of the distance of Jupiter System from the Sun, ice is stable under direct solar illumination thus water can stand only in a solid state. Some additional volatile ices such as ammonia may also be present. The abundance of ice, strongly reflecting incident radiation from the Sun, is the main reason of the relatively high value of Ganymede's geometric albedo equal to 0,43.

Ganymede is the only moon in the Solar System known to generate an internal magnetic field. Moreover, there are indications of an induced field component suggesting a deep internal ocean of briny water.

According to its dimension, Ganymede stands in a band of transition between geological active planets like Earth, Venus and Mars and the smaller bodies, getting cold too rapidly to permit a geological activity such as volcanism and tectonism with a "rejuvenation" of surface, deformation and fracture of external crust.

Images of Ganymede's surface show craters, furrows, dark and bright areas greatly differentiated and structures characteristic only of it. In particular, its surface consists of about 35% ancient dark terrain, remaining of an ancient crust which is relatively heavily cratered, and about 65% bright grooved terrain, which is more recent and heavily tectonized, both signs of a tumultuous past and of a geological activity.

Spectral study has shown that the bright terrain is ice-rich, while dark terrain contains a greater fraction of rocky material. Though its lithosphere is primarily water ice, the geology of bright terrain finds analogy with terrestrial rift zones. Impact features of different forms and ages are present mainly on dark terrain but also on the bright ones and sometimes can overlay or be crosscut by the groove systems. No significant relieves have been found on the surface with a maximum altitude estimated to be less than 1000 m.

Areas of heavily cratered dark terrain can have an apparent analogy with lunar *mari*, that are also them darker than *terrae* regarding the Moon, but in this case the darker character is index of more ancient terrains, also for the great number of impact craters that they present [2].

Differently from dark terrain, the main part of bright terrain does not extend on vast regions but forms across 2/3 of Ganymede's surface a global intricate network of large interconnected lanes tens to hundreds of kilometres wide, which separate the dark areas that appear like isolated patches.

Voyager observations before and those of Galileo after have shown continuous diffuse bright deposits of water frost without relieves, known as the polar caps, which mantle regions poleward of about ±40° latitudes [2].

A very great number of impact craters have been observed on Ganymede's surface on both types of terrain. Ganymede displays also the greatest diversity of primary impact morphologies on one planetary surface in the Solar System. Generally, ganymedian craters have a flatter and more subdued character than those on the Moon and Mercury, probably because on Ganymede large craters were unable to retain their topography over geological time in a relatively weak icy crust, which could flow and thereby soften the relief [2].

At our best knowledge data about the slope and roughness distribution on the Ganymede surface are not available in literature. For slope, it is just available the range of variation assumed (0°÷20°); therefore data will be parameterized in this range.

Regard the roughness, we refer to the works of *Gardner* [3] and *Harding* [4] which report tables with the values of roughness corresponding to several types of terrestrial terrains: we consider the terrestrial values of the

structures which have geomorphic characteristics similar to those expected on Ganymede surface. It has been set a mean value of roughness for low-to-mid latitudes zones and one for high latitudes polar zones: for polar regions, which extends over latitudes major than ±60°, a value of 1,0 m considering that these areas are quite totally covered by ice caps; for low-to-mid latitudes areas, where tectonics structures are more common in bright grooved terrains and they cover about 2/3 of Ganymede surface, a mean value in the range 1,2 – 4,7 , that is 3 m.

According to data available in literature, for dark terrains the albedo range is from 0,12 (the darkest area of surface) to 0,34 with a medium value of 0,25. For bright terrains, at our best knowledge no data are available in literature. The need information has been extrapolated considering Ganymede's mean albedo and the percentage of dark and bright terrains. In this way, the range for albedo is 0,34 ÷ 0,72 (with a mean of 0,53) for bright terrains.

3 Orbital Propagator

In this chapter it will be introduced the orbital propagator developed in MATLAB® environment.

The orbital propagator is intended for mission analysis of JGO, particularly for the operative stage of the mission that consists in 180 days in circular orbit around Ganymede.

The orbital propagator is based on the equations of the orbital dynamics. The equations of two-body motion are the simplest way to describe the orbital motion of a satellite. They are obtained from Newton's law of gravitation and Kepler's law of orbit motion under the hypothesis of point mass or mass with spherical symmetrical distribution. It can be expressed as in Eq. (1)

$$\frac{d^2\vec{r}}{dt^2} = -\frac{\mu}{r^3}\vec{r} \qquad (1)$$

where $\vec{r}$ is the position vector of the satellite evaluated from the centre of the primary body, and μ is the gravitational constant [5].

This equation can be used only as an approximation of the actual motion because there are a lot of perturbing forces such as planet gravity harmonics due to deviations from a perfect sphere, the mutual gravitational attractions among planets and/or natural satellites with sizable mass, atmospheric drag, radiation pressure, and tides. The equation, in this case, result

$$\frac{d^2\vec{r}}{dt^2} = -\frac{\mu}{r^3}\vec{r} + \overrightarrow{a_p} \qquad (2)$$

with $\overrightarrow{a_p}$ that includes all the perturbing accelerations [5].

The perturbing effects that have been considered in this orbital propagator are:

- the deviation from a sphere of Ganymede, within the limits of known data about its gravitational field; this influences the orbit with secular perturbations (right ascension of ascending node, argument of perifocus, and mean anomaly) and periodic perturbations (all orbital parameters excepted semi-major axis);
- third body perturbation that can not be neglected due to the proximity of Jupiter; this affects the orbits in the same manner of the previous perturbing effect.

Some of the equations governing the first perturbing effect are [5] [6]:

$$\dot{\Omega} = -\frac{3}{2}\frac{J_2 R^2}{p^2}\bar{n}\cos(i) \qquad (3)$$

$$\dot{\omega} = \frac{3}{2}\frac{J_2 R^2}{p^2}\bar{n}\left(2 - \frac{5}{2}\sin^2(i)\right) \qquad (4)$$

$$\bar{n} = \sqrt{\frac{\mu}{a_0^3}}\left[1 + \frac{3}{2}\frac{J_2 R^2}{p^2}\left(1 - \frac{3}{2}\sin^2(i)\right)\left(1 - e^2\right)^{\frac{1}{2}}\right] \qquad (5)$$

where $\dot{\Omega}$ is the rate of ascending node, $\dot{\omega}$ is the argument of perifocus rate, $\bar{n}$ is the orbit mean motion with J_2 correction, J_2 is the coefficient of the second zonal harmonic, R is the planet equatorial radius, i is the orbit inclination, μ is the gravitational constant, a_0 is the semi-major axis at epoch, e is the eccentricity, and p is the semilatus rectum.

Some of the equations that can be used to obtain third body effects are [6]

$$\dot{\Omega}_{sec} = -\frac{3\mu_3\left(1-e^2\right)^3\left(1+5e^2\right)\left[2-3\sin^2(i_3)\right]}{8r_3^3 n\left(1-e_3^2\right)^{3/2}}\cos(i) \quad (6)$$

$$\dot{\omega}_{sec} = \frac{3\mu_3\left(1-e^2\right)^3\left[2-3\sin^2(i_3)\right]}{16r_3^3 n\left(1-e_3^2\right)^{3/2}}\left\{4-5\sin^2(i)+5e^2\left(3-\frac{7}{2}\sin^2(i)\right)\right\} \quad (7)$$

where i_3 is the inclination with respect to central body equatorial plane.

Schematically the software is organized in the following conceptual macro-blocks:

- *orbital data input, check and time conversion* in which the user inserts the orbital parameters, initial and final dates and the step of temporal variation. This block verifies the correct introduction of inputs and provides the conversion from Gregorian to Julian;
- *orbital propagation* which allows to know the position and the velocity of the spacecraft in a particular reference system in a given instant of time;
- *search in Ganymede's virtual map* which returns the topographical features of the portion of Ganymede's surface visited by the spacecraft.

The tool allows calculating the position and the velocity of a spacecraft, moving along a known orbit, after a certain time starting from given initial conditions. Furthermore it permits to evaluate Sun and Jupiter eclipse duration, mean β angle, global coverage and revisit, percentage of Jovian and anti-Jovian coverage, percentage of day and night coverage.

From the knowledge of the orbit and the spacecraft's position on it at the initial time , the propagation algorithm returns the position and velocity vectors, the true anomaly and the altitude at the instant together with the longitude and latitude of the point of surface the spacecraft is pointing (along nadir-direction).

Since the scope is to analyze temporal profiles of laser altimeter performance during the entire mission, the outputs of the orbital propagator will be provided for a series of successive times.

The knowledge of the current values of longitude and latitude allows individuating, in each instant of time, the precise point of Ganymede's surface pointed by the laser altimeter along the nadir-direction. The software provides a map of the surface and it is possible to obtain a characterization of the lit-up area in terms of albedo, slope and roughness, fundamental quantities for a performance analysis of the instrument.

4 Laser altimeter performance model

The evaluation of laser altimeter performance is based on the concept that a portion of the reflected optical pulse is collected by the receiver telescope, focused into a detector, constituted, by an APD with an interferential filter centred on laser wavelength, and then analysed in width by an Electronic Unit.

We assume that the ground target is a lambertian surface and no specular reflections have been considered [7].

The main parameters useful to characterize the performances of a laser altimeter for topographical analysis of a planetary surface are:

- *received pulse width* (*ns*) representing how much the pulse has been spread in time due to the reflection by the surface;
- *probability of detection and probability of false alarm*
- *Signal-to-Noise Ratio* representing the ratio between the signal and noise power;
- *range accuracy* (*m*) representing the range error in the measurement.

4.1 Echo pulse width

The backscattered laser pulse, respect to the transmitted one, is distorted by the surface topographical variations within the laser footprint.

For this reason, backscattered pulse has a minor peak power since the pulse transmitted energy, unless small absorption losses, is spread along a major interval of time and has a more irregular shape which cannot be predicted *a priori* because it depends on surface profile and, as consequence of that, contains information about roughness, slope and albedo.

An analytic expression for the mean square detected pulse width adapted for nadir pointing is provided by *Gardner* [7]:

$$\sigma_p^2 = \left(\sigma_1^2 + \sigma_h^2\right) + \frac{4Var(\Delta\xi)}{c^2} + \frac{4R^2\tan^2\theta}{c^2}\left(\tan^2\theta + \tan^2 S\right) \qquad (8)$$

where σ_1 is the rms laser pulse width (s), $\sigma_h = \left(\tau/2\sqrt{2\ln 2}\right)$ the rms width of receiver impulse response (s) with τ low-pass filter impulse width, $Var(\Delta\xi)$ the variance (m^2) of surface roughness, c is the velocity of light in vacuum (m/s), R is the range to surface (m), θ the laser beam divergence angle (rad), S the surface slope (rad). The terms of this equation refer respectively to system effect, effect connected to the roughness of Ganymede surface, and effects due to beam curvature and slope of the surface.

Detected pulse with is essentially functions of three terms: system effects, roughness effects, and beam curvature and slope effects.

4.2 Probability of detection and probability of false alarm

This model predicts detection probability P_D and probability of false alarm P_{FA}, which is the probability that the receiver noise will exceed the threshold and generate a false measure of TOF. To evaluate P_D, P_{FA}, and the threshold the Neyman-Pearson method is used.

With the Neyman-Pearson method the P_{FA} is set to a fixed value and the threshold is consequently evaluated. This approach for estimation of P_D and P_{FA} is known as "near exact analysis".

The final expressions for P_{FA} and P_D are [7]:

$$P_{FA} = \frac{1}{2}\int_{\frac{n_b G}{\sigma_b}}^{+\infty} dz P\left(z|n_b\right) erfc\left(\frac{x_{th} - \sigma_b z - n_b G - I_s\tau/Ge}{\sigma}\right) \qquad (9)$$

$$P_D = \frac{1}{2}\int_{\frac{(n_s+n_b)G}{\sigma_{sb}}}^{+\infty} dz P\left(z|n_s+n_b\right) erfc\left(\frac{x_{th} - \sigma_{sb} z - \left(n_s+n_b\right)G - I_s\tau/Ge}{\sigma}\right) \qquad (10)$$

where $P\left(z|n_b\right)$ is the conditional density probability in presence of background photons (Eq. (11)), x_{th} is the threshold, n_b is the total number of background photons (Eq. (13)), G is the APD gain, I_S is the APD leakage noise current (A), τ is the filter response width (s), e is the electron charge, $P\left(z|n_s+n_b\right)$ is the conditional density probability in presence of signal photoelectrons and background photons, n_s is the number of signal photoelectrons generated per pulse (Eq. (17)). The following equations present the expression for terms of Eq. (9) and Eq. (10).

$$P\left(z|n_b\right) = \frac{1}{\sigma_b\sqrt{2\pi}\left[1+\frac{G(F-1)}{\sigma_b}z\right]^{3/2}} \exp\left[-\frac{z^2}{2\left(1+\frac{G(F-1)}{\sigma_b}z\right)}\right] \qquad (11)$$

$$\sigma_b = \sqrt{n_b G^2 F} \qquad (12)$$

$$n_b = \left(\lambda_{SUN} + \lambda_{GAN} + \lambda_{JUP} + \frac{I_D}{e}\right)\cdot\tau \qquad (13)$$

$$\sigma^2 = \frac{I_a^2\tau}{e^2} + \frac{I_s\tau}{eG} \qquad (14)$$

$$P(z|n_b+n_s)=\frac{1}{\sigma_{sb}\sqrt{2\pi}\left[1+\frac{G(F-1)}{\sigma_{sb}}z\right]^{3/2}}\exp\left[-\frac{z^2}{2\left(1+\frac{G(F-1)}{\sigma_{sb}}z\right)}\right] \quad (15)$$

$$\sigma_{sb}=\sqrt{(n_s+n_b)G^2F} \quad (16)$$

$$n_s=\frac{E_t}{h\nu}\cdot\frac{\rho}{\pi}\cdot\frac{A_R}{R^2}\eta_{RX}\eta_{APD}\eta_{TX}\eta_f \quad (17)$$

where λ_{SUN} is the number of solar background photoelectron per second (s^{-1}), λ_{GAN} is the Ganymede radiation photon rate (s^{-1}), λ_{JUP} is the number of background photoelectrons per second due to Jupiter's emission (s^{-1}), I_D is the APD dark current (A), I_a is preamplifier noise current ($A/Hz^{1/2}$), $F=k_{eff}G+(2-1/G)(1-k_{eff})$ is the excess noise factor with k_{eff} ionization coefficient ratio, E_t is the transmitted laser pulse energy (J), $h\nu$ is the photon energy with h the Plank's constant (J·s) and ν the photon frequency (Hz), ρ is the surface albedo, A_R is the receiving area (m^2), R is the range from target, η_{RX} is the receiver transmission, η_{APD} is the APD quantum efficiency, η_{TX} is the optics transmittance, η_f is the interference filter transmittance.

4.3 Signal-to-Noise Ratio

An adequate Signal-to-Noise Ratio (SNR) is fundamental in order to obtain high-quality ranging data, then altimeter performances analysis has to verify its values and how they influence the ranging resolution.

In particular, laser altimeter signal strength depends on laser pulse power backscattered from the target surface and collected by the receiver telescope. Processes like optical shot noise, background noise, detector noise and pre-amplifier noise reduces the capacity to discriminate the received pulse from background and electronic noise.

Evaluating these components of noise, it is possible to compute SNR.

The laser altimeter receiver subsystem is formed by a telescope, an APD detector and an interference filter centred on laser wavelength. Assuming that the fluctuation of background noise intensity is not at frequencies falling within the filter bandwidth, the signal-to-noise ratio at the output of the filter (for a nadir-pointing altimeter) can be written as follows [8]:

$$SNR=\frac{G^2(en_S/\sigma_p)^2}{2eBFG^2\left[e(\lambda_{SUN}+\lambda_{GAN}+\lambda_{JUP})+en_S/\sigma_p\right]+2eBI_D+I_A^2B} \quad (18)$$

where σ_p is the received pulse width (s) (Eq.(8)), B is the detection bandwidth (Hz).

4.4 Ranging resolution

The ranging measurement resolution for a laser altimeter operating with the *TOF* single shot technique results inversely proportional to *SNR* due to the timing uncertainty with a fixed threshold level and directly proportional to the pulse spreading, according to the following relation [8]:

$$R_v=\frac{c\sigma_p}{2SNR} \quad (19)$$

where σ_p is the received pulse width given by Eq. (8), and *SNR* is the Signal-to-Noise Ratio given by Eq. (18).

5 Laser altimeter performance analysis

In this paragraph the most significant results obtained implementing the software described in previous paragraphs will be presented.

The laser altimeter for *JGO* will operate during the *Ganymede Science Phase* and particularly during the circular sub-phase, more suitable than the elliptic one for topographic studies of the surface. Indeed, the low altitude (200-500 km) and very high inclination (87.5°) of the circular orbit guarantees a good coverage in a relatively short time and the possibility to

investigate the surface from a very reduced distance then with a major accuracy.
Before to analyse the laser altimeter performance, let we make some considerations on the orbit results.
From the orbital propagator it results that the coverage in longitude is practically total, while that in latitude is limited by the inclination of the orbit, then areas with latitudes major than ±87.5° are not interested by observations. This can be seen in Fig. 1.

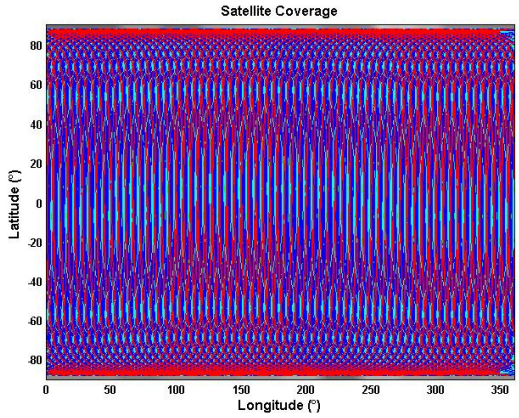


Fig. 1 Satellite Coverage map

For this orbit, the altitude remains constant with time, with important consequences for performance which will not be influenced by the changing in satellite height. Moreover, also the velocity doesn't change during the orbit and the mission, keeping always tangential, resulting in an uniform horizontal mapping accuracy.

Given the physical features (albedo, roughness and slope) of the portions of surface flown over during the chosen interval of time and through the instrument and mission parameters, performance are evaluated for each instant of time (for each taken sample) thanks to the current values related to it. In particular, for each time, the roughness is fixed while the slope is variable with a step set by the user, so that, for the same point, we have the values of performance for different values of slope within the interval of interest (0° ÷ 20°).

5.1 Performance in night-light conditions

For night light conditions the following parameter setting has been considered:

- temperature of Ganymede $T = 70K$ [13]
- solar spectral irradiance at Ganymede surface $I_{sun} = 0$
- surface reflectivity of 0.12
- false alarm rate n_{FA} of 0.01.

In order to verify in which conditions the laser altimeter satisfied *the requirements proposed* for detection probability P_D (95%) and ranging resolution R_v (10 cm), the Eq. (9), (10), (18), and (19) have been applied to evaluate P_D and R_v versus slope surface S, for different values of altitude (200-500 km) and at variance surface roughness of 3 m. the result are showed in Fig.2 and Fig. 4.

The graph in Fig.2 is obtained considering an ideal case where the low-pass filter impulse width τ is always equal to the echo pulse width and neglecting the receiver impulse response contribution.

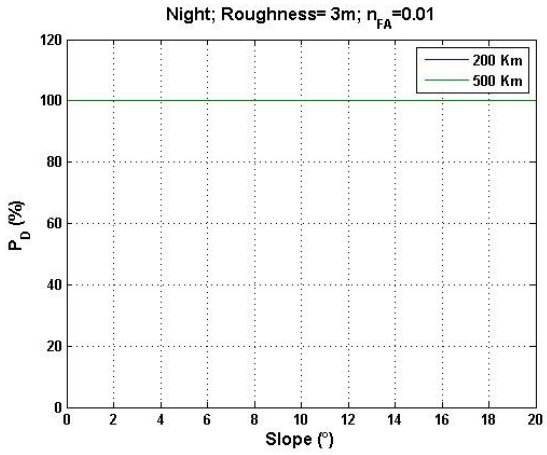


Fig. 2 Probability of Detection versus slope to surface for n_{FA}=0.01, roughness variance Var(Δξ)=3m and altitude of 200 and 500 km in nightlight conditions

From Fig.2, it is possible to see that the requirements for the detection probability are fulfilled because for both the altitudes P_D is higher than 95%.

Fig. 4 shows that the for an altitude of 200 km ranging resolution is always better than the threshold value of 10 cm *fixed by requirements* while for an altitude of 500 km this is true only for slope values lower than 2.9°.

The ranging resolution has been estimated from the value of the SNR using the expression (19). SNR versus slope is reported in Fig. 3.

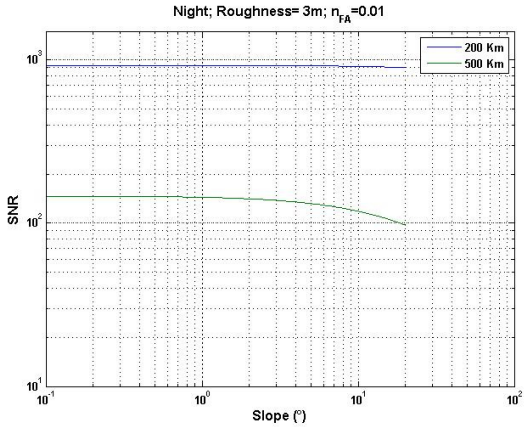


Fig. 3 Signal-to-Noise Ratio versus surface slope for different heights and roughness variance of 3m in nightlight conditions

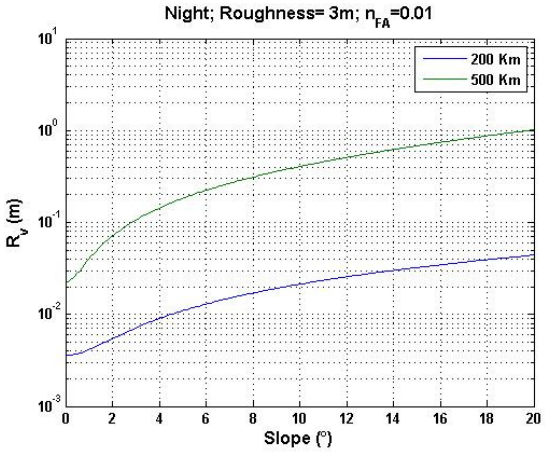


Fig. 4 Ranging resolution versus surface slope for different heights and roughness variance of 3m in nightlight conditions

5.2 Performance in daylight conditions

For daylight conditions the following parameter setting has been considered:

- day temperature of Ganymede $T = 152K$
- solar spectral irradiance at Ganymede surface $I_{sun} = 0.02355923 \mathrm{Wm}^{-2} nm^{-1}$
- surface reflectivity of 0.12
- false alarm rate n_{FA} equal to 0.01.

Also for daylight condition, the same method has been used to evaluate the detection probability and the ranging resolution and the results are showed in Fig. 5 and Fig. 7.

Fig. 5 shows that for both altitudes the P_D is always higher than 95% that is the value fixed as requirement for the probability of detection.

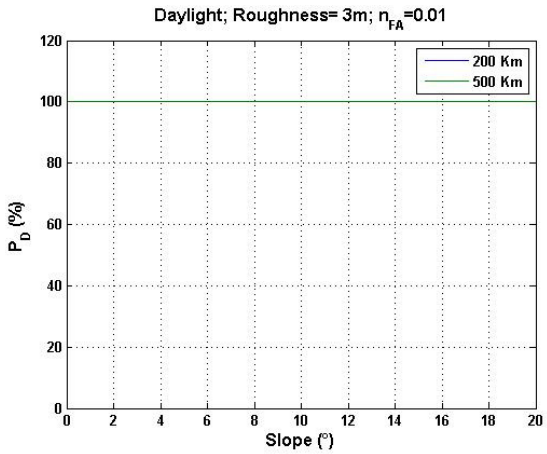


Fig. 5 Probability of Detection P_D versus slope to surface S for n_{FA}=0.01, roughness variance Var(Δξ)=3m and altitude of 200 and 500 km in daylight conditions

The results of Fig. 7 show that the requirement about ranging resolution is always fulfilled for an altitude of 200 km while it is fulfilled only for a slope lower than 2.8° for an altitude of 500 km. This very low difference from the night-light condition is due to the high distance from the sun that cause a weak solar irradiance.

The ranging resolution has been estimated from the value of the SNR using the expression (19). SNR versus slope is reported in Fig. 6.

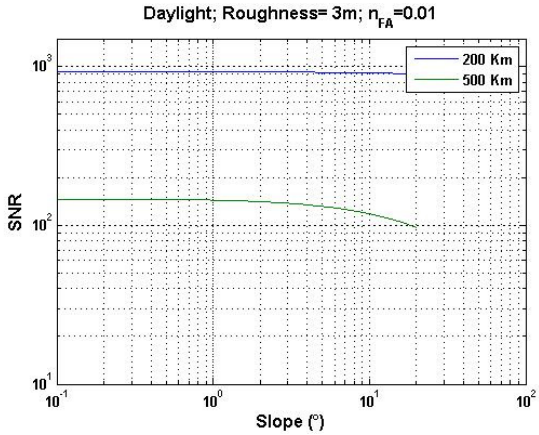


Fig. 6 Signal-to-Noise Ratio versus surface slope for different heights and roughness variance of 3m in daylight conditions

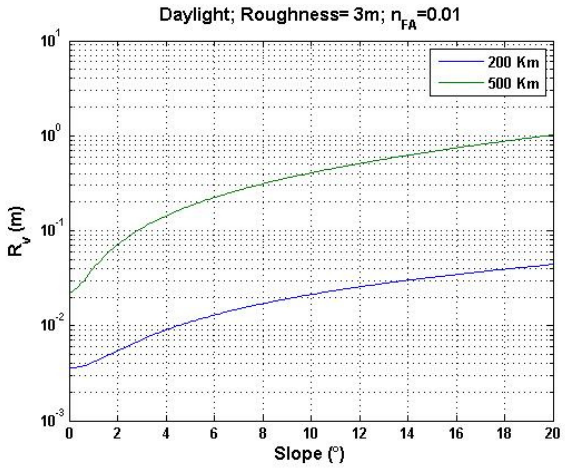


Fig. 7 Ranging resolution versus surface slope for different heights and roughness variance of 3m in daylight conditions

6 Conclusion

The performance for the laser altimeter which will board on *Jupiter Ganymede Orbiter* have been analysed by means a software developed in MATLAB®. The study has pointed out how instrument performances respect completely the requirements for a satellite altitude of 200 km while for orbit altitude of 500 km the ranging resolution falls in the requirements only for low slope angles.

The considered scenario appears much favourable for observations in the optic-near infrared spectrum. Indeed, the distance of Ganymede from the Sun reduces background noise at this frequencies and the absence of an atmosphere around Ganymede avoids absorptions of the transmitted light beam.

The possibility of *JGO* to collect a very high-quality altimetric data will improve in a significant way the knowledge of the topography of Ganymede.

Acknowledgments

This work has been funded by the Italian Space Agency (ASI) in the context of EJSM/Laplace mission.

References

[1] Joint Jupiter Science Definition Team, NASA/ESA Study Team, "EJSM NASA/ESA joint summary report", NASA/ESA, 20 Jan 2009.

[2] Pappalardo Robert T. et al., *Jupiter, The Planet, Satellites and Magnetosphere*, 1st ed., Cambridge University Press, Cambridge, 2004, Chap. 16.

[3] Gardner C. S., "Ranging performance of satellite laser altimeters", *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 30, No. 5, Sept 1992, pp. 1061-1072.

[4] Harding D.J., Bufton J. L., Frawley J. J., "Satellite Laser Altimetry of Terrestrial Topography: Vertical Accuracy as a Function of Surface Slope, Roughness and Cloud Cover", *IEEE Transaction on Geoscience and Remote* Sensing, Vol. 32, No. 2, March 1994, pp. 329-339.

[5] Chobotov Vladimir A., *Orbital Mechanics*, 3rd ed., American Institute of Aeronautics and Astronautics, 1801 Alexander Bell Drive, Reston, Virginia, 2002, Chaps. 3, 8, 9, 10.

[6] Vallado David A., *Fundamentals of Astrodynamics and Applications*, 3rd ed., Microcosm Press, Hawthorne, CA, Springer, New York, NY, 2007, Chaps. 2, 3, 8, 9.

[7] Santovito Maria Rosaria et al., "A laser altimeter for BepiColombo Mission: Instrument design and performance model", *Planetary and Space Science*, Vol. 54, No. 7, July 2006, pp. 645-660.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Preliminary Orbit Determination: Uncertainty Management via Taylor Expansions

R. Armellin, P. Di Lizia, M. Lavagna
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano
Via La Masa 34, 20156, Milano, Italy

Keywords: Preliminary orbit determination, Differential algebra, Uncertainty mapping, High-order methods

Abstract

A method for high-order treatment of uncertainties in preliminary orbit determination is presented. The observations consist in three couples of topocentric right ascensions and declinations at three observation epochs. The goal of preliminary orbit determination is to compute a trajectory that fits with the observations in two-body dynamics. The uncertainties of the observations are usually mapped to the phase space only when additional observations are available and a least square fitting problem is set-up. A method based on Taylor differential algebra for the analytical treatment of observation uncertainties is implemented. Taylor differential algebra allows for the efficient computation of the arbitrary order Taylor expansion of a sufficiently continuos multivariate function. This enables the mapping of the uncertainties from the observation space to the phase space as high-order multivariate Taylor polynomials. These maps can then be propagated forward in time to predict the observable set at successive epochs. This method can be suitably used to recover newly discovered objects when a scarce number of measurements is available. Real asteroids with simulated topocentric observations are used to show the performances of the method.

1 Introduction

Angles-only preliminary orbit determination (OD) of celestial body is an old problem. Gauss' [11] and Laplace's [13] methods are commonly used to determine a Keplerian orbit that fits with three astrometric observations. These two methods have been revisited and analyzed by a large number of authors [14; 16; 7; 15] and new ones introduced more recently also tailored for artificial satellites. The Double r-iteration technique of Escobal [10] and the approach of Gooding [12] are two examples of angles-only methods. When more than three observations become available a least squares method is implemented to compute a refined solution in which the two-body approximation is lifted. This solution is optimal in the sense of minimizing the observation residuals. As an additional result, the covariance matrix of the measurements is mapped in the asteroids orbital elements. The knowledge of the so-called confidence region is of fundamental importance when the computation the impact probability of a near Earth object with our planet is of interest [18]. The uncertainties associated to the observations is usually propagated forward in time by the state transition matrix or Monte Carlo simulations [8; 17].

For newly discovered asteroids studying the effect of observations uncertainties in preliminary OD can be crucial, as it strongly impacts

on the capability of recovering the body at the sky in future epochs. A high-order method for 1) solving the preliminary OD problem, 2) analytically mapping the observation uncertainties into the phase space, and 3) propagating the uncertainty set forward in time is presented. The differential algebra (DA) implemented in COSY-Infinity [5] is the tool that enables the method. Differential algebra allows for the automatic computation of the Taylor coefficients of a function up to a specified order k, along with the function evaluation, with a fixed amount of effort. Available algorithms for polynomial composition, inversion, and evaluation, enable the expansion of the solution of implicit equations with respect to an arbitrary set of parameters. The method presented in this paper is based on a DA version of Lambert's and Kepler's problems, in which DA is used for the expansion of the solution of implicit equations with respect to observation uncertainties. A high-order iterative algorithm is then implemented to find the exact solution to the preliminary OD problem and to analytically map the uncertainties from the observation to the phase space. This set can be propagated forward in time using either the DA version of Kepler's problem or in an n-body dynamical model adopting a DA integration scheme, as already shown in [1].

2 Notes on Differential Algebra

DA techniques, exploited here to obtain k-th order Taylor expansions of the solution of implicit equations, find their origin in the attempt to solve analytical problems by an algebraic approach [4]. Historically, the treatment of functions in numerics has been based on the treatment of numbers, and the classical numerical algorithms are based on the mere evaluation of functions at specific points. DA techniques rely on the observation that it is possible to extract more information on a function rather than its mere values. The basic idea is to bring the treatment of functions and the operations on them to the computer environment in a similar way as the treatment of real numbers. The straightforward implementation of differential algebra in a computer allows to compute the Taylor coefficients of a function up to a specified order k, along with the function evaluation, with a fixed amount of effort. The Taylor coefficients of order k for sums and product of functions, as well as scalar products with reals, can be computed from those of summands and factors; therefore, the set of equivalence classes of functions can be endowed with well-defined operations, leading to the so-called truncated power series algebra [3]. Similarly to the algorithms for floating point arithmetic, the algorithms for functions followed, including methods to perform composition of functions, to invert them, to solve nonlinear systems explicitly, and to treat common elementary functions. In addition to these algebraic operations, the DA framework is endowed with differentiation and integration operators, therefore finalizing the definition of the DA structure. The differential algebra sketched in this section was implemented in the software COSY-Infinity [5].

3 Iterative Correction to the Preliminary Orbit

A high-order DA-based method for the iterative correction of the preliminary orbit is presented. The iterative procedure has the goal of both refining the approximate solution obtained by Gauss' 8-th degree polynomial [9] and providing the functional dependence of the solution of the preliminary OD problem with respect to observation uncertainties in terms of a high-order Taylor polynomial. We start by initializing the observations as DA variables by

$$\begin{aligned}[\boldsymbol{\alpha}] &= \boldsymbol{\alpha} + \delta\boldsymbol{\alpha} \\ [\boldsymbol{\delta}] &= \boldsymbol{\delta} + \delta\boldsymbol{\delta}.\end{aligned} \tag{1}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ are the vector of measured right ascensions and declinations (the observations are performed at time t_1, t_2, and t_3). The square brackets indicate Taylor polynomials and $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\delta}$ accounts for measurement uncertainties. The line of sight vectors at t_2 and t_3 are

$$\begin{aligned}[\hat{\boldsymbol{\rho}}_2] &= \hat{\boldsymbol{\rho}}_2 + \mathcal{M}_{\hat{\boldsymbol{\rho}}_2}(\delta\alpha_2, \delta\delta_2) \\ [\hat{\boldsymbol{\rho}}_3] &= \hat{\boldsymbol{\rho}}_3 + \mathcal{M}_{\hat{\boldsymbol{\rho}}_3}(\delta\alpha_3, \delta\delta_3).\end{aligned} \tag{2}$$

with

$$\hat{\boldsymbol{\rho}}_i = (\cos\delta_i \cos\alpha_i, \cos\delta_i \sin\alpha_i, \sin\delta_i)^T. \tag{3}$$

Similarly, we initialize DA variables on the slant ranges at t_2 and t_3

$$\begin{array}{l}[\rho_2]^{1^-} = \rho_2^{1^-} + \delta\rho_2 \\ [\rho_3]^{1^-} = \rho_3^{1^-} + \delta\rho_3,\end{array} \tag{4}$$

where the superscript 1^- indicates the first step of the iterative procedure, and $\rho_2^{1^-}$ and $\rho_3^{1^-}$ are the starting values associated to the solution of Gauss's 8-th degree polynomial.

The DA representation of asteroid's state at t_2 and t_3 is

$$\begin{array}{l}[\boldsymbol{r}_2] = \boldsymbol{r}_2 + \mathcal{M}_{\boldsymbol{r}_2}(\delta\alpha_2, \delta\delta_2, \delta\rho_2) \\ [\boldsymbol{r}_3] = \boldsymbol{r}_3 + \mathcal{M}_{\boldsymbol{r}_3}(\delta\alpha_3, \delta\delta_3, \delta\rho_3).\end{array} \tag{5}$$

Having $[\boldsymbol{r}_2]$ and $[\boldsymbol{r}_3]$ and the time of flight τ_3 a Lambert's problem can be set-up. The implicit equation that appears in the Lambert problem is solved by the algorithm described in [2], in which the vector of parameters is given by $(\alpha_2, \alpha_3, \delta_2, \delta_3, \rho_2, \rho_3)$. The result is the Taylor polynomial of asteroid's velocity at t_2

$$[\boldsymbol{v}_2] = \boldsymbol{v}_2 + \mathcal{M}_{\boldsymbol{v}_2}(\delta\alpha_2, \delta\delta_2, \delta\alpha_3, \delta\delta_3, \delta\rho_2, \delta\rho_3). \tag{6}$$

The state at t_2 is fully available from the first line of (5) and (6), and it can be propagated analytically backward in time up to t_1 by solving the DA-based Kepler's problem [2]. The position vector of the object at t_1 is then

$$[\boldsymbol{r}_1] = \boldsymbol{r}_1 + \mathcal{M}_{\boldsymbol{r}_1}(\delta\alpha_2, \delta\delta_2, \delta\alpha_3, \delta\delta_3, \delta\rho_2, \delta\rho_3). \tag{7}$$

Finally, the observables associated to asteroid's position vector at time t_1 are computed, yielding

$$\begin{array}{l}[\tilde{\alpha}_1] = \tilde{\alpha_1} + \mathcal{M}_{\tilde{\alpha}_1}(\delta\alpha_2, \delta\delta_2, \delta\alpha_3, \delta\delta_3, \delta\rho_2, \delta\rho_3) \\ [\tilde{\delta}_1] = \tilde{\delta_1} + \mathcal{M}_{\tilde{\delta}_1}(\delta\alpha_2, \delta\delta_2, \delta\alpha_3, \delta\delta_3, \delta\rho_2, \delta\rho_3).\end{array} \tag{8}$$

Note that a tilde is used to indicate that the constant parts of these polynomials are not the exact right ascension and declination at t_1, as the slant ranges associated to the solution of Gauss's 8-th degree polynomial have been used to initialize the procedure.

We now want to find the exact values of ρ_2 and ρ_3 and the high-order Taylor expansion of asteroid's state at t_2 with respect to observation uncertainties. We start by subtracting (1) to (8)

$$\begin{array}{rcl}[\Delta\tilde{\alpha}_1] &=& [\tilde{\alpha_1}] - [\alpha_1] = \Delta\tilde{\alpha}_1 + \\ &+& \mathcal{M}_{\Delta\tilde{\alpha}_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_2, \delta\rho_3) \\ [\Delta\tilde{\delta}_1] &=& [\tilde{\delta_1}] - [\delta_1] = \Delta\tilde{\delta}_1 + \\ &+& \mathcal{M}_{\Delta\tilde{\delta}_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_2, \delta\rho_3),\end{array} \tag{9}$$

which is an high-order Taylor map representing the observations defects. Note that for the exact values of ρ_2 and ρ_3 the constant part of maps (9) would be zero. We now need a) to find the variations $\delta\rho_2$ and $\delta\rho_3$ necessary to cancel out these constants and b) to express $\boldsymbol{r}_2$ and $\boldsymbol{v}_2$ as Taylor polynomials in $\delta\boldsymbol{\alpha}$ and $\delta\boldsymbol{\delta}$ only.

The first step is to work with an origin preserving map

$$\begin{array}{rcl}[\Delta\alpha_1] &=& [\Delta\tilde{\alpha}_1] - \Delta\tilde{\alpha}_1 = \\ &=& \mathcal{M}_{\Delta\alpha_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_2, \delta\rho_3) \\ [\Delta\delta_1] &=& [\Delta\tilde{\delta}_1] - \Delta\tilde{\delta}_1 = \\ &=& \mathcal{M}_{\Delta\delta_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_2, \delta\rho_3)\end{array} \tag{10}$$

and to build an augmented Taylor polynomial by adding identities in observation deltas

$$\begin{bmatrix}\Delta\alpha_1 \\ \Delta\delta_1 \\ \delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\delta}\end{bmatrix} = \begin{bmatrix}\mathcal{M}_{\Delta\alpha_1} \\ \mathcal{M}_{\Delta\delta_1} \\ \mathcal{I}_{\boldsymbol{\alpha}} \\ \mathcal{I}_{\boldsymbol{\delta}}\end{bmatrix} \begin{bmatrix}\delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\delta} \\ \delta\rho_2 \\ \delta\rho_3\end{bmatrix}. \tag{11}$$

This polynomial map can be inverted using ad-hoc algorithms implemented in COSY-Infinity, yielding

$$\begin{bmatrix}\delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\delta} \\ \delta\rho_2 \\ \delta\rho_3\end{bmatrix} = \begin{bmatrix}\mathcal{M}_{\Delta\alpha_1} \\ \mathcal{M}_{\Delta\delta_1} \\ \mathcal{I}_{\boldsymbol{\alpha}} \\ \mathcal{I}_{\boldsymbol{\delta}}\end{bmatrix}^{-1} \begin{bmatrix}\Delta\alpha_1 \\ \Delta\delta_1 \\ \delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\delta}\end{bmatrix}. \tag{12}$$

Extracting the two last lines we obtain

$$\begin{bmatrix}\delta\rho_2 \\ \delta\rho_3\end{bmatrix} = \begin{bmatrix}\mathcal{M}_{\rho_2} \\ \mathcal{M}_{\rho_3}\end{bmatrix} \begin{bmatrix}\Delta\alpha_1 \\ \Delta\delta_1 \\ \delta\boldsymbol{\alpha} \\ \delta\boldsymbol{\delta}\end{bmatrix}. \tag{13}$$

We now evaluate the map (13) in $[\Delta\alpha_1] = -\Delta\tilde{\alpha_1}$ and $[\Delta\delta_1] = -\Delta\tilde{\delta_1}$, obtaining

$$\begin{array}{l}[\rho_2]^{1^+} = \rho_2^{1^+} + \mathcal{M}_{\rho_2}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\rho_3]^{1^+} = \rho_3^{1^+} + \mathcal{M}_{\rho_3}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}),\end{array} \tag{14}$$

where the subscript 1^+ indicates the Taylor polynomial of the corrections of the slant ranges to be applied at the end of the first iteration. Note that this step is the high-order counterpart of classical Newton's method.

The second iteration starts with the Taylor polynomials of the slant ranges given by

$$\begin{aligned}[\rho_2]^{2^-} &= [\rho_2]^{1^-} + [\rho_2]^{1^+} + \delta\rho_2 = \\ &= \rho_2^{2^-} + \mathcal{M}_{\rho_2}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_2) \\ [\rho_3]^{2^-} &= [\rho_3]^{1^-} + [\rho_3]^{1^+} + \delta\rho_3 = \\ &= \rho_3^{2^-} + \mathcal{M}_{\rho_3}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \delta\rho_3),\end{aligned} \tag{15}$$

where now the explicit dependence on the entire set of observations appears. Thus, from the second iteration, the Taylor polynomials (5)–(8) depend on all $(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}, \rho_2, \rho_3)$. The iterative procedure ends when the values of $\Delta\tilde{\alpha}_1$ and $\Delta\tilde{\delta}_1$ are smaller that a prescribed tolerance. The Taylor polynomials of the slant ranges at the last iteration k are

$$\begin{aligned}[\rho_2] &= [\rho_2]^{k^-} + [\rho_2]^{k^+} = \rho_2 + \mathcal{M}_{\rho_2}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\rho_3] &= [\rho_3]^{k^-} + [\rho_3]^{k^+} = \rho_3 + \mathcal{M}_{\rho_3}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}).\end{aligned} \tag{16}$$

Using these expressions the asteroid position vector and velocity at t_2 assume the form

$$\begin{aligned}[\boldsymbol{r}_2] &= \boldsymbol{r}_2 + \mathcal{M}_{\boldsymbol{r}_2}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\boldsymbol{v}_2] &= \boldsymbol{v}_2 + \mathcal{M}_{\boldsymbol{v}_2}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}).\end{aligned} \tag{17}$$

Note that, as a result of the iterative procedure, $\boldsymbol{r}_2$ and $\boldsymbol{v}_2$ exactly satisfy (in the two-body model) the nominal observation set $(\boldsymbol{\alpha}, \boldsymbol{\delta})$. Furthermore, for any displaced value of the observables, the solution of the preliminary determination problem is computed by evaluating the polynomial (17) in the corresponding values of $(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta})$.

4 Observable set propagation

The algorithm described in Sect. 3 represents the core of the method. The map (17) is a high-order analytical representation of the functional dependence between asteroid's observable set and its state vector at t_2. The analogous map at t_1 and t_2 is gained by the DA-based solution of a backward and forward Kepler's problem. The result is

$$\begin{aligned}[\boldsymbol{r}_1] &= \boldsymbol{r}_1 + \mathcal{M}_{\boldsymbol{r}_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\boldsymbol{v}_1] &= \boldsymbol{v}_1 + \mathcal{M}_{\boldsymbol{v}_1}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta})\end{aligned} \tag{18}$$

and

$$\begin{aligned}[\boldsymbol{r}_3] &= \boldsymbol{r}_3 + \mathcal{M}_{\boldsymbol{r}_3}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\boldsymbol{v}_3] &= \boldsymbol{v}_3 + \mathcal{M}_{\boldsymbol{v}_3}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}).\end{aligned} \tag{19}$$

Furthermore, in a similar way, the asteroid state can be propagated to an arbitrary epoch t_4, obtaining

$$\begin{aligned}[\boldsymbol{r}_4] &= \boldsymbol{r}_4 + \mathcal{M}_{\boldsymbol{r}_4}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\boldsymbol{v}_4] &= \boldsymbol{v}_4 + \mathcal{M}_{\boldsymbol{v}_4}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}).\end{aligned} \tag{20}$$

In principle the propagation of the asteroid can be performed using an arbitrary dynamical model including all the sources of perturbations that may be relevant on asteroid motion, as shown in [1]. Once the asteroid state at the arbitrary epoch t_4 is available, the high-order Taylor map between the observables at t_4 and the observations uncertainties is obtained as

$$\begin{aligned}[\alpha_4] &= \alpha_4 + \mathcal{M}_{\alpha_4}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}) \\ [\delta_4] &= \delta_4 + \mathcal{M}_{\delta_4}(\delta\boldsymbol{\alpha}, \delta\boldsymbol{\delta}).\end{aligned} \tag{21}$$

The Taylor polynomials (21) analytically describe how the measurement uncertainties $(\boldsymbol{\alpha}, \boldsymbol{\delta})$ are mapped into the set of observables at an arbitrary epoch t_4. Thus, Eq. (21) analytically describes the region of the sky to be observed to recover the asteroid, for given estimates of observation accuracies.

5 Expansion Accuracy Analysis

The algorithms introduced in Sec. 3 and 4 expand the solution of preliminary OD and propagate the observable set as Taylor polynomials of order k in the measurement uncertainties. The convergence radius of these Taylor polynomial maps depends on both the expansion order k, the uncertainty values, and the problem at hand (i.e., asteroid's orbital parameters, asteroid–observer relative configuration, and time intervals between observations). To analyze these aspects the observation of two asteroids are considered in Sec. 6. For the first one the observation is almost in opposition, whereas

for the second a condition close to quadrature is analyzed. For each asteroid two different sets of observation epochs and different values of measurements accuracies are examined.

The accuracy analysis of the Taylor expansions is performed by comparing the outputs of two Monte Carlo runs, one based on classical pointwise Gauss' method and the other on the DA-based algorithm presented in Sec. 3. The procedure is:

1. Define the test case: asteroid, observatory, measurements statistics, epoch of the three observations, and the date of the fourth planned observation t_4.
2. Compute the nominal observation set, i.e. $(\boldsymbol{\alpha}, \boldsymbol{\delta})$.
3. Perform a single DA-based OD algorithm run and propagate the solution up to t_4, obtaining the maps (17), (18), (19), and (20).
4. Generate N "pointwise" observation deviations compatible with measurements statistics, $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}$.
5. Evaluate maps (17), (18), and (19) in $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}$, and compute the associated "differential algebra" deviations $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{DA}$.
6. Evaluate map (20) in $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}$, obtaining $\boldsymbol{r}_{4,i}^{DA}$.
7. For each $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}$ solve the OD problem with pointwise Gauss' method and propagate each solution up to t_4 by solving N pointwise Kepler's problems, obtaining $\boldsymbol{r}_{4,i}^{PW}$.
8. Compute the defects $\epsilon_{\boldsymbol{\alpha},\boldsymbol{\delta}} = \max_{i=1,N} ||(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{DA} - (\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}||$ and $\epsilon_{\boldsymbol{r}_4} = \max_{i=1,N} ||\boldsymbol{r}_{4,i}^{DA} - \boldsymbol{r}_{4,i}^{PW}||$.

The accuracy of the DA-based OD algorithm is evaluated by $\epsilon_{\boldsymbol{\alpha},\boldsymbol{\delta}}$. The accuracy of the Taylor expansion of asteroid's state at t_4 is assessed by $\epsilon_{\boldsymbol{r}_4}$. It is worth mentioning that Gauss' method is assumed to deliver the exact solution for all $(\delta\boldsymbol{\alpha}_i, \delta\boldsymbol{\delta}_i)^{PW}$. This is true as far as the iterative Gauss procedure converges for all considered deviations. Furthermore, $\epsilon_{\boldsymbol{r}_4}$ includes the expansion errors of both the DA-based OD algorithm and the DA expansion of Kepler's problem necessary for the propagation of the solution up to t_4.

6 Test Cases

The preliminary OD of two known asteroids is analyzed. For the two asteroids, the simulation considers observations made by the observatory of the official discovery. Asteroid J95E01K was discovered at Oizumi observatory and Asteroid K09S19T at Santa Maria de Montmagastrell. Observation dates close to those of the discovery are considered. These dates are slightly modified to reproduce an observation close to opposition for Asteroid J95E01K and to quadrature for Asteroid K09S19T. Two test cases are presented for each asteroid. In the first one $t_2 - t_1 = t_3 - t_2 = 1$ day (labeled as case A), whereas in the second one $t_2 - t_1 = 2$ hr and $t_3 - t_1 = 1$ day (labeled as case B).

6.1 Asteroid J95E01K

For case A, Gauss' 8-th degree polynomial has three real positive roots. The DA-based preliminary OD algorithm converges in two iterations to the true solution starting from $r_2 = 1.41701$ AU (Note that with the same guess 52 iterations are necessary for the pointwise Guass' method described in [9].)

Figures 1 and 2 show the accuracy of the DA-based algorithm for the solution of the preliminary OD and the successive propagation up to $t_4 = 30$ days, respectively. The procedure described in Sec. 5 is applied to assess the accuracy of the expansion for observations standard deviations $\sigma = 0.15, 0.3, 0.45$ arcsec (these values are compatible with those collected by Bykov et al. [6]). A same value of standard deviation is used for declination and right ascension angles. A 6-th order computation is highly accurate for both the solution of the OD problem and the propagation of the asteroid state up to 30 days. This value is chosen as it is the observation window from the Earth.

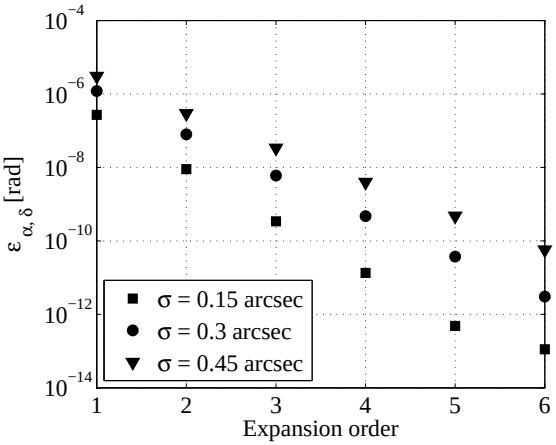


Figure 1: Accuracy of the DA-based preliminary OD algorithm as function of order and observation uncertainties.

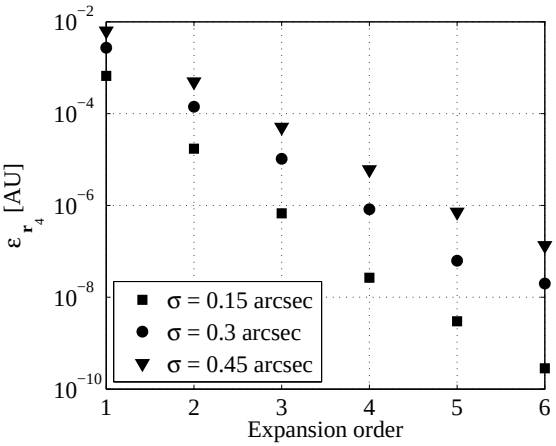


Figure 2: Accuracy of the DA-based propagation (30 days) algorithm as function of expansion order and observation uncertainties.

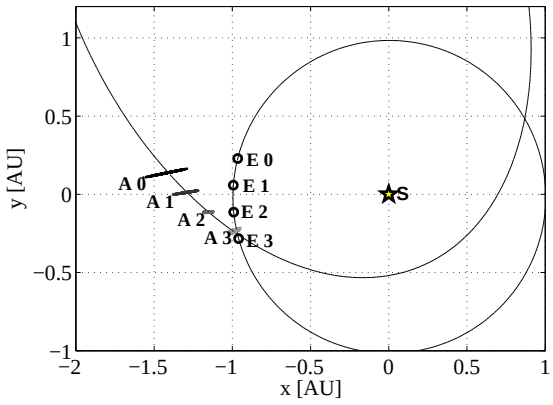


Figure 3: Propagated asteroid position sets. Sixth order expansion and propagation up to 30 days for $\sigma = 0.3$ arcsec.

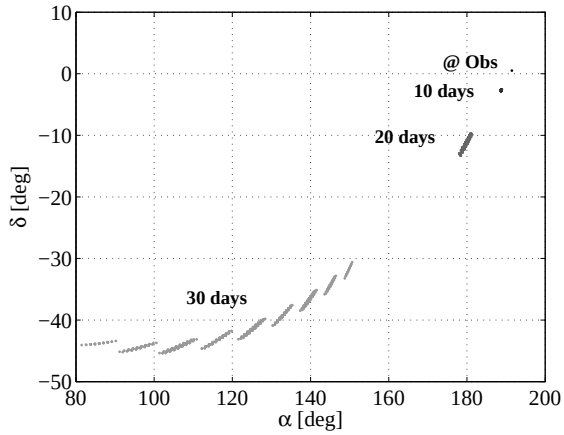


Figure 4: Propagated observable sets. Sixth order expansion and propagation up to 30 days for $\sigma = 0.3$ arcsec.

In Figure 3 and 4 the Taylor polynomials of the solution and the forward propagation are plotted, for $\sigma = 0.3$ arcsec. The polynomials are plotted by evaluating each variable in $-3\sigma, 0, 3\sigma$ arcsec, thus 3^6 polynomial evaluations are necessary. Note that the solution is plotted every 10 days up to the final value $t_4 = 30$ day (the label 0 is associated to the central observation; labels 1, 2, and 3 are separated by 10 days).

Figure 5 and 6 report the accuracy analysis for case B. In this case the Gauss 8-th degree polynomial has only one positive real root, $r_2 = 1.415469$. The DA-based algorithm converges in 2 iterations, whereas the pointwise Gauss method takes 25 iterations. Due to closer observations, the solution is much more sensitive to the observation accuracy. As a result, a maximum value of $\sigma = 0.3$ arcsec can be managed, requiring a seventh order expansion to achieve sufficient accuracy. The convergence radius of the Taylor expansion can be estimated by inspecting the coefficients of the Taylor expansion. This can be done in two steps. First, the maximum absolute value among the coefficients of monomials of the same order is identified. Then, an exponential fitting of these values is computed with respect to the expansion order. The con-

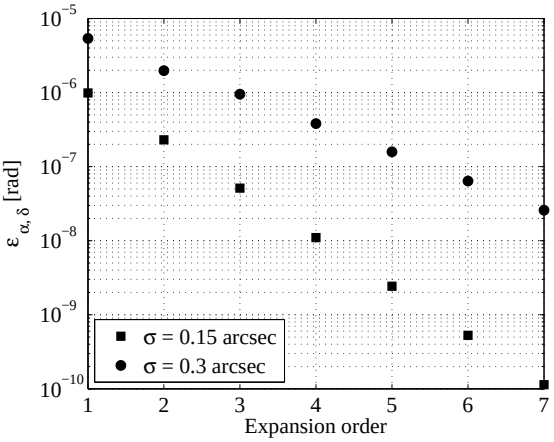


Figure 5: Accuracy of the DA-based preliminary OD algorithms as function of order and measurement uncertainties.

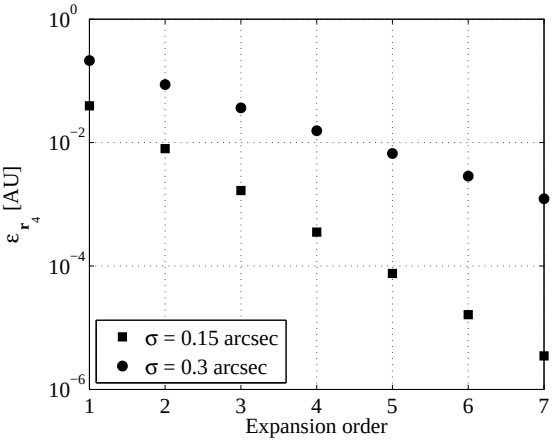


Figure 6: Accuracy of the DA-based propagation (30 days) algorithm as function of order and measurement uncertainties.

vergence radius is estimated by $r = \frac{c_{max}^{k-1}}{c_{\max}^{k}}$, where $c_{\max}^{k}$ is the maximum absolute value among the coefficients of monomials of order k, evaluated by the curve fitting. For this test case the estimate of the convergence radius is 1.3 arcsec, which is smaller than perturbations associated to $\sigma = 0.45$ arcsec.

In figures 7 and 8 the Taylor polynomials of the solution and the forward propagation are plotted, for $\sigma = 0.15$ arcsec. The solution is plotted at the central observation epoch t_2 and at the final value $t_4 = 30$ days. The increase in the size of the observable set is appreciable.

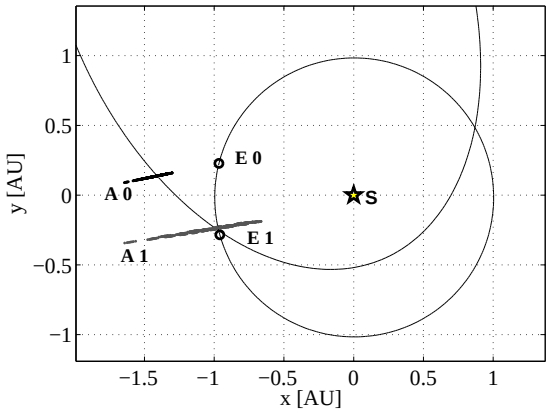


Figure 7: Propagated asteroid position sets. Sixth order expansion and propagation up to 30 days for $\sigma = 0.15$ arcsec.

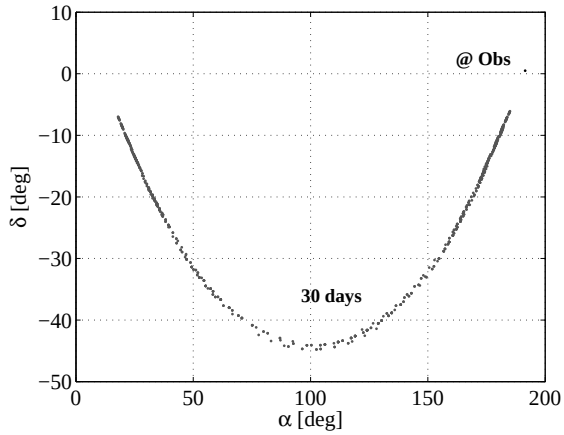


Figure 8: Propagated observable sets. Sixth order expansion and propagation up to 30 days for $\sigma = 0.15$ arcsec.

6.2 Asteroid K09S19T

For case A, Gauss' 8-th degree polynomial has three real positive roots. The DA-based preliminary OD algorithm converges to the true solution starting from the value $r_2 = 1.03270$ AU in two iterations (the poitnwise Gauss method in this case does not converge).

Figures 9 and 10 show the accuracy of the DA-based algorithm for the solution of the preliminary OD and the successive propagation up to $t_4 = 100$ days, respectively. Values of $\sigma = 0.15, 0.3, 0.45$ arcsec are considered to assess the accuracy of the expansion using the procedure

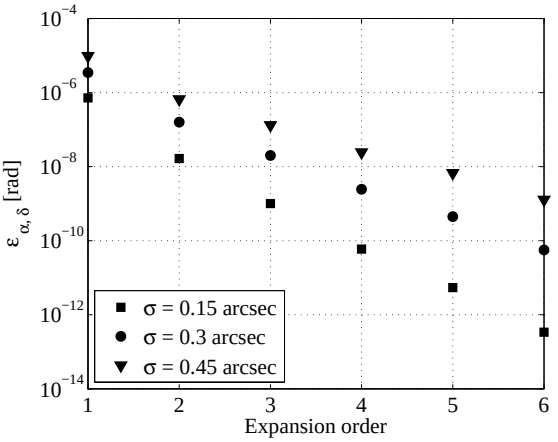


Figure 9: Accuracy of the DA-based preliminary OD algorithms as function of order and measurement uncertainties.

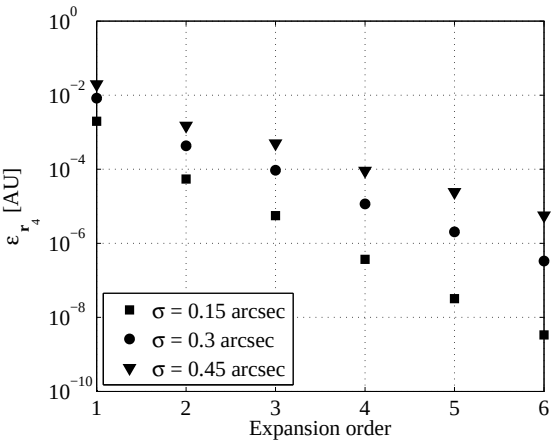


Figure 10: Accuracy of the DA-based propagation (100 days) algorithm as function of order and measurement uncertainties.

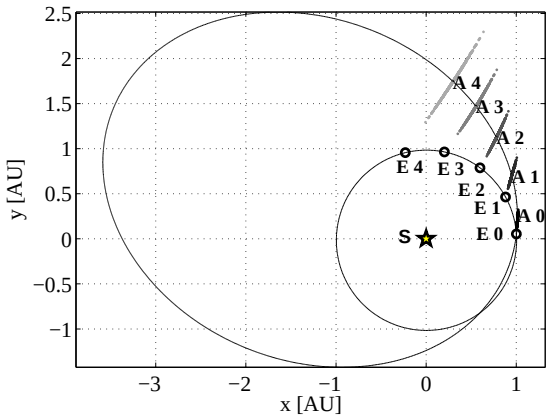


Figure 11: Propagated asteroid position sets. Sixth order expansion and propagation up to 100 days for $\sigma = 0.3$ arcsec

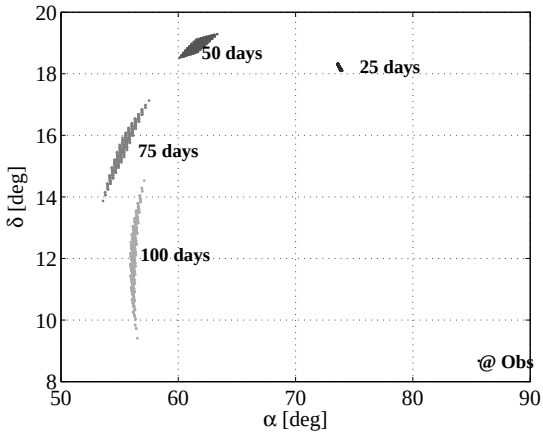


Figure 12: Propagated observable sets. Sixth order expansion and propagation up to 100 days for $\sigma = 0.3$ arcsec.

described in Sec. 5. Note that, as the pointwise Gauss algorithm does not compute the true solution, Figure 10 accounts for the accuracy of the DA expansion of Kepler's problem only. A 6-th order computation guarantees sufficient accuracy for both the solution of the preliminary OD problem and the propagation of asteroid state up to 100 days.

In Figure 11 and 12 the Taylor polynomials of the solution and the forward propagation are plotted for $\sigma = 0.3$ arcsec. The solution is plotted at the central observation epoch t_2 and at every 25 days up to $t_4 = 100$ days.

In Figures 13–16 analogous results are plotted for the case of $t_2 - t_1 = 2$ hr and $t_3 - t_1 = 1$ day.

7 Conclusions

A high-order method for solving angles-only preliminary OD problems and analytically mapping uncertainties from the observations to the phase space has been presented. The method relies on the tool of differential algebra, which allows for the automatic computation of the Taylor coefficients of a function up to a specified order k, along with the function evaluation. As the uncertainty set in the phase space can be nonlin-

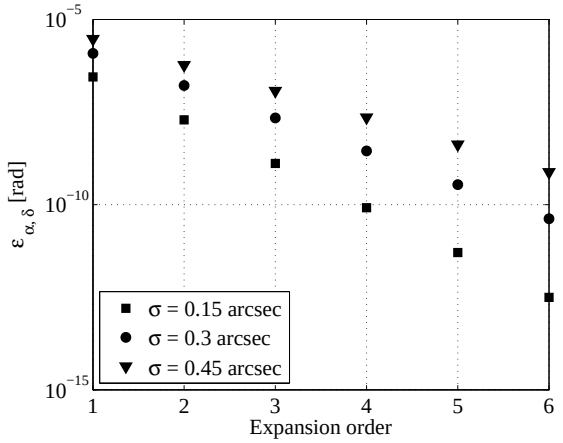


Figure 13: Accuracy of the DA-based preliminary OD algorithms as function of order and measurement uncertainties.

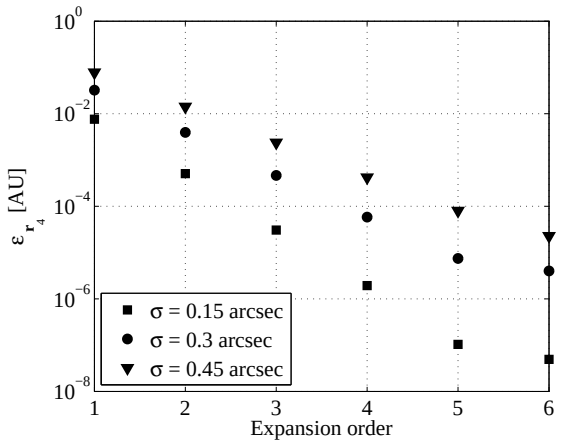


Figure 14: Accuracy of the DA-based propagation (100 days) algorithm as function of order and measurement uncertainties.

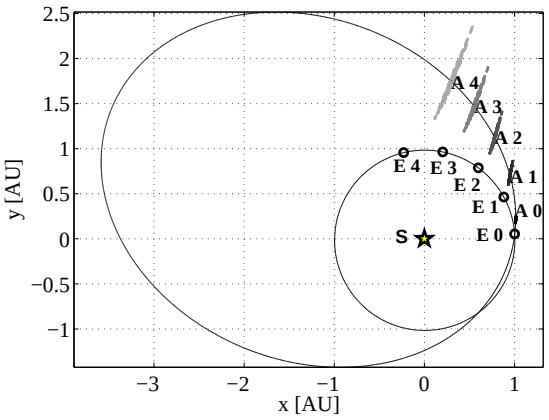


Figure 15: Propagated asteroid position sets. Sixth order expansion and propagation up to 100 days for $\sigma = 0.3$ arcsec

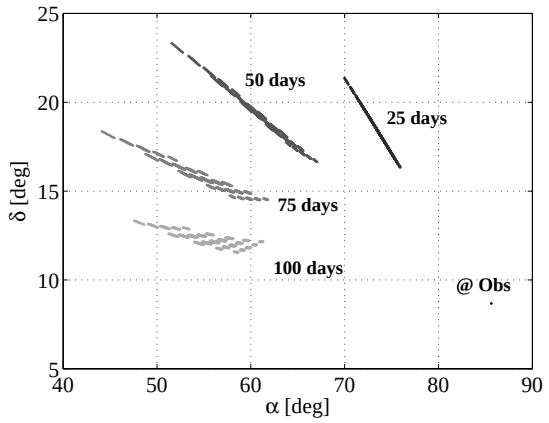


Figure 16: Propagated observable sets. Sixth order expansion and propagation up to 100 days for $\sigma = 0.3$ arcsec.

early mapped forward in time up to an arbitrary epoch, the method can be potentially used to recover asteroids' for which scarce observations are available.

Simulated observations of two known asteroids have been used to assess the method performances. For the considered cases the number of iterations necessary for the method to converge is limited to two or three. Furthermore, for one case, convergence is reached even when Gauss' method does not. Throughout the manuscript more emphasis was given to the capability of mapping uncertainties rather than on the convergence properties. Future works will be devoted to a detailed analysis of the convergence properties by running a large test campaign. Furthermore, the use of the solution of Gauss' 8-th degree polynomial as first guess for the iteration procedure will be lifted to gain independency from the hypothesis of sufficiently close and not aligned observations (being based on Lambert's and Kepler's problem, the iterative procedure is independent from observation gaps and configuration). A study on suitable starting values for the slant ranges $\rho_2^{1^-}$ and $\rho_3^{1^-}$ would be then mandatory. The final goal is to deliver a software for application to real scenarios.

8 Acknowledgment

The authors are grateful to Mario Carpino for helpful discussions on asteroids orbit determination. Special thanks go to Martin Berz and Kyoko Makino for having introduced us to differential algebra techniques as well as for the implementation of COSY-Infinity, without which this work would not have been possible.

References

[1] Armellin, R., P. Di Lizia, Bernelli-Zazzera, F., and Berz, M.,"Asteroid close encounter characterization using differential algebra: the case of Apophis", *Celestial Mechanics and Dynamical Astronomy*, Vol. 107, 2010, pp.451–470.

[2] Armellin R., Di Lizia P., Makino K., Berz M., "Rigorous Global Optimization of Impulsive Planet-to-Planet Transfers in the Patched-Conics Approximation", *Engineering Optimization*, in press, doi:10.1080/0305215X.2011.570757.

[3] Berz, M., "The method of power series tracking for the mathematical description of beam dynamics", *Nuclear Instruments and Methods* A258, 1987, pp. 431–436.

[4] Berz, M. *Modern Map Methods in Particle Beam Physics*, Academic Press, New York, 1999b.

[5] Berz, M., and Makino, K., "COSY INFINITY version 9 reference manual", MSU Report MSUHEP-060803, Michigan State University, East Lansing, MI 48824, 2006, pp. 1–84.

[6] Bykov, O.P., L'vov, V.N. Izmailov, I.S. and Kastel, G.R., "An accuracy estimation of the World CCD asteroid observations in the years 1999-2005", *Planetary and Space Science*, Vol. 56, 2008, pp. 1847–1850.

[7] Celletti, A., Pinzari, G., "Dependence on the observational time intervals and domain of convergence of orbital determination methods", *Celest. Mech. Dyn. Astron.*, Vol. 95, 2006, pp. 327–344.

[8] Chesley, S.R., and Milani, A., "An automatic Earth-asteroid collision monitoring system", *Bull. Am. Astron. Soc.*, Vol. 32, 2000, pp. 862.

[9] Curtis, H.D., *Orbital Mechanics for Engineering Students*, Elsevier, Oxford, 2005.

[10] Escobal, P.R., *Methods of Orbit Determination*, John Wiley and Sons, Inc., New York, 1965.

[11] Gauss C. F., *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, reprinted by Dover publications, New York, 1963.

[12] Gooding, R.H.: "A new procedure for the solution of the classical problem of minimal orbit determination from three lines of sight", *Celestial Mechanics Dynamical Astronomy*, Vol. 66, 1997, pp. 387–423.

[13] Laplace, P.S., *Mémoires de l'Académie Royale des Sciences*, Paris, Reprinted in Laplace's Collected Works, 1780.

[14] Leuschner, A.O., "On the Laplacean orbit methods", *Proceedings of the ICM*, 1912, pp. 209[U+FFFD]217.

[15] Milani, A. and Gronchi, G.F., *Theory of orbit determination*, Cambridge University Press, Cambridge, 2010.

[16] Merton, G., "A modification of Gauss's method for the determination of orbits", *Monthly Notices of the Royal Astronomical Society*, Vol. 85, 1925, pp. 693–732.

[17] Milani, A., Chesley, S.R., and Valsecchi, G.B., "Asteroid close encounters with Earth: Risk assessment ", *Planet. Space Sci.*, Vol. 48, 2000, pp. 945–954.

[18] Milani, A., Chesley, S.R., Chodas, P.W. and Valsecchi, G.B., "Asteroid close approaches: analysis and potential impact detection", Asteroids III, Vol. 1, 2002, pp. 89–101.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Optimal Low-Thrust Station Keeping of Geostationary Satellites

F. Topputo and F. Bernelli-Zazzera
Dipartimento di Ingegneria Aerospaziale
Politecnico di Milano, Milano, Italy

(francesco.topputo,franco.bernelli)@polimi.it

Keywords: *station keeping, geostationary satellite, optimal control, low-thrust propulsion.*

Abstract

In this note we design optimal low-thrust station keeping maneuvers for geostationary satellites. As the orbital perturbations would move the satellite from its nominal position, station keeping maneuvers are needed to stabilize the satellite about a given longitude. The system of equations we use models the satellite's motion through spherical coordinates. The optimal control problem is solved with the approximating sequence of Riccati equations. Results show the feasibility of the proposed application as well as the advantages intrinsic in low-thrust systems.

1 Introduction

The most important mission requirement for a geostationary (GEO) satellite is position control, in order to have extremely small deviations with respect to the nominal position for the entire lifetime of the satellite. As the number of GEO satellites is constantly growing, with a great congestion in the most interesting position windows, the reduction in the acceptable deviations from nominal positions becomes necessary. It is then important to program an effective control strategy, called station keeping (SK), capable of counteracting the perturbations that act on a satellite.

In an ideal model the satellite would exactly remain on the GEO orbit and no corrections would be necessary to preserve its position. Nevertheless, the real motion is characterized by the presence of several perturbations which may have gravitational nature, as those given by the Earth oblateness and by the gravitational forces of the Moon and Sun, or non-gravitational nature, as the solar radiation pressure [1,2]. Compared to the gravitational attraction of the Earth, these perturbations have a small magnitude but they cannot be neglected as their effect is considerable in the lifetime of a satellite, which is typically on the order of years. This asks for the introduction of SK maneuvers.

The position of GEO satellites is usually controlled with chemical, high-thrust engines. Chemical thrusters are typically fired once every two weeks. Given the small ratio between the maneuver time and the GEO orbital period, chemical thrusts is considered with good approximation as impulsive. More recently, the use of low-thrust systems is being considered as a powerful alternative to the chemical propulsion [3]. The advantage offered by low-thrust systems is twofold: 1. they guarantee a good pointing of the thrust vector, so achieving

an accurate execution of the SK maneuver; 2. they have the specific impulse which is approximately ten times that of chemical thrusters, so allowing a considerable saving of propellant mass [4]. The low-thrust SK represents a controlled motion (their effect is no longer modeled as producing impulses), and it can be approached with the optimal control theory. One recent theory applicable to the control of nonlinear systems is the approximating sequence of Riccati equations (ASRE), which allows solving optimal control problems for nonlinear, time varying, systems. It is based on the iterative solution of a sequence of linear quadratic control problems, which converges to the exact solution of the problem.

In this work we use the ASRE to solve the nonlinear optimal control that models the low-thrust SK of GEO satellites.

2 Dynamical Model

Among the possible formulations of the satellite dynamics, the model in spherical coordinates has been adopted since it allows to; 1) immediately adapt to the ASRE control scheme; 2) make the angular coordinates coincide with position error, for easy understanding of the results; 3) simplified models of perturbations are available in this reference system.

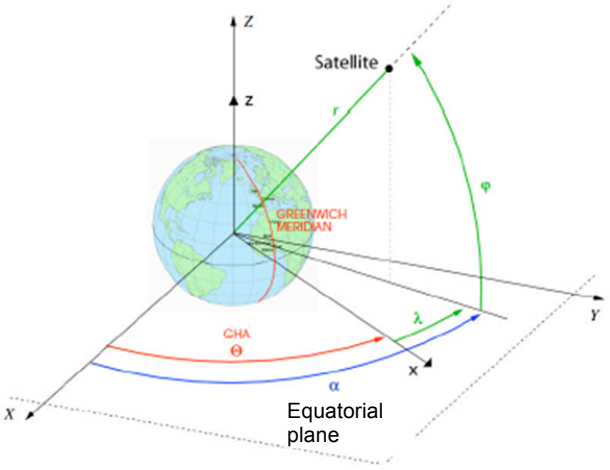


Fig. 1 Spherical coordinates adopted

With reference to Figure 1, the system of equations used to model the satellite's motion is

$$\begin{aligned}
\ddot{r} &= -\frac{\mu}{r^2} + r\dot{\varphi}^2 + r\left(\dot{\lambda}+\omega\right)\cos^2\varphi + a_{p_r} \\
\ddot{\lambda} &= 2\dot{\varphi}\left(\dot{\lambda}+\omega\right)\tan\varphi - 2\frac{\dot{r}}{r}\left(\dot{\lambda}+\omega\right) + \frac{a_{p_\lambda}}{r\cos\varphi} \\
\ddot{\varphi} &= -2\frac{\dot{r}}{r}\dot{\varphi} - \left(\dot{\lambda}+\omega\right)^2 \sin\varphi\cos\varphi + \frac{a_{p_\varphi}}{r}
\end{aligned} \qquad (1)$$

where (r, λ, φ) is a set of spherical coordinates (r is the radius, λ is the longitude deviation from the nominal position, φ is the latitude), μ and ω are the Earth's gravitational parameter and angular velocity, respectively. Eqs. (1) embed the orbital perturbations, a_p, which are decomposed into their radial, tangential, and normal terms (a_{pr}, $a_{p\lambda}$, $a_{p\varphi}$). These terms model the following perturbations: 1. the Earth's oblateness; 2. the luni-solar gravitational attraction; 3. the solar radiation pressure. The orbital perturbations alter the two-body dynamics, and deviate the satellite from its nominal position, so producing an uncontrolled motion.

All perturbations are modeled and evaluated through a potential function method. This is common practice to model the non-uniform gravity field of the central body, and it is here adopted also for the other perturbations [5,6] for ease of use within the nonlinear state-space model that will be adopted for the control design.

The gravity potential is modeled taking into consideration the first three terms in the series

$$U = \frac{\mu}{r}\sum_{l=2}^{\infty}\sum_{m=0}^{l}\left(\frac{R_T}{r}\right)^l P_{l,m} \cdot \cdot\left(C_{l,m}\cos m\lambda + S_{l,m}\sin m\lambda\right) \qquad (2)$$

where $P_{l,m}$ are Legendre polynomials, function of $\sin\varphi$, $C_{l,m}$ and $S_{l,m}$ are known coefficients.

Third body gravitational perturbations are evaluated with the second order potential

$$U_3 = \frac{\mu_3}{r_3}\left[1+\left(\frac{r}{r_3}\right)^2\frac{3\cos^2\psi_3 - 1}{2}\right] \qquad (3)$$

where the variable r_3 indicates the distance between the Earth and the third body and ψ_3

indicates the angle between the satellite position and the third body position (see Fig.2). In the present work the Sun and the Moon have been considered as third bodies.

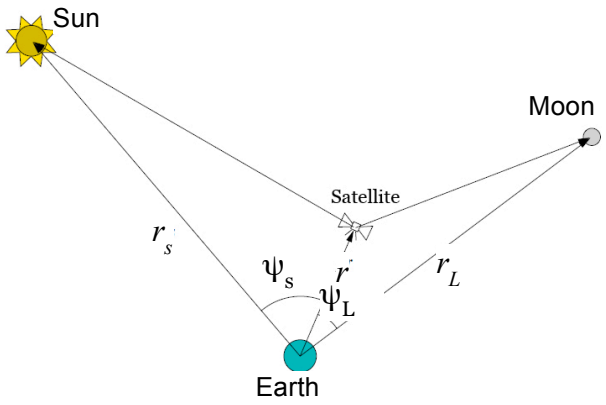


Fig. 2 Third body parameters

Solar radiation perturbation has been modeled through the pseudo-potential function

$$\begin{aligned} U_{sp} &= -P_{sr}(1+\beta)\frac{A}{m} r\cos\psi_S \\ a_{sp} &= P_{sr}(1+\beta)\frac{A}{m}\hat{u}_{Ssat} \\ P_{sr} &= \frac{C_S}{c} \end{aligned} \qquad (4)$$

where, with reference to Fig.3, A/m is the area to mass ratio of the satellite, β takes into account the surface properties of the satellite (β=1 indicates total reflection, β=0 total absorption), ψ_S indicates the angle between the satellite position and the Sun position, $\hat{u}_{Ssat}$ indicates the Sun-satellite direction, C_S is the solar flux intensity and c is the speed of light.

The perturbation accelerations acting on the spacecraft are then evaluated as the gradient of the potential functions above mentioned. The expressions of the accelerations due to each perturbation are collected in the Appendix.

When used for numerical analysis, the dynamic model is formulated normalizing some units for better numerical consistency. The unit of length is taken equal to the GEO orbit radius, the gravitational parameter of the Earth μ is taken as equal to 1, and accordingly the unit of time is such that the orbit period is equal to 2π units of time, which means that the unit of time is approximately 3.82 hours.

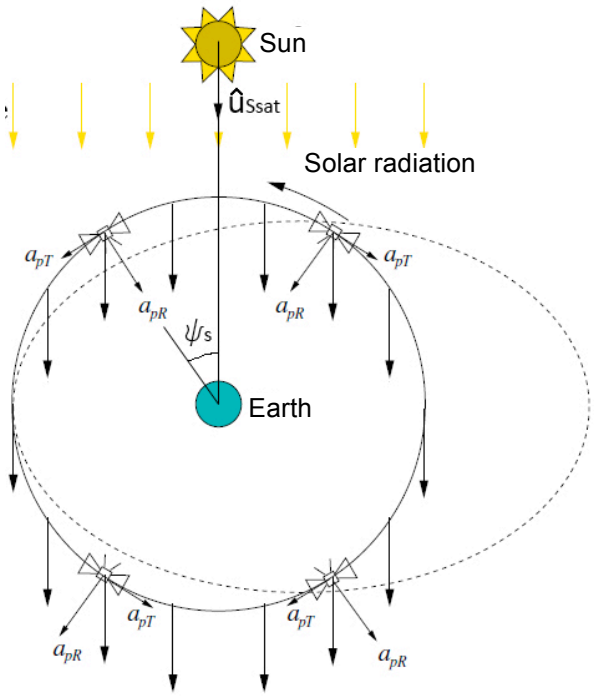


Fig. 3 Solar radiation perturbation

3 Optimal Control

The controlled dynamics is defined by adding the terms u_r, $u_\lambda/(r\cos\varphi)$, and u_φ/r to the three Eqs. (1), where $\mathbf{u} = (u_r; u_\lambda; u_\varphi)$ is the low-thrust control vector. The controlled dynamics can be written in the form

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x},t)\mathbf{x}(t) + \mathbf{B}(\mathbf{x},t)\mathbf{u}(t) \qquad (5)$$

where $\mathbf{x}$ is the six-element state vector (positions and velocities). The nonlinear objective function is

$$J = \frac{1}{2}\mathbf{x}_f^T\mathbf{S}\mathbf{x}_f + \frac{1}{2}\int_{t_0}^{t_f}\left(\mathbf{x}^T\mathbf{Q}(\mathbf{x})\mathbf{x} + \mathbf{u}^T\mathbf{R}(\mathbf{x})\mathbf{u}\right)dt \qquad (6)$$

where $\mathbf{S}$, $\mathbf{Q}$ are two positive-semidefinite 6×6 matrices; $\mathbf{R}$ is a 3×3 positive-definite matrix.

3.1 The ASRE Method

The nonlinear, control-affine system (5), together with the performance index (6), define an optimal control problem (the initial state, $\mathbf{x}_0$, is assumed given). This problem is solved with a method called approximating sequence of Riccati equations (ASRE). This method is a series of time-varying linear quadratic regulators whose solution is proven to converge to the optimal solution of the nonlinear problem [7,8]. Once the problem is factorized into the form (5), the successive approximation scheme is initiated by evaluating the state and control pseudo-matrices using the problem's initial condition and zero control. The associated time-varying linear quadratic regulator problem is solved and its solution is used to evaluate the state and control pseudo matrices in the successive step.

$$\dot{\mathbf{x}}^{[0]} = \mathbf{A}(\mathbf{x}_0,t)\mathbf{x}^{[0]}(t) + \mathbf{B}(\mathbf{x}_0,t)\mathbf{u}^{[0]}(t) \qquad (7)$$

$$J = \tfrac{1}{2}\mathbf{x}_f^{[0]T}\mathbf{S}\mathbf{x}_f^{[0]} + \\ + \tfrac{1}{2}\int_{t_0}^{t_f}\left(\mathbf{x}^{[0]T}\mathbf{Q}(\mathbf{x}_0)\mathbf{x}^{[0]} + \mathbf{u}^{[0]T}\mathbf{R}(\mathbf{x}_0)\mathbf{u}^{[0]}\right)dt \qquad (8)$$

At the [i]-th step, the following problem has to be solved

$$\dot{\mathbf{x}}^{[i]} = \mathbf{A}(\mathbf{x}^{[i-1]},t)\mathbf{x}^{[i]}(t) + \mathbf{B}(\mathbf{x}^{[i-1]},t)\mathbf{u}^{[i]}(t) \qquad (9)$$

$$J = \tfrac{1}{2}\mathbf{x}_f^{[i]T}\mathbf{S}\mathbf{x}_f^{[i]} + \\ + \tfrac{1}{2}\int_{t_0}^{t_f}\left(\mathbf{x}^{[i]T}\mathbf{Q}(\mathbf{x}^{[i-1]})\mathbf{x}^{[i]} + \mathbf{u}^{[i]T}\mathbf{R}(\mathbf{x}^{[i-1]})\mathbf{u}^{[i]}\right)dt \qquad (10)$$

whose solution produces $x^{[i]}$, $u^{[i]}$ ($x^{[i-1]}$ is the solution of the previous step). Iterations go on until a certain convergence criterion is satisfied; i.e., $|x^{[i]}-x^{[i-1]}|\leq\varepsilon$, where ε is a prescribed tolerance.

At each step, the solution is found by solving a two-point boundary value problem, derived by assembling the Euler-Lagrange equation associated to the optimal control problem

$$\begin{Bmatrix}\dot{\mathbf{x}}^{[i]}\\ \dot{\boldsymbol{\lambda}}^{[i]}\end{Bmatrix} = \mathbf{H}(\mathbf{x}^{[i-1]},t)\begin{Bmatrix}\mathbf{x}^{[i]}\\ \boldsymbol{\lambda}^{[i]}\end{Bmatrix} \qquad (11)$$

$$\mathbf{H}(\mathbf{x}^{[i]},t) = \\ = \begin{bmatrix}\mathbf{A}(\mathbf{x}^{[i]},t) & -\mathbf{B}(\mathbf{x}^{[i]},t)\mathbf{R}^{-1}(\mathbf{x}^{[i]})\mathbf{B}^T(\mathbf{x}^{[i]},t)\\ -\mathbf{Q}(\mathbf{x}^{[i]}) & -\mathbf{A}^T(\mathbf{x}^{[i]},t)\end{bmatrix} \qquad (12)$$

The general solution can be found by evaluating the state transition matrix associated to problem (11)-(12), and solving for the adjoint variable $\boldsymbol{\lambda}$. This is obtained with the following algorithm:

- Partition the augmented state vector in order to separate the states that are fixed at final time, along with their associated adjoint variables

$$\mathbf{y}(t) = \begin{cases}\mathbf{x}_1(t)\\ \hat{\mathbf{x}}(t)\\ \boldsymbol{\lambda}_1(t)\\ \hat{\boldsymbol{\lambda}}(t)\end{cases} \quad\rightarrow\quad \mathbf{x}_1(t_f) = \mathbf{x}_{1f} \qquad (13)$$

- write the general solution using the state transition matrix

$$\dot{\mathbf{y}} = \mathbf{H}\mathbf{y} \rightarrow \mathbf{y}(t) = \boldsymbol{\Phi}(t,\tau)\mathbf{y}(\tau) \\ \dot{\boldsymbol{\Phi}} = \mathbf{H}\boldsymbol{\Phi} \quad ; \quad \boldsymbol{\Phi}(0^+) = \mathbf{I} \qquad (14)$$

$$\boldsymbol{\Phi} = \begin{bmatrix}\boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12} & \boldsymbol{\Phi}_{13} & \boldsymbol{\Phi}_{14}\\ \boldsymbol{\Phi}_{21} & \boldsymbol{\Phi}_{22} & \boldsymbol{\Phi}_{23} & \boldsymbol{\Phi}_{24}\\ \boldsymbol{\Phi}_{31} & \boldsymbol{\Phi}_{32} & \boldsymbol{\Phi}_{33} & \boldsymbol{\Phi}_{34}\\ \boldsymbol{\Phi}_{41} & \boldsymbol{\Phi}_{42} & \boldsymbol{\Phi}_{43} & \boldsymbol{\Phi}_{44}\end{bmatrix}$$

- write the solution at $t=t_f$ and evaluate $\boldsymbol{\lambda}_1(t_0)$

$$\boldsymbol{\lambda}_1(t_0) = \left[\boldsymbol{\Phi}_{13}(t_f,t_0)\right]^{-1}\cdot \\ \cdot\left[\mathbf{x}_1(t_f) - \boldsymbol{\Phi}_{11}(t_f,t_0)\mathbf{x}_1(t_0) - \right. \\ \left. - \boldsymbol{\Phi}_{12}(t_f,t_0)\hat{\mathbf{x}}(t_0) - \boldsymbol{\Phi}_{14}(t_f,t_0)\hat{\boldsymbol{\lambda}}(t_0)\right] \qquad (15)$$

- write the linear combination $\hat{\mathbf{S}}\hat{\mathbf{x}}(t_f) - \hat{\boldsymbol{\lambda}}(t_f)$, where $\hat{\mathbf{S}}$ is the partition of $\mathbf{S}$ associated to the unconstrained final state

$$\begin{aligned}&\left[\hat{\mathbf{S}}\boldsymbol{\Phi}_{21}(t_f,t_0)-\boldsymbol{\Phi}_{41}(t_f,t_0)\right]\mathbf{x}_1(t_0)+\\&+\left[\hat{\mathbf{S}}\boldsymbol{\Phi}_{22}(t_f,t_0)-\boldsymbol{\Phi}_{42}(t_f,t_0)\right]\hat{\mathbf{x}}(t_0)+\\&+\left[\hat{\mathbf{S}}\boldsymbol{\Phi}_{23}(t_f,t_0)-\boldsymbol{\Phi}_{43}(t_f,t_0)\right]\boldsymbol{\lambda}_1(t_0)+\\&+\left[\hat{\mathbf{S}}\boldsymbol{\Phi}_{24}(t_f,t_0)-\boldsymbol{\Phi}_{44}(t_f,t_0)\right]\hat{\boldsymbol{\lambda}}(t_0)\end{aligned} \qquad (16)$$

- assuming that the initial state is known, evaluate $\hat{\boldsymbol{\lambda}}(t_0)$ from Eq.(16). Now the entire initial condition vector is known, so that the system can be integrated forward in time and the control law becomes

$$\mathbf{u}(t)=-\mathbf{R}^{-1}\mathbf{B}^T\boldsymbol{\lambda}(t) \qquad (17)$$

More details on the procedure can be found in [9]. This method, which represents a variant of the standard ASRE method, allows us to consider a mix of boundary conditions on state and controls, therefore it allows us to solve a wide class of control problems.

4 Control Requirements

The position control requirement for a GEO satellite can be in short summarized as the need to maintain a position within a window centered in the nominal longitude and defined by the maximum errors in longitude and latitude (see Fig. 4)

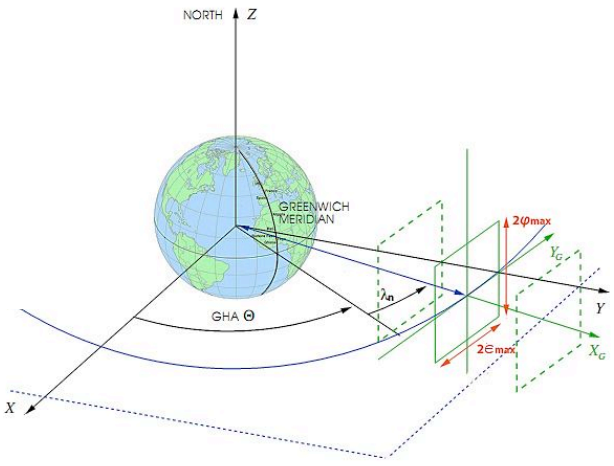


Fig. 4 SK control requirements

$$\begin{aligned}-\varepsilon_{max} \le \varepsilon \le \varepsilon_{max}\\ -\varphi_{max} \le \varphi \le \varphi_{max}\end{aligned} \qquad (18)$$

In a typical mission the maximum deviation is of 0.05 degrees.

The SK position control is characterized by three distinct operations:

- orbit determination, that estimates the actual position of the satellite on the basis of available measurements;
- maneuver planning, in which the actual control commands are evaluated;
- maneuver execution, in which the control is effectively applied to the satellite.

The three phases are normally all synchronized with an overall control period T_{SK}. The first two phases are within a free drift period T_{FD}, while the third phase has a period T_C (see Fig. 5). The free drift period is normally defined on the basis of the disturbances, so that the satellite remains within the prescribed control window.

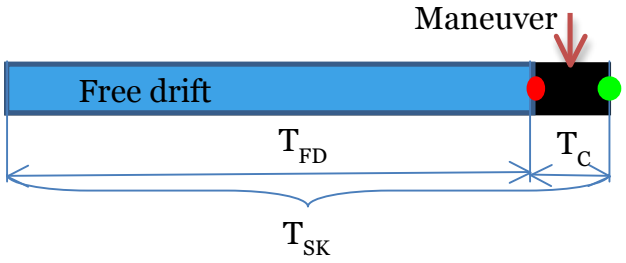


Fig. 5 SK maneuver planning

5 Application Example

The SK control algorithm has been applied to a sample test case, considering a satellite with overall mass m_0=3000 kg, surface exposed to Sun equal to 100 m^2, optical reflection coefficient equal to 0.5, positioned nominally at a longitude λ_n = 60° East. Since this position is unstable, the satellite would drift eastward, essentially because of the non-uniform Earth's gravity field. In order to limit the drift to the limit value ε_{max} = 0.05 degrees, a free drift period of 1.5 days is the maximum possible. In order to increase the free drift period, in the present example the nominal longitude λ_n has been shifted westward by 0.04 degrees, so

allowing to consider a free drift period T_{FD} of 2.5 days. The control period T_C has been fixed at 0.5 days. The final time constraint has been explicitly posed only to latitude and longitude drift, setting them to zero, while the remaining states of the system have no hard constraint at final time, their errors are the result of the minimization of the control cost function.

Figure 6 shows the trajectory during 1 control cycle. The performances are satisfactory in terms of position control, well within the assumed bounds.

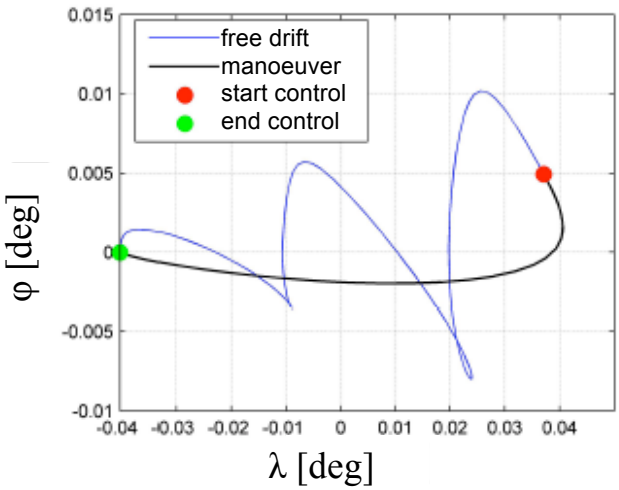


Fig. 6 Trajectory during 1 control cycle

Figure 7 reports the same results for a mission time of 1 month, shoving the same average performances. During one month, the overall control requirements, measured as velocity increment, are ΔV=6.59 m/s.

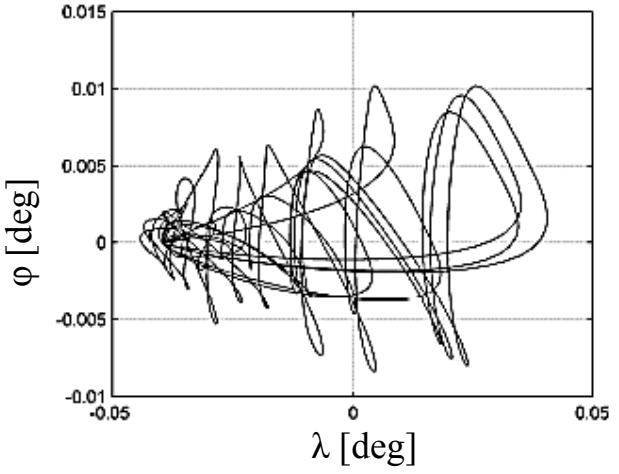


Fig. 7 Trajectory during 1 month

It is remarked that the control cycle is not the same at each cycle due to the change in position of the Moon and the Sun, with consequent change in the perturbations acting on the satellite.

Figure 8 reports the same results for a mission time of 1 year, shoving the same average performances and highlighting the differences between controlled and uncontrolled motion. The overall control requirements, for 1 year, are ΔV=79.69 m/s. Assuming the control is performed with a low thrust propulsion system with specific impulse I_{sp}=3000 s, the propellant mass required would be 8.11 kg/year, evaluated according to the rocket equation

$$m_p = m_0\left(1 - e^{-\frac{\Delta V}{I_{sp} g_0}}\right) \qquad (19)$$

This result is extremely interesting, since the position control is effective with low fuel consumption. A comparison has been made with two other methods available in the literature, for similar test cases, and the results are collected in Tab. 1. The ASRE method has been compared to a fixed horizon optimal control (FHOC) and to a receding horizon optimal control (RHOC). It is pointed out that ASRE is actually a fixed horizon control method, and the comparison of Tab. 1 shows the improvements achieved by ASRE over FHOC, almost reaching the performance of the RHOC.

Tab. 1 Comparison of results

Method	ΔV (m/s)	Yearly propellant consumption (kg)
ASRE	79.69	8.11
FHOC [10]	180.48	18.34
RHOC [10]	69.56	7.08

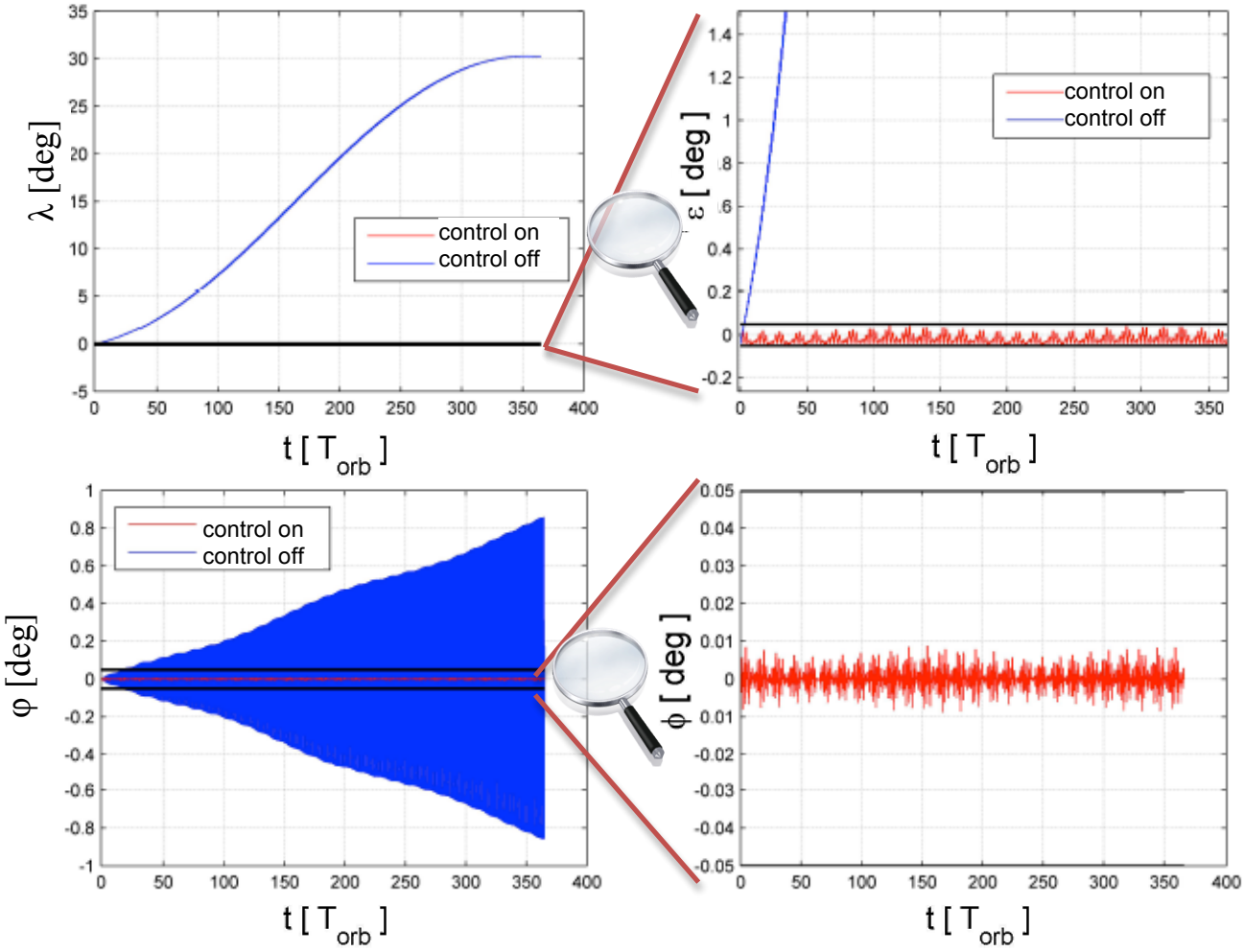


Fig. 8 Latitude and longitude errors during 1 year

6 Conclusion

This paper has shown how a modern ASRE control theory can be applied to the station keeping of geostationary satellites. One variant of the ASRE method has been elaborated in order to find a solution as a sequence of TPBVP, allowing to solve problems with a mix of constrained and unconstrained terminal states.

Numerical results and simulations on one realistic sample case confirm the effectiveness and accuracy of the proposed method.

The method and the results obtained allow us to define some future enhancements of the work. Special attention should be devoted to the optimization of the factorization and the selection of appropriate weight functions, for improved numerical performances of the algorithm. Furthermore, the control law should be in some way adapted to the actual thrust available, and finally it would be interesting to separate EW station keeping from NS station keeping, for improved performances.

References

[1] Vallado D. A., *Fundamental of Astrodynamics and Applications*, 2nd ed., Space Technology Library, 2004.

[2] Chong-Hung Zee, *Theory of Geostationary Satellites*, Kluwer Academic Publisher, 1989.

[3] Saccoccia G., Gonzalez del Amo J., Estublier D., "Electric Propulsion: A key Technology for Space Missions in the New Millennium", *ESA Bulletin*, Vol. 101, 2000.

[4] Martinez-Sanchez M., Pollard J., "Spacecraft Electric Propulsion – An Overview", *Journal of Propulsion and Power*, Vol. 14, No.5, 1998, pp. 688-699.

[5] Bertachini Prado A.F., "Third-Body Perturbation

in Orbits Around Natural Satellite", *Journal of Guidance, Control and Dynamics*, Vol. 26, No. 1, 2003.

[6] Campan G., Alby F., Gautier H., "Station Keeping Techniques for Geostationary Satellites", *Spaceflight Dynamics*, Vol.2, CNES 1995.

[7] Çimen T., Banks S.P., "Global Optimal Feedback Control for General Non Linear System with Non-Quadratic Performance Criteria", *System & control letters*, Vol.53, 2004, pp.327-346.

[8] Çimen T., Banks S.P., "Nonlinear Optimal Tracking Control with Application to Super-Tankers for Autopilot Design", *Automatica*, Vol.40, 2004, pp.1845-1863.

[9] Pasta M., "Station keeping di satelliti geostazionari mediante controllo ottimo non lineare" (in Italian) M.Sc. Dissertation, Politecnico di Milano, Italy, 2010.

[10] Losa D., "High vs Low Thrust Station Keeping Maneuver Planning for Geostationary Satellites", Ph.D. Thesis in Real Time Computer Science, Robotics, Control Engineering at the Ecole Nationale Superieure des Mines de Paris, France, 2007

Acknowledgements

The authors would like to acknowledge Marco Pasta whose MSc Thesis inspired part of the work described in this note. The authors are also grateful to Ashraf Owis of Cairo University for the support and the valuable suggestions given.

Appendix

The perturbation terms in the satellite dynamics are derived using a potential function model. The complete expressions of the perturbing accelerations are:

Radial acceleration due to Earth's oblateness

$$
\begin{aligned}
&a_{g_r}(r,\lambda,\varphi) = -3\frac{\mu R^2}{r^4}\left[\left(\tfrac{3}{2}\sin^2\varphi - \tfrac{1}{2}\right)C_{20} + \right.\\
&\left. + \left(3\cos^2\varphi\right)\left(C_{22}\cos 2\lambda + S_{22}\sin 2\lambda\right)\right] + \\
&- 4\frac{\mu R^3}{r^5}\left\{\left(\tfrac{5}{2}\sin^3\varphi - \tfrac{3}{2}\sin\varphi\right)C_{30} + \right.\\
&+ \left(\tfrac{15}{2}\sin^2\varphi\cos\varphi - \tfrac{3}{2}\cos\varphi\right)\left(C_{31}\cos\lambda + S_{31}\sin\lambda\right) + \\
&+ \left(15\sin\varphi\cos^2\varphi\right)\left(C_{32}\cos 2\lambda + S_{32}\sin 2\lambda\right) + \\
&+ 15\cos^3\varphi\left[C_{33}\cos\lambda\left(1 - 4\sin^2\lambda\right) + \right.\\
&\left.\left. + S_{33}\sin\lambda\left(4\cos^2\lambda - 1\right)\right]\right\}
\end{aligned}
$$

Longitude acceleration due to Earth's oblateness

$$
\begin{aligned}
&a_{g_\lambda}(r,\lambda,\varphi) = \\
&\frac{\mu R^2}{r^4\cos\varphi}\left(3\cos^2\varphi\right)\left(2S_{22}\cos 2\lambda - 2C_{22}\sin 2\lambda\right) + \\
&+ \frac{\mu R^3}{r^5\cos\varphi}\left\{\left(\tfrac{15}{2}\sin^2\varphi\cos\varphi - \tfrac{3}{2}\cos\varphi\right)\cdot\right.\\
&\cdot\left(S_{31}\cos\lambda - C_{31}\sin\lambda\right) + \\
&+ \left(15\sin\varphi\cos^2\varphi\right)\left(2S_{32}\cos 2\lambda - 2C_{32}\sin 2\lambda\right) + \\
&+ 15\cos^3\varphi\left[C_{33}\left(12\sin^3\lambda - 9\sin\lambda\right) + \right.\\
&\left.\left. + S_{33}\left(12\sin^3\lambda - 9\sin\lambda\right)\right]\right\}
\end{aligned}
$$

Latitude acceleration due to Earth's oblateness

$$
\begin{aligned}
&a_{g_\varphi}(r,\lambda,\varphi) = \frac{\mu R^2}{r^4}\left[3\sin\varphi\cos\varphi C_{20} + \right.\\
&\left. - 6\sin\varphi\cos\varphi\left(C_{22}\cos 2\lambda + S_{22}\sin 2\lambda\right)\right] + \\
&+ \frac{\mu R^3}{r^5}\left\{\left(\tfrac{15}{2}\sin^2\varphi\cos\varphi - \tfrac{3}{2}\cos\varphi\right)C_{30} + \right.\\
&+ \left[\tfrac{3}{2}\sin\varphi + \tfrac{15}{2}\left(2\sin\varphi\cos^2\varphi - \sin^3\varphi\right)\right]\cdot\\
&\cdot\left(C_{31}\cos\lambda + S_{31}\sin\lambda\right) + \\
&+ \left(15\cos^3\varphi - 30\sin^2\varphi\cos\varphi\right)\cdot\\
&\cdot\left(C_{32}\cos 2\lambda + S_{32}\sin 2\lambda\right) + \\
&- 45\sin\varphi\cos^2\varphi\left[C_{33}\cos\lambda\left(1 - 4\sin^2\lambda\right) + \right.\\
&\left.\left. + S_{33}\sin\lambda\left(4\cos^2\lambda - 1\right)\right]\right\}
\end{aligned}
$$

Radial acceleration due to third body

$$a_{3_r} = \frac{\mu_3}{r_3^3} r\left(3\cos^2\psi_3 - 1\right)$$

Longitude acceleration due to third body, calling δ the declination and α the right ascension

$$a_{3_\lambda} = -3\frac{\mu_3}{r_3^3} r\cos\psi_3 \cos\delta_3 \sin\left(\alpha - \alpha_3\right)$$

Latitude acceleration due to third body

$$a_{3_\varphi} = 3\frac{\mu_3}{r_3^3} r\cos\psi_3\left[\cos\varphi\sin\delta_3 - \right.$$
$$\left. - \sin\varphi\cos\delta_3\cos\left(\alpha - \alpha_3\right)\right]$$

Radial acceleration due to solar radiation

$$a_{sp_r} = -P_{sr}\left(1+\beta\right)\frac{A}{m}\cos\psi_S$$

Longitude acceleration due to solar radiation

$$a_{sp_\lambda} = P_{sr}\left(1+\beta\right)\frac{A}{m}\left[\cos\delta_S \sin\left(\alpha - \alpha_S\right)\right]$$

Latitude acceleration due to solar radiation

$$a_{sp_\varphi} = -P_{sr}\left(1+\beta\right)\frac{A}{m}\left[\sin\delta_S\cos\varphi - \right.$$
$$\left. - \cos\delta_S \sin\varphi\cos\left(\alpha - \alpha_S\right)\right]$$

CEAS 2011

Safety and Security

Ethical Aspects in Airport Security - A Passenger Survey

M. Cole and A. Kuhlmann
Bauhaus Luftfahrt e.V., Germany

Keywords: *Ethics, airport security, passenger survey, assessment of future security measures*

Abstract [1]

A broad range of security technologies and processes is installed at the airport to prevent threats to aviation from being successfully performed. New measures are often implemented as direct reactions to aviation security incidents, adding layer after layer to the security system.

The initial desire for more security inevitably entails a conflict of interests as more security often leads to less liberty, justice and privacy. Regarding this trade off, a comprehensive assessment of ethical aspects of airport security should always precede the development of new technologies and the implementation of new security standards. In the paper the development of a questionnaire for a passenger survey is described, which was conducted to investigate the connection of ethical aspects and airport security measures. Findings will be presented that allow evaluating current measures as well as anticipating reactions to possible future technologies, processes or standards.

1 Motivation

During the last decades civil aviation has been a preferred target for terrorists' attacks. Airports are the gateways through which most potential attackers have to pass through in order to reach an aircraft and consequently attract maximum attention. A broad range of security technologies and processes is installed to prevent this scenario from happening. At current passenger checkpoints and baggage control stations objects that could possibly cause trouble on board are focused on. The measures passengers face today have been subsequently introduced, adding layer after layer to the security system. Implementations of new technologies or processes often took place as direct reactions to aviation security incidents. One example is the 2006 transatlantic aircraft plot leading to an immediate liquid ban, followed by the introduction of detection devices for liquids, aerosols and gels (LAGs) at European airports [1].

[1] The results of this work are part of the research project SiVe, which is (partially) financed by the German Federal Ministry of Education and Research (BMBF). SiVe is part of the Security Research Program initiated in the context of the High-Tech Strategy by the German Federal Government.

The technological benefit derived from new measures regarding the identification of potentially dangerous objects usually is the main concern during the development and implementation phase. The fact that passengers – the customers of both airline and airport – have to pass through more and more invasive security checks is readily pushed into the background. But these technologies and processes are not only employed on luggage or cargo, but on humans, which has far reaching effects for their public acceptability and thus for air travel as such. The initial desire for more security inevitably entails a conflict of interests: more security often affects other basic rights, thus leading to e.g. less liberty, justice and privacy for the passenger.

So far the literature addressed general issues on ethics and security [2], ethical issues concerning specific security technologies [3] as well as the situation on specific airports [4]. This paper proposes an approach for the assessment of ethical issues in aviation- or airport-security in a systematic way and offers a way to link the upcoming ethical problems to specific airport security processes and technologies

2 Introduction

In the following passage the two basic terms in this context, ethics and security, are defined. The significance of both concepts for research in the field of airport security will be discussed.

First of all, ethics is based on the Greek word *ethos*, meaning ‘character’ or ‘custom’ [5]. The term relates to principles of human conduct that “provide the basic guidelines for determining how conflicts in human interests are settled and for optimizing mutual benefits of people living together in groups” [6]. Thus, by studying ethics one critically examines moral practices and human conduct resulting from a body of unwritten rules and habits underlying society.

The term security stems from the Latin *securitas*, which consists of *sine* (‘without’) and *cura* (‘worry’/ ‘sorrow’). The notion of security greatly depends on the context of its use [7]. In terms of airport security and security ethics it can be defined as “the integrity of the individual, particularly his or her body, and the integrity of the infrastructure critical to the functioning society” [3].

A basic level of security is an important prerequisite for cultural development and social life, as a constant feeling of insecurity limits individual freedom [3] and possibilities for development on both a personal and a societal level. Today security is no longer viewed as naturally given but as something that has to be actively created and maintained. In the course of globalization a shift has taken place in the security landscape: internal and external security cannot be seen separately anymore. A worldwide network of trade, commerce and travel leads to shared vulnerabilities but also to a global solidarity. A “merging of global threats and local fears” can be observed [8]. The feeling of insecurity is further encouraged by a development Baumann named ‘liquid life’ and defined as a life lived in “a society in which the conditions under which its members act change faster than it takes the ways of acting to consolidate into habits and routines” [9]. In times of high insecurities people accept severe restrictions of their fundamental rights to restore a certain level of subjective security (such as the “Patriot Act” in the aftermath of 9/11).

Security technologies and processes are introduced at the airport to protect people from unlawful behavior of others and to consequently reduce the subjective feeling of insecurity. What level of security seems desirable and appropriate is defined anew each time new dangers, threats and risks arise [3]. But, as mentioned above, measures meant to improve security may simultaneously violate other fundamental rights and values. Health, mobility, freedom of action, privacy, dignity as well as many other rights and values need to be carefully balanced as no fundamental right can be abandoned in favor of another [3].

In a performance driven environment, such as the field of airport security, the major concern is the effectiveness of implemented measures. Ethics are taken into account only in second instance or at an ad hoc basis [5]. Innovative measures will, however, never be accepted by a widespread public if ethically relevant issues are not solved beforehand [3].

Thus, a thorough assessment of relevant ethical aspects should always precede the development of new technologies and the implementation of new security standards for aviation security.

In this context, questions that should be addressed are: How does the passenger experience this measure? Will he perceive the innovation as affecting his sense of security in a positive way? Or will he rather feel his privacy or physical integrity intruded and consequently refrain from flying whenever possible? To tackle these questions related to acceptance of security measures, the passengers themselves have to be asked for their opinion.

3 The Passenger Survey Methodology

The following sections will give an overview on the content of a passenger survey conducted in 2010 on the acceptance of airport security measures and the subsequent structuring of the research topic. Parameters of the survey will be described to outline the conditions under which the survey took place.

3.1 Survey Parameters

The aim of the survey was to assess passengers' attitude towards security technologies and processes at the airport today and also on the ones passengers might face in the future. The representative sample of n=2000 was drawn from a travel panel of the German market research institute Gesellschaft für Konsumforschung (GfK). All respondents had at least flown once during the last year, thereby making sure that the respondents were familiar with current airport checkpoint procedures. The survey method was split between an online survey and postal questionnaires. This variation was done to equally reach different target groups and thus ensure the representativity of the survey. The respondents answered questions concerning their attitudes towards ethically relevant characteristics of security measures that will be described in detail in the next chapter.

3.2 Survey Setup

Ethical attitudes manifest themselves mostly in rather subtle feelings such as discomfort. People feel awkward or even hurt if ethical standards are violated. The more subtle a feeling the more difficult it is to clearly articulate the personal perception. The next logical step would be to specify what caused this feeling to be able to link ethical concerns with specific technologies and processes. This kind of abstract connection is normally not established. It is, thus, challenging to investigate the interrelation of ethical hurdles and security measures in a predefined survey, where face-to-face interaction or a subjective description of feelings is not possible. In the preparation phase of the survey this was a central issue that had to be addressed and solved.

In order to assess possible future measures it is necessary to identify general properties of security measures, which affect the controlled person, and in a second step link them to general dimensions of acceptance. If specific technologies are directly inquired in a survey one does not gain insights into how the underlying ethically relevant features of the technology might be assessed. Existing technologies can only be improved in an ethical sense if there is enough transparency on the critical properties.

To trace what technological or processual aspect caused a feeling of discomfort the approach chosen for this study was to decompose security measures into ethically relevant dimension and to ask people about their feelings towards these dimensions and not towards actual measures. Security measures were decomposed in six top level dimensions, which were partially split further in sub-categories:

Movement. The regular path passengers normally pursue through the airport (including today's checks) can be diverted if one is e.g. asked to follow the security personnel into a special room for additional checks to be conducted. This diversion of the regular path is indicated by the dimension "movement" and is split into two sub-categories: "movement diverted" and "movement not diverted".

Who is controlled? The dimension "who is controlled?" covers if there are any differentiations made between the passengers, meaning if some passengers are asked to undergo a number of checks not normally conducted in the standard procedure. Passengers are not informed ahead of the process if they are selected or not. "Differentiation" and "no differentiation" are the two sub-categories in this dimension.

Screening. There are different technological possibilities to screen a person. These are reflected in the sub-categories of this dimension: "electromagnetic (active)", "acoustic" and "sniffing". The measure for this category is "physical integrity" meaning intactness of the body.

Who controls how? This dimension accounts for the fact that security checks can be conducted by different means. The 'who' in "who controls how?" can either be human beings, animals or technologies. The 'how' indicates that the passengers might be touched or not touched during the process. Thus, this dimension has six sub-categories combining the 'who' and the 'how': e.g. "human being, no touch", "human being, touch" and so on.

Noticeability. In security checks today the passenger notices when he is controlled. Nevertheless, it is imaginable that in the future not every process is perceived by the passenger. Thus the two sub-categories of the dimension are "noticeable" and "not noticeable".

Details of passenger visible? This dimension concerns a specific type of technology that performs imaging scans of the human body. The result is some kind of image of the human body that can be assessed by responsible security personnel. This dimension is split into two sub-categories: "details of body are visible for personnel" and "details of body are not visible for personnel". The second sub-category can e.g. be realized by indicating suspicious areas on an abstract figure that can then be further examined by a directed pat-down search.

The dimensions can be matched with existing security measures as each measure is composed from some of them. By combining the survey's ratings of the relevant dimensions, single measures can be assessed with regard to their influence on the passengers' perception. A walk-through metal detector, for example, combines the ethical dimensions "movement not diverted", "no physical contact", "conducted by technology", "noticeable" and "no details of body visible".

The advantage of this approach is twofold: First, the decomposition allows inquiring into subtle personal attitudes in a way that otherwise would have only been possible in face-to-face interviews. Second, the approach offers an opportunity to anticipate the public's opinion concerning possible future security measures, as they most likely will be composed of the same fundamental dimensions.

4 Survey Results

This chapter will illustrate results of the conducted survey. In the following section 4.1 the structuring of the results according to the aforementioned dimensions of security measures will be discussed. In section 4.2 it will be demonstrated on the basis of three examples how the data can be employed to assess different technologies.

4.1 Ethical Assessment of Security Dimensions

The survey results can be analyzed according to the dimensions of the security measures described in section 3.2. Some distinctive features of the data set – structured according to the security dimensions – are presented in the following paragraphs to provide an overview on the results, which are then summarized in Table 1.

Movement. A vast majority of male (81.2%) as well as female (80.4%) participants indicated that they would find it acceptable if their regular path was unexpectedly diverted. However, the data indicates that the attitude towards this security dimension changes with age: Only 69% of the youngest respondents (14-24 years) felt that a diversion was acceptable whereas 87.6% of the elderly people (60+ years) would readily accept a respective measure. A steady shift can be observed between these two extremes.

Who is controlled? A high number of respondents found it acceptable if asked to pass through additional checks not everybody has to undergo. No major distinction can be noted concerning age, sex or income of the participant. For example, 79% of all male respondents and 82.9% of all female participants felt that a differentiation at the security checks is acceptable.

Screening. Of the three different possible screening technologies only the electromagnetic screening is rated critical by the majority (51.3%). Sniffing as well as acoustic screening is accepted by more than two-thirds. As far as electromagnetic and acoustic screening are concerned, women are somewhat more skeptical than men (45.3% vs. 52.3% for electromagnetic screening and 66.1% vs. 73.6% for acoustic screening).

Who controls how? Answers to questions addressing this dimension suggest that checks conducted by technological means are preferred over controls by human beings which in turn are preferred over animal checks. Generally, processes where the passengers are not touched are more agreed on than checks where passengers are touched (56.7% vs. 46.7%).

Noticeability. Concerning the dimension noticeability no major distinction can be found between the different groups of respondents: neither age nor sex nor income can be seen to have a relevant influence on the attitude. 46.3% of the respondents said they would not like to undergo security checks that are unnoticeable.

Details of passenger visible? A vast majority of passengers would readily accept a screening method where no details of the body become visible for security personnel. No relevant distinction can be found between different groups, except for the variable income: 86.6% of the respondents with an income above 2500€ p.a. indicated that they would accept the measure, whereas only 76.7% of those interviewed with an income below 2499€ p.a. agreed on the measure. The responses change when we take a closer look at the level of acceptance for measures where details of the body are visible for security personnel. In this case only 26.6% of the respondents feel they can accept the measure.

Table 1 summarizes the survey results for the entire sample. The Table gives an overview on the attitude the participants expressed towards the different dimensions of security measures and is, thus, a precondition to evaluate the ethical criticality of today's as well as tomorrow's security measures. The higher the percentage in the cells, the higher the overall level of approval for these particular dimensions.

Tab. 1 Survey Results: Security and Ethics

Dimension of Security Measures \ Ethical Dimension		Acceptance	Physical Integrity	Feeling of Security
Movement	movement diverted	81%		10%
Who is controlled?	differentiation	81%		14%
Screening	electromagnetic (active)		49%	
	acoustic		70%	
	sniffing		82%	
Who controls how?	human being, touch	38%		58%
	human being, no touch	51%		72%
	animal, touch	37%		55%
	animal, no touch	46%		62%
	technology, touch	65%		67%
	technology, no touch	72%		82%
Noticeability	not noticeable	54%		51%
Details of passenger visible?	details of body are not visible for personnel	82%		69%
	details of body are visible for personnel	27%		38%

For example, 81% of those who responded felt that it would be acceptable to undergo additional checks (see row "who is

controlled?"). Checks that are not noticeable would be accepted by 54%.

The column on the right side of the Table headed "feeling of security" provides additional information to evaluate measures. The higher the percentage in this column, the more passengers feel their subjective feeling of security to be improved by the respective security dimension. This factor can of course not outweigh the level of acceptance but helps to get a more holistic impression.

4.2 Mapping of Security Technologies and Processes to General Security Dimensions

In order to evaluate security processes and technologies according to these results, it is necessary to define the relevant properties for each security measure in terms of the six defined security dimensions. In most cases these dimensions are applicable and one of the options in each of the categories is valid. It is then possible to extract the value for the respective level of acceptance, physical integrity, or feeling of security using Table 1. If none of the mentioned sub-categories are applicable (e.g. "movement - movement not diverted") the level of acceptance for the dimension is assumed to be 100%. With this approach a security measure can provided with a specific value for each of the six dimensions.

For the calculation of a comparable, quantitative ethical measure the following procedure was applied. The "ethics score" of a security measure is calculated as the product of the specific values for the six relevant security dimensions, using either the value for "acceptance" or "physical integrity". The third dimension of ethical relevance chosen in this paper, "feeling of security", is assessed in a second step.

In the following paragraphs the application of this approach for the ethical evaluation of security measures will be presented. Examples will be the walk-through metal detector, the security scanner and a new kind of technology, the bee-scanner.

The metal detector is a technology that detects metal objects carried by the passenger without touching the body. It is a standard technology in security checks at airports today. The sub-category "technology, no touch" of the dimension "who controls how?" can be easily identified. 72% of the survey's participants felt that such a dimension would be acceptable (see Table 2).

Tab. 2 Walk-Through Metal Detector

Dimension of Security Measures / Ethical Dimension		Acceptance	Physical Integrity	Level of Acceptance
Movement	movement not diverted	100%		100%
Who is controlled?	no differentiation	100%		100%
Screening	electromagnetic (active)			
	acoustic			
	sniffing			
	not applicable		100%	100%
Who controls how?	human being, touch			
	human being, no touch			
	animal, touch			
	animal, no touch			
	technology, touch			
	technology, no touch	72%		72%
Noticeability	noticeable	100%		100%
Details of passenger visible?	details of body are not visible for personnel			
	details of body are visible for personnel			
	no imaging technology applied	100%		100%
Ethics Score				**72%**

The normal path one pursues through the airport is not diverted by this check and everybody has to undergo the control. Consequently, the sub-categories for "movement" and "who is controlled?" as shown in Table 1 are not applicable. Each sub-category is thus rated with a 100% acceptance.

The screening technologies mentioned in the survey did not include a passive electromagnetic technology. And as walk-through metal detectors are extensively used today with no public reservation, a 100% acceptance is assumed. Noticeable checkpoints are today's standard and highly accepted (100%) and as no imaging technology is applied there are no issues to be taken into account concerning the last dimension in the Table (100%). Now the "ethics score" can be calculated by multiplying the specific values for this technology. The result is an ethics score of 72%, which indicates a rather high level of overall acceptance by the passengers.

The figure for "feeling of security" corresponding with this evaluation can be derived from Table 1. As the only critical sub-category for the walk-though metal detector is "technology, no touch" the value for "feeling of security" can be read off Table 1: 82% of the respondents felt that checks by means of a non-touching technology would improve their personal feeling of security. The two high scores (72% Ethics Score and 82% Feeling of Security) indicate a very positive attitude of the public towards these scanners.

Another example for the application of the described approach is discussed in the following paragraph. Here a new generation security scanner is analyzed (see Table 3). In some areas the walk-through metal detector and the security scanner can be allocated the same sub-categories. This holds true for "movement", "who is controlled?", "noticeability" and "who controls how?". However, the security scanner operates with active electromagnetic waves. 49% of those who responded felt this to be no harm for their physical integrity. Furthermore, only 27% agree on checks that employ imaging technology showing details of the body to the personnel. By calculating the product of these dimensions the ethics score is generated. In this case it is very low because the security scanner touches different rather critical dimensions of security measures simultaneously.

Tab. 3 Security Scanner

Dimension of Security Measures \ Ethical Dimension		Acceptance	Physical Integrity	Level of Acceptance
Movement	movement not diverted	100%		100%
Who is controlled?	no differentiation	100%		100%
Screening	electromagnetic (active)		49%	49%
	acoustic			
	sniffing			
	not applicable			
Who controls how?	human being, touch			
	human being, no touch			
	animal, touch			
	animal, no touch			
	technology, touch			
	technology, no touch	72%		72%
Noticeability	noticeable	100%		100%
Details of passenger visible?	details of body are not visible for personnel			
	details of body are visible for personnel	27%		27%
	no imaging technology applied			
Ethics Score				**10%**

The two corresponding levels for "feeling of security" can be derived from Table 1: 38% o the respondents would state that a security check showing details of the body to the

security personnel would make them feel more secure at the airport. And 82% rely on a non-touching technology to secure the airport. Taking these different scores into account one can easily assess the security scanner as being very poorly accepted and not improving the feeling of security substantially.

To demonstrate that this method can also be used to evaluate possible future technologies the third example will focus on a technology in the development phase: an unconventional sniffer that uses conditioned sniffer bees to detect e.g. improvised explosives (e.g. Inscentinel [http://www.inscentinel.com/]). The basic concept is very closely related to standard handheld sniffers, but inside the shell there is no piece of high tech but living bees with their reflexes conditioned to react to stimuli e.g. smells of explosives.

If one presumes that bee scanners are integrated in the standard passenger process at the airport they neither would divert the movement of passengers nor would the passengers be subject to differentiation (see Table 4). The scanner screens by sniffing which 82% of the respondents feel would have no influence on their physical integrity. This sniffing is done by animals that do not touch the passengers (46% of acceptance). As the checks would be noticeable and there would be no imaging technologies applied the screening would otherwise be uncritical from an ethical point of view. The overall ethics score for this security measure is 38%. 62% of those who responded felt their subjective feeling of security positively influenced by a non-touching animal check. The scanner would thus probably be moderately accepted and the feeling of security would be in principal positively influenced.

Tab. 4 Future Technology: Bee Scanner

Dimension of Security Measures \ Ethical Dimension		Acceptance	Physical Integrity	Level of Acceptance
Movement	movement not diverted	100%		100%
Who is controlled?	no differentiation	100%		100%
Screening	electromagnetic (active)			
	acoustic			
	sniffing		82%	82%
	not applicable			
Who controls how?	human being, touch			
	human being, no touch			
	animal, touch			
	animal, no touch	46%		46%
	technology, touch			
	technology, no touch			
Noticeability	noticeable	100%		100%
Details of passenger visible?	details of body are not visible for personnel			
	details of body are visible for personnel			
	no imaging technology applied	100%		100%
Ethics Score				**38%**

5 Discussion

An ideal security measure would be composed only of ethically uncritical dimensions that, at the same time, also improve the passengers feeling of security. Two dimensions investigated in the survey fulfill these criteria: checks that are conducted by a

non-touching technology (acceptance 72%, feeling of security 82%) and imaging technologies that show no details of the body (acceptance 82%, feeling of security 69%). Security checks that are conducted by technologies but touch the body during the checks still receive a fairly high acceptance rate (65%) and positively influence the feeling of security (67%). The diversion of movement during checks and a differentiation between the passengers both received a high level of acceptance from the respondents (81%) and can thus be rated as uncritical from an ethical point of view. Concerning their influence on the feeling of security they are very poorly rated: only 10% and 14% of the respondents felt an improvement of security by these dimensions.

Taking these results into account, recommendations for new security measures and processes can be derived. Generally, measures that control by technological means and that avoid body contact during the checks are judged positively by the passengers. Imaging technologies that do not show any body details are also well accepted. Furthermore, if the industry invented measures that would divert the passenger's way or if additional processes would be randomly imposed, they would be widely accepted. From today's point of view the latter technologies do not improve the feeling of security, but if new procedures containing these properties arise and experts propose a high efficiency in terms of risk reduction, they will be applicable.

In the introduction of this paper it has been mentioned that security is something that has to be actively created and that we accept drawbacks concerning other basic rights to reach a certain level of security. The more a specific security measures improves the subjective feeling of security, the more people generally are willing to accept violations of other basic rights.

6 Summary and Conclusions

The active creation of security at an airport through different processes and technologies often entails a violation of other basic rights such as freedom or justice. The effects different security measures might have on these rights can lead to a very low level of acceptance. The passenger's attitude has to be taken into account when planning and implementing new measures. To assess relevant ethical aspects a passenger survey was conducted. The development and the results are described in this paper.

Recommendations derived from the survey can be summarized as follows: From an ethical standpoint and on the basis of the described survey data, the ideal security measure is

- a technology,
- which scans the passengers without any body contact (preferably with a sniffing technology) for dangerous devices
- without displaying any details of the body to the involved security staff.

All these aspects are not only well accepted, they also improve the passengers subjective feeling of security. Furthermore, the diversion of the standard path and a certain unpredictability concerning who is controlled are both aspects, which can be rated as not critical from an acceptance point of view.

References

[1] Feakin, T., "Insecure Skies? Challenges and Options for Change in Civil Aviation," Occasional Paper, Royal United Services Institute, London, United Kingdom, 2011.

[2] den Boer, M. and Kolthoff, E, (eds.), *Ethics and Security*, 1st ed., Eleven International Publishing, The Hague, 2010.

[3] Ammicht Quinn, R. and Rampp, B., "The Ethical Dimension of Terahertz and Millimeter-Wave Imaging Technologies – Security, Privacy and Acceptability," SPIE Proceedings: Defense, Security, and Sensing, Optics and Photonics in Global Homeland Security V, Orlando, 2009.

[4] Lemke, P. "Versicherheitlichung - Zur Akzeptanz von Sicherheitsmaßnahmen in Theorie und Praxis," 1. SIRA Conference Series, 2011.

[5] Kolthoff, E. "The Importance of Integrity in the Security Profession: Bringing in Human Rights", den Boer, M. and Kolthoff, E, (eds.), *Ethics and Security*, 1st ed., Eleven International Publishing, The Hague, 2010, pp. 39-55.

[6] Wittmer, D.P., "Ethical Decision-Making", Cooper, T.L. (ed.), *Handbook of Administrative Ethics*, 2nd ed., Marcel Dekker, New York, 2000,

pp. 481-508.
[7] den Boer, M. “Revolving Doors: Ethics in a Shifting Security Paradigm,” den Boer, M. and Kolthoff, E, (eds.), *Ethics and Security*, 1st ed., Eleven International Publishing, The Hague, 2010, pp. 15-37.
[8] Aas, K.F., *Globalization & Crime*, 1st ed., Sage, London, 2007.
[9] Baumann, Z., *Liquid Life*, 1st ed., Polity Press, Cambridge and Malden, 2005.

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Aircraft and Rotorcraft Pilot Coupling: a survey of recent research activities within the European project ARISTOTEL

Marilena D. Pavel
Delft University of Technology, The Netherlands

Jacek Malecki
PZL-Swidnik, Poland

Binh DangVu
ONERA, France

Pierangelo Masarati and Giuseppe Quaranta
Politecnico di Milano, Italy

Massimo Gennaretti
Università Roma Tre, Italy

Michael Jump
University of Liverpool, England

Hafid Smaili
NLR, The Netherlands

Achim Ionita
STRAERO, Romania

Larisa Zaicek
TsAGI, Russian Federation

Keywords: Aircraft and Rotorcraft Pilot Couplings

Abstract

Fixed and rotary wing pilots alike are familiar with potential instabilities or with annoying limit cycle oscillations that arise from the effort of controlling aircraft with high response actuation systems. Understanding, predicting and suppressing these inadvertent and sustained aircraft oscillations, respectively known as Aircraft- and Rotorcraft-Pilot Couplings (A/RPCs), is a challenging problem for the designers. Recent experiences show that especially modern designs are being confronted with an increasing degree of dangerous A/RPCs. The reason is that modern aircraft feature a significant level of automation in their Flight Control System (FCS). The FCS is generally intended to relieve pilot workload and allow operations in degraded weather and visibility conditions. Especially in modern rotorcraft, there seem to be embedded tendencies predisposing the FCS system towards triggering and sustaining dangerous RPCs. As the level of automation is likely to increase in future designs, extending to smaller aircraft and to different kinds of operation, the consequences of the pilot 'fighting' the FCS system and inducing A/RPCs needs to be eradicated. In Europe, the ARISTOTEL project (2010–2013) has been launched with the aim of understanding and predicting modern aircraft's susceptibility to A/RPC. The present paper gives an overview of the current status in RPCs and what can be expected in future designs.

Nomenclature

AAIB	Air Accidents Investigation Branch
ACAH	Attitude Command/Attitude Hold
ACARE	Advisory Council for Aeronautic Research in Europe
ADOCS	Advanced Digital/Optical Control System
AFCS	Automatic Flight Control System
AGARD	Advisory Group for Aerospace Research and Development
APC	Aircraft-Pilot Coupling
ARTI	Advanced Rotorcraft Technology Integration
ATTHeS	Advanced Technologies Testing Helicopter System
DFBW	Digital Fly-by-Wire
EASA	European Aviation Safety Agency
EMD	Engineering Manufacturing Development
EMS	Emergency Medical Services
FCS	Flight Control System
FDR	Flight Data Recorder
FSD	Full-Scale Development
HHC	Higher Harmonic Control
IBC	Individual Blade Control
J-AIB	Japan Accident Investigation Board
NATO	North Atlantic Treaty Organization
NoE	Nap of the Earth
NTSB	National Transportation Safety Board
OLOP	Open Loop Onset Point
PAO	Pilot Assisted Oscillations
PIO	Pilot Induced Oscillations
PVS	Pilot-Vehicle System
PLE	Position Limiting Element
QDR	Quick Data Recorder
RLE	Rate Limiting Element
RPC	Rotorcraft-Pilot Coupling
SAS	Stability Augmentation System
SCAS	Stability and Control Augmentation System
TCL	Thrust Control Lever
VNE	Velocity Never Exceed

1 Introduction

Today's high performance aircraft and rotorcraft are a product of the ever increasing demands of operator requirements. They are faster and more capable, but are consequently more complex than their predecessors. As their complexity increases, it appears that both engineers and pilots must be prepared to deal with an associated increased incidence of unfavourable so-called "Aircraft- and Rotorcraft-Pilot Couplings" (A/RPC). Generally speaking, A/RPCs are oscillations or divergent vehicle responses originating from adverse pilot-vehicle couplings. These undesirable couplings can range in severity from benign to catastrophic; benign A/RPCs affect the operational effectiveness of a mission, degrading the aircraft handling qualities; catastrophic A/RPCs result in the loss of the aircraft and lives. Until 1995, A/RPCs were usually known under the name of Pilot Induced/Pilot Assisted Oscillations (PIO/PAO), or Pilot in-the loop/Pilot-out-of-the-loop oscillations. The reason for this was that, in the past, the key causal factor in A/RPCs appeared to be the pilot, and the term 'Pilot Induced Oscillations' (PIOs) considered that he/she was mainly responsible for any such issue. Generally, for modern aircraft, it has become increasingly clear that the pilot is not at fault and that actually it is the rapid advance in the field of Flight Control Systems (FCS) that has increased the sensitivity of the pilot-vehicle system to the appearance of unfavourable A/RPC events. According to McRuer [1], "as a matter of fact, almost every aircraft equipped with a partial or total fly-by-wire FCS has, at one time or another of development process, experienced one or more A/RPC events." Moreover, "actuator rate limiting has been a key contributor to many of the PIOs experienced in flight test programs and must be considered in the development and evaluation of the flight control system," as reported by Tischler et al. [2].

In other words, in the FCS of any modern aircraft there seem to be some embedded tendencies that predispose the pilot-aircraft system towards A/RPC occurrence. Recently, the European Commission launched, under the umbrella of the 7th Framework Programme (FP7), the ARISTOTEL project (Aircraft and Rotorcraft Pilot Couplings — Tools and Techniques for Alleviation and Detection (`http://www.aristotelproject.eu/`), whose aim is to advance the state-of-the-art of A/RPC prediction and suppression. With a duration of 3 years, starting October 2010, and involving partners from across Europe — Delft University (TUD) as coordinator and NLR from The Netherlands, ONERA from France, Politecnico di Milano (POLIMI) and Università Roma Tre (UROMA3) from Italy, University of Liverpool (UoL) from the UK, STRAERO from Romania, PZL-Swidnik from Poland, TsAGI from Russia and EURICE from Germany, the project's objectives are to improve the physical understanding of present A/RPCs, foresee that of future ones, and define criteria to quantify an aircraft's susceptibility to A/RPC. The ARISTOTEL project is mainly based on the experience of previous European GARTEUR projects related to the area of A/RPCs (i.e. FM AG-12 "Pilot-in-the-Loop Oscillations — analysis and test techniques for their prevention, phase I", 1999–2001 [3], FM AG-15 "Pilot-in-the-Loop Oscillations — analysis and test techniques for their prevention, phase II", 2004–2007 [4] and HC AG-16 "Rigid body and aeroelastic rotorcraft-pilot coupling — prediction tools and means of prevention", 2005–2008 [5–8]) and is advised by main helicopter manufacturers in Europe such as Eurocopter and AgustaWestland. ARISTOTEL considers also the findings of the excellent work performed between 1995 and 1996 in the United States by the US NRC/ASEB Study Committee under the leadership of D. T. McRuer [1], one of the broadest and best documented open literature investigations especially on APC problems. Further, the workshops organised on APC problems by the NATO nations under the AGARD Flight Vehicle Integration Panel, Work Group 17 "Flying Qualities of Unstable Highly Augmented Aircraft" (1991) [9], Flight Vehicle Panel Workshop on "Active Control Technology" (1995) [10] and "Pilot Induced Oscillations" (1995) [11], NASA research conducted in 1995 [12] and 2001 [13] and the research of the US Air Force [14] were also analysed by ARIS-

TOTELians. The present paper is a synthesis of initial ARISTOTEL's investigations in the area of APCs and RPCs.

2 Defining and Understanding A/RPCs

The first dilemma that one needs to solve when analysing aircraft oscillatory behaviour is whether or not a particular event is an A/RPC. According to Mitchell and Klyde [15], ten different definitions seem to exist in the open literature and many times the aerospace community is unable to agree upon whether or not a particular event is an A/RPC. Also, it has been argued by some experts that renaming of PIO/PAOs as APCs in the case of aircraft or RPCs in that of rotorcraft is even more confusing.

"The introduction of the term 'Aircraft-Pilot Coupling' (APC) in the mid-1990s contributed to the obscuration of the obvious: while the intent of this new term was to capture both oscillatory and non-oscillatory adverse behaviors of the aircraft-pilot system, it has further factionalized the debate as there are now questions like, "Was this event a PIO or just APC?" and "What's the difference between PIO and APC?" to be addressed in the ongoing debates" [15].

Presently, PIO/PAO are considered subclasses of A/RPCs. In this context, the basic definitions of the old terminology PIO/PAO for an A/RPC event must be recalled. PIOs (Pilot Induced Oscillations) occur when the pilot inadvertently causes divergent oscillations by applying control inputs which are essentially in the wrong direction, or have a significant phase lag with respect to the aircraft/rotorcraft response. High gain tasks, long time delays introduced by the pilot while controlling the aircraft or changes in the pilot control behaviour introduced by FCS and control interfaces can all be the cause of a PIO. PAOs (Pilot Assisted Oscillations) are higher frequency phenomena related to involuntary control inputs given from the pilot, which may destabilize the aircraft. Figure 1 presents the closed loop Pilot-Vehicle System (PVS) for a modern rotorcraft. The input into the system is the Mission Task. This can be anything from a tracking task, manoeuvre or forcing on the stick. The pilot uses the task to control the integrated

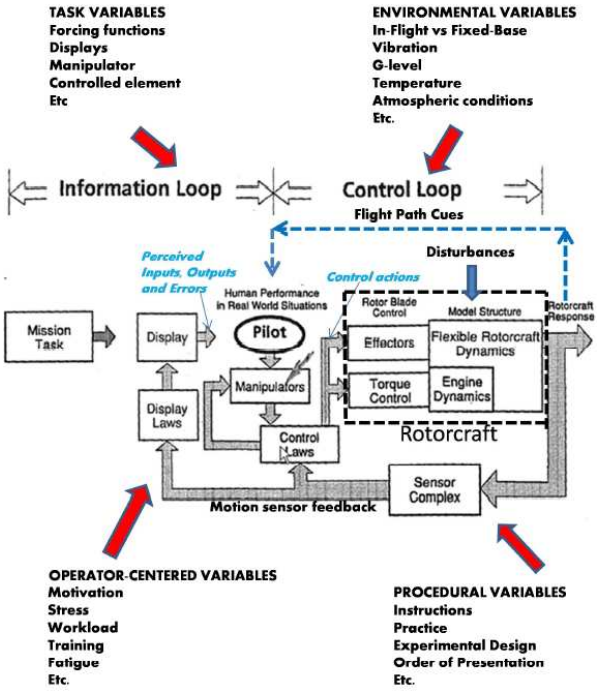


Figure 1: Pilot in the loop system for RPC analysis (adapted from [16]).

vehicle system which comprises 'inceptors' (manipulators), 'effectors' (actuators controlling the vehicle control surfaces, i.e. blade pitching system in the case of rotorcraft), 'sensors', 'displays', 'software interfaces' (control laws in the form of SAS/SCAS — Stability and Control Augmentation System, digital filters of a FCS etc., and display laws) and the inherent vehicle dynamics (the 'Rotorcraft' block in the figure). The output of the vehicle system is fed back to the pilot and the various automatic systems of the vehicle itself. The Pilot is of course the essential element in the PVS; he is the one that ultimately has to handle and evaluate all the complex vehicle systems.

In order to understand all possible factors that come into play during an A/RPC event, Figure 1 (adapted from [16] in the context of helicopter RPC) added four types of variables which, according to McRuer and Jex [17], affect pilot decisions.

- **Task variables** comprise all the system inputs and control system elements external to the pilot and which enter directly and explicitly into the pilot's control task. Four

of these — forcing function, display, manipulator, and controlled element dynamics — have a major effect on the pilot dynamics.

- **Environmental variables** are the variables external to the pilot such as ambient illumination, temperature, vibration, and G-loading (to the extent that this is superimposed on, rather than controlled by the pilot).
- **Operator-centred variables** include such things as training, fatigue, and motivation.
- **Procedural variables** (for a given experimental series) include such things as instructions, practice, order of presentation, etc., which can be very important to the accuracy and generality of experimentally based conclusions.

A closer look to Figure 1 reveals typical problems for rotorcraft that may induce RPCs:

1. In Key flight conditions, rotorcraft are inherently dynamically unstable, i.e. the vehicle does not stabilize itself and return to a steady flight condition after an upset. Moreover, rotorcraft are often required to execute demanding manoeuvres such as precision landings, hovering (with or without slung-loads), tracking tasks or autorotation. These are all high-gain tasks.

2. Rotorcraft are prone to vibrations transmitted from the rotors/prop-rotors to the airframe. The vibrations are generated by the inherent nature of the rotating-blade, that transmit periodic loads to the hub fixed frame, and also by the aerodynamic interference between the rotors and the fuselage. Once in the cabin, vibrations can afflict the pilot, resulting in wrong command inputs or biodynamic coupling (unaware inputs).

3. Rotorcraft are characterized by many control couplings resulting from the interactions between the dynamics of the rotating system, the rotor, and those of the fixed system, the airframe.

4. In conventional fixed-wing aircraft, control moments are transmitted directly from the control surfaces to the aircraft. In contrast, in rotorcraft, control inputs are transmitted through the swashplate to the blade in form of pitch changes, causing the rotor to flap and thence transmitting moments to the vehicle. It is well-known that cyclic inputs are applied at 1/rev-frequency through the swashplate mechanism. Thus, low-frequency pilot inputs generate high-frequency blade excitations. Clearly, rotor blade excitations, in the form of flap and lag motion, can be transformed back to the fixed airframe system, where eventually a new 1/rev-frequency shift may occur with positive or negative sign. In order to comprehend this airframe-rotor transformation mechanism of multiblade rotor systems, the concept of rotor modes is helpful. For example, collective rotor mode dynamics are transferred directly without frequency shift and cyclic rotor mode dynamics (so-called progressive and regressive modes) are transformed with a ± 1/rev frequency shift. Based on flight experience with modern helicopters, it appears that the RPCs of special interest are associated mainly with the multiblade rotor frequencies. For example, excitation of the low-damped main rotor regressive-inplane mode by cyclic inputs results in aircraft roll and pitch vibrations, or in the excitation of the low frequency pendulum mode of external slung loads by delayed collective and/or cyclic control inputs due to couplings of the load dynamics via elastic cables.

5. In rotorcraft, there exists a high inherent phase lag between inceptor input and vehicle body response due to the time required for actuator and rotor responses. Table 1 from [18] presents the typical equivalent time delays that are the result of implementing a digital FCS in a helicopter (see also Figure 4). One can see that the rotorcraft accounts for total effective time delays of more than 200 ms resulting from 50–70 ms inherent rotor response delays (66

ms in the Table 1), as rotor flapping responds to pitch input with about 90 deg phase delay, which corresponds to a quarter of the rotor period and time constant $\tau = \pi/(2\Omega)$, some 30 ms actuator delay and additional delays due to digital computing, sensor signal shaping and filtering. This rotor inherent delay is not present in control loops in fixed wing aircraft. The delay typically amounts to about 100 ms with conventional flight controls (actuators included). With FBW and filtering, the total rotorcraft delay can amount to 250 ms [19]. The total effective time delay of the rotorcraft-pilot system is directly related to the higher-order dynamics of actively controlled rotorcraft and will result in reduced system bandwidth and increased system phase delay. System bandwidth and effective time delay are two of the most important flight control design and specification parameters of the US Army's Handling Qualities Requirements for Military Rotorcraft, ADS-33E [20]. Figure 2 from [6] illustrates what happens to the phase lag of the helicopter dynamic response when the time delay is increased. The figure presents the bode plot for the pitch response to a swashplate (control) deflection input (note that time delays do not influence the magnitude plot). Looking at this figure, two observations can be made: a) the slope becomes steeper; this so-called phase roll-off or rate at 180 deg crossover frequency increases the equivalent time delay, and b) the phase bandwidth (crossover frequency at 135 deg) decreases. The combined effect of these two trends is that, due to the larger decrease in rate of the phase lag at a lower frequency, the phase margin decreases quicker for increasing input frequencies. In other words, the system destabilizes earlier. Concluding, total effective time delays of more than 200 ms may reveal poor handling qualities due to high gain tasks. Such time delays can be a strong cause of mental mismatch for the pilot with vehicles' dynamics.

Following Figure 1, the first goal of the ARIS-

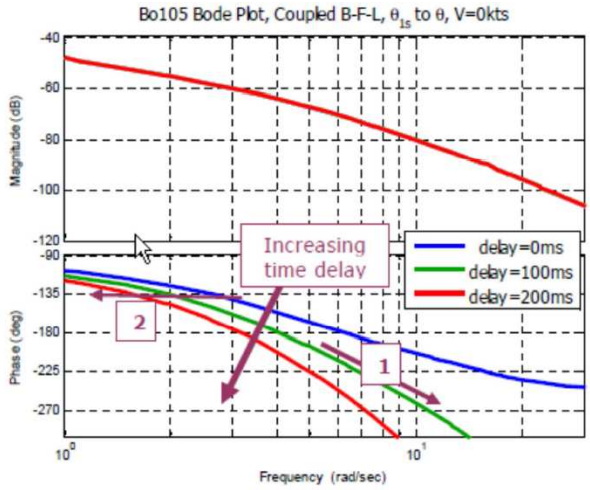


Figure 2: Effects of increasing the time delay on phase lag for a BO105 helicopter [6].

TOTEL project was to give a unified definition on A/RPCs to be used throughout the project.

"This is the best way to be sure that we are all talking about the same phenomenon, even if there is wide variation in the details of its causes and characteristics" [15].

In ARISTOTEL, an exhaustive review of the history of the studies related to A/RPC was performed, together with an analysis of the involved literature. A database of accidents occurred in the last 60 years both regarding APC and RPC was then put together (see Appendices A and B). The following definition was then proposed to be used throughout the project:

Definition: *An Aircraft- or Rotorcraft-Pilot Coupling (A/RPC) is an unintentional (inadvertent) sustained or uncontrollable vehicle oscillation characterized by a mismatch between the pilot's mental model of the vehicle dynamics and the actual vehicle dynamics. The result is that the pilot's control input is out-of-phase with the response of the vehicle, possibly causing a diverging motion.*

This definition is the result of an exhaustive discussion between the project partners where every single word has been given the right weight in order not to judge too severely or too negligently the risk of actual A/RPC happening. Three keywords can be highlighted from this definition, which can be found in almost every A/RPC accident:

Table 1: Equivalent time delays for rotorcraft [18].

Element	Delay (ms)	% of total
Rotor	66	30
Actuators	31	14
Control laws	17	8
Computations	22	10
Notch filter	11	5
Stick dynamics and filtering	76	34
Total delay	223	

- **Oscillatory behaviour:** every A/RPC event is related to oscillatory behaviour perceived (and then induced) by the pilot. If not leading to crash and catastrophic consequences, these events are related to extreme discomfort and are reported as dangerous happenings.
- **Mental Mismatch:** it is always related to A/RPC in that the pilot shows a wrong mental model of the dynamics of the system he/she is leading/lagging and increasing levels of unawareness of the command input he/she is giving. One can see that both descriptions of PIO and PAO as given above can be related to mental mismatch, leading to either a wrong command input in the PIO or to a totally unaware command input in the case of PAO.
- **Out-of-phase behaviour:** every fully developed A/RPC reports an out-of-phase input-output response witnessing the loss of control on the vehicle.

As discussed by McRuer [1], there are three crucial ingredients in an A/RPC event:

1. a change in pilot behaviour in his/her closed-loop control of the aircraft,
2. a change in the dynamic state or configuration of the aircraft, and
3. an initiation mechanism commonly referred to as 'trigger'.

Figure 3 presents the necessary and sufficient conditions for A/RPC development. The general cause of an A/RPC is commonly accepted to

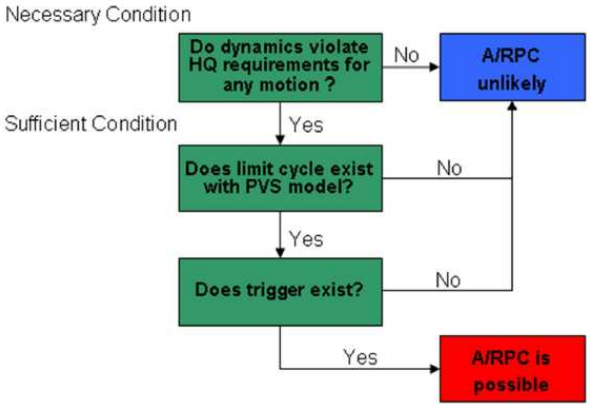


Figure 3: Three necessary conditions for A/RPCs [22].

be due to a trigger event. The trigger causes the pilot to quickly alter his/her control strategy. The trigger can occur in a number of different situations such as wind, gust (exogenous trigger), changes in FCS mode or in aircraft functioning, discontinuities in the pilot perception or in the behaviour of the vehicle, etc. (endogenous trigger) [1, 9, 21]. Trigger events may lead to A/RPC; however, not all trigger events will necessarily develop into A/RPC. Figure 3 from [22] shows that A/RPCs occur because aircraft dynamics allows them. For A/RPC to occur, aircraft must respond to pilot inputs in a manner that propagates them.

The triggers may develop under different conditions such as atmospheric turbulence, sudden change in the closed loop dynamics of the aircraft-pilot system, a nonlinear effect in flight control system, all these requiring a rapid change in pilot's control strategy. The trigger event has its effect on the pilot, but, as stated earlier, the aircraft must respond to the pilot's

input in a manner that propagates an A/RPC. Aircraft characteristics that are known to facilitate A/RPC behaviour include sluggish response modes, lightly damped modes, excessive phase lag or time delay, sensitive stick gradient, unusual coupling responses, and unstable modes [23].

Definition: *the trigger is an inseparable element of A/RPC that activates a transition of vehicle motion from steady state to oscillatory or divergent motion when the pilot applies corrective control.*

According to [3], three classes of triggers exist: environmental, vehicle and pilot triggers. The environmental and pilot triggers have historically been found to be the most frequent. However, for modern configurations, vehicle triggers have become also a threat for vehicles' safety.

3 A Short History of Rotorcraft-Pilot Couplings with a Recent Database of Events

Adverse A/RPC problems have manifested themselves since the early days of manned flight. The earliest recorded examples of PIOs date back to the Wright Brothers first aircraft [1, 24, 25]. According to [10], the earliest video record dates from just before World War II, with the XB-19 aircraft which suffered a pitch PIO on touchdown. Despite decades of work to develop methods for their prevention, unfavourable A/RPCs continue to occur. In order to better understand the incidence and the nature of A/RPCs, a database of A/RPCs cases was collected and updated from open literature, along with accidents investigation reports of the National Transportation Safety Board (NTSB) and Air Accidents Investigation Branch (AAIB). Table A.1 in Appendix A presents the database of APC events (fixed wing aircraft, gliders, and Space Shuttle), while Table B.1 in Appendix B presents the database of RPC events (helicopters, tiltrotor and gyrocopters). As can be seen from Table B.1, to date most RPC events involve large rotorcraft with conventional (non-digital) flight controls. Furthermore, many earlier recordings of RPCs were mostly associated with external underslung loads, regardless of the size of the rotorcraft. This is true, as it is now well known that unfortunate combinations of helicopter and external load dynamics can introduce new lightly damped modes which are easily excitable by the pilot. If the pilot enters the loop, these oscillations amplify and a classical divergent RPC develops. The typical solution was to drop the load, which eliminated the problem. For normal (internal load) operations, earlier recordings of RPCs were not really mentioned in the literature. Indeed, RPCs have not typically been an issue for earlier helicopters and not many were reported during operation. According to Ockier [26], three reasons may explain this.

The first main reason is that, until recently, there were no Fly-By-Wire (FBW) operational helicopters (with the exception of research vehicles). Figure 4 from [26] presents typical equivalent time delays for some early FBW research helicopters. Four configurations are presented in this figure:

1. **ADOCS (Advanced Digital/Optical Control System)**: FBW system mounted on a specially modified UH-60 helicopter, named "Light Hawk" during an Army-sponsored programme in 1980s;

2. **ARTI (Advanced Rotorcraft Technology Integration)** demonstrator using an AH-64A Apache fitted with an advanced digital FCS. This demonstrator helicopter first flew in October 1985 and was used to experiment with new technology that allowed 'hands-off' flying. The idea was that using advanced FCS, the traditional cyclic, collective and foot pedal controls were removed and replaced with a four-way sidearm controller. Using this advanced FCS, the pilot could put the helicopter into, for example, a 60 degree bank or a 2 g turn and maintain constant altitude and airspeed. With computers compensating for different torques, engine speeds etc., the pilot could concentrate outside the helicopter; he only needed to flex his hand on the sidearm controller for full manoeuvrability;

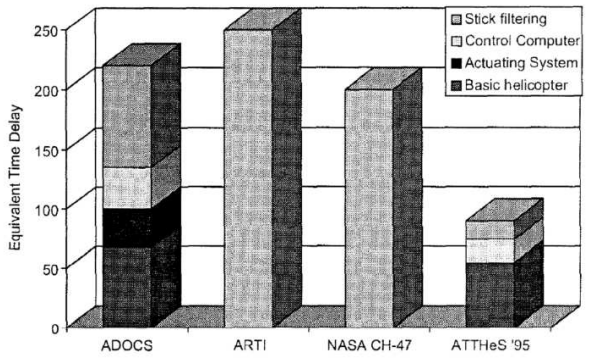


Figure 4: Typical equivalent time delays for fly-by-wire helicopters (from [26]).

3. **NASA CH-47 system**, a model-following control implemented for the first time on a tandem helicopter, the Boeing Vertol CH-47, in 1985 by NASA with the goal to improve different flying characteristics (increase task performance and simultaneously reduce pilot workload) on the helicopter on low-speed flight (later, that system was also used in the Sikorsky/Boeing RAH-66 Comanche for both low and high speed);

4. **ATTHeS (Advanced Technologies Testing Helicopter System)**, DLR's in-flight simulator, a modified BO105 helicopter equipped with a full authority non-redundant FBW control system for the main rotor and fly-by-light system for the tail rotor. In the 1990s different programmes at DLR such as automatic hovering stabilization above a moving object or use of active side-stick during different tasks [27] used this system.

Looking at Figure 4, one can see that many of these early FBW research helicopters did have high equivalent time delays (more than 200 ms, usually as the result of stick dynamics (input filtering), see also Table 1), and therefore were prone to RPCs. Indeed, for example in 1982, during the flight tests undertaken with the ADOCS system, rather severe PIOs were noted in high gain tasks such as slope landings [28].

The second reason for the absence of RPCs in earlier helicopters was that the typical APCs problems from fixed wing aircraft (e.g. fuselage bending, long control cables with bobweights, etc. [10]) were never an issue in helicopters.

The third reason probably has to do with the differences in piloting technique between helicopter and fixed wing (fighter) pilots: "highly unstable helicopters are the rule, rather than the exception and helicopter pilots are used to flying these unstable vehicles (the BO105 in hover, for instance, has an unstable phugoid with a time to double amplitude of just over 2 seconds). Helicopter pilots are generally aware of the dangers of 'getting into the loop' and tend to fly less aggressive than fixed wing (fighter) pilots. With the increase of controller sophistication in helicopters and the increase in simulator training, the helicopter pilots of the future may lose that gentle touch on the controls. When they do, PIOs may be just as frequent in helicopters as in fixed wing aircraft" [26].

Looking at Appendix B, one important conclusion appears, namely that in modern helicopters RPCs have become evident and can often be associated with couplings between the pilot and the lower flexible vehicle modes. Such was the case of the V-22 tiltrotor FSD in 2003–2004 where a divergent PAO happened, caused by a pilot who coupled with the 2.3 Hz asymmetric drive train torsion mode, primarily through the lateral stick to lateral cyclic gearing control path in the lateral axis. A secondary coupling responsible for this PAO appeared to be through the lateral cyclic to differential collective pitch control path, although this path had been treated previously in 1986 during FSD of the V-22 with a notch filter due to drive system transient loads issues [29]. To better understand the incidence and the nature of RPCs, a brief statistical survey has been performed based on the incidents of Appendices A and B. Figure 5 shows the A/RPCs cases as a function of the PIO/PAO cases. Based on this figure it can be seen that there is still a major difference between APCs and RPCs: 77% of APCs are related to PIO events not involving elasticity, whereas the RPC situation is much more entangled. At least 50% of the reports, in fact, involve aeroservoelastic phenomena (sections named PAO, PAO/PIO, flexible modes, slung loads). Moreover, a deeper analysis of the

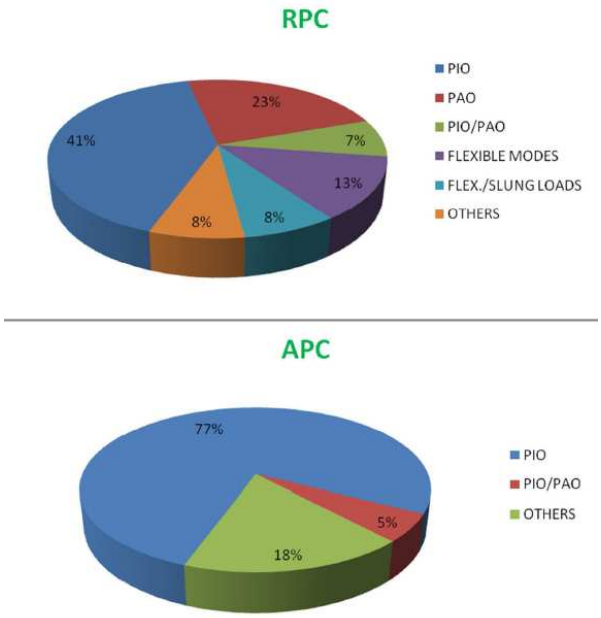


Figure 5: A/RPCs statistical analysis.

reported cases shows that also during PIOs flexibility is often inherently present, due to the interaction between the rotating frame of the rotor and the fixed frame of the airframe; thus, it is very hard to give a precise classification, as rigid and elastic phenomena are connected. As the level of automation is likely to increase and full-authority Fly-By-Wire are likely to be more common place in operational helicopters (at present operational on the NH-90, V-22 and BA609, but in the future probably also in commercial rotorcraft that hitherto have relied on manual control), it follows that more rather than fewer future RPCs are expected in the future. This situation appeared also in the 1990s for fixed-wing aircraft. McRuer ([1], pag. 87), discussing the findings of the NRC Committee in 1995, mentioned that "[...] most new flight-by-wire commercial aircraft have experienced one or more events during development, some of them severe. The sophisticated flight test instrumentation fitted to development aircraft enabled those APCs to be identified and the problems eliminated before the aircraft was put into operation. Once in service, however, the aircraft FDRs (Flight Data Recorder) and QDRs (Quick Data Recorder) cannot detect PIO problems, except in the most fortuitous circumstances (for example, the appendices of [30] report digital FDR measurements during a PIO incident occurred to an MD-11). Therefore, investigations of commercial accidents seldom mention PIO as a contributory factor. Nevertheless, there may have been a few APC-related incidents in operational service. Airbus, which has more than 700 FBW aircraft in airline operation, has more FBW experience than any other manufacturer. In all the flight hours accumulated by this fleet to date, 10 possible PIO incidents have been identified. Although Airbus acknowledges only three as genuine PIOs, the problems associated with these 10 incidents have been identified and fixed."

4 Approach to A/RPCs in ARISTOTEL

The primary focus of the ARISTOTEL project is Rotorcraft Pilot Couplings (RPCs), mainly because

- rotorcraft design tends to lag aircraft design by up to two decades, and their designers currently lack reliable criteria for specifically predicting RPC problems; furthermore,
- the particular characteristics and missions of rotorcraft may make them more prone to adverse couplings than fixed-wing aircraft.

The APC expertise of the project's fixed-wing partners (TsAGI and NLR) is leveraged to support the rotorcraft effort and to develop a general in-flight APC/RPC prediction method to be used in the early development phase of new designs. Based on GARTEUR HC AG-16 experience [5], A/RPC phenomena have been divided into two groups based on the characteristic frequency range of such phenomena, i.e. low frequency and high frequency. In the high frequency range one can discuss about 'rigid body' A/RPCs — the realm of flight dynamics — and 'aeroelastic' A/RPCs — the realm of aeroservoelasticity (see Table 2). It is assumed that a certain overlap between these two A/RPC categories exists. As an extension to the GARTEUR's approach, the frequency range consid-

ered for rigid body RPCs analysis has been raised from 1 Hz to 3.5 Hz. The reason for this, very specific to helicopters, is that, especially for hingeless rotor configurations, "the body motion 'speeds up' and the rotor dynamics enter into body dynamics" [31]. Therefore, the rotorcraft flight mechanics low frequency dynamics and the 'active' pilot concentrating on performing his/her mission task in the closed loop are affected by the low frequency rotor modes (especially regressive flapping and regressive lagging). Aeroelastic RPCs are oscillations in the bandwidth between 2 Hz and 8 Hz and correspond to higher helicopter frequency dynamics associated to elastic airframe and main rotor blade modes. Usually, for aeroelastic RPCs, a 'passive' pilot subjected to vibrations is considered. The passive pilot is usually modelled considering the pilot impedance in the form of a transfer function relating cabin movement to pilot input displacements. It is considered that an active pilot model for an aeroelastic RPC would require modelling of the pilot neuromuscular and cognitive system and this approach would not enhance the understanding of how vibrations affect pilot controls as they are too high in frequency to be adequately reacted by a human. For the frequency overlap range, the above mentioned classifications of RPCs correspond to a merger of the phenomena, with a mix of models and procedures for A/RPC detection. In ARISTOTEL's approach as presented above, the main rotor blade modes are seen from a 'rigid body' point of view in the non-rotating system, and as 'aeroelastic' modes in the rotating system. It is thought that a parallel rigid body/aeroelastic approach may enhance the understanding of RPC phenomena in the critical range of 1–3.5 Hz, where many accidents have been observed [29].

5 Example of a Classic RPC Problem

One "classical" adverse coupling characteristic observed in rotorcraft is the excitation of the low-damped main rotor regressive inplane mode by pilot's cyclic inputs which results in body roll and pitch aircraft vibrations affecting blade strength limits. This problem, which is actually an air resonance of the inplane regressive/blade bending modes, was categorized in some references as PAO [29] and sometimes as PIO (as it does not involve biodynamic couplings, see [5]). The instability was observed for the first time in 1967 on the H-46D Sea Knight, a tandem rotor helicopter manufactured by Boeing-Vertol [29]. The instability existed at 3.2 Hz at airspeeds near the never exceed speed (VNE) and at low speed with high descent rates. It was created by the lightly damped main rotor regressive lag mode at 3.2 Hz, causing out-of-phase behaviour of the forward and aft rotors [32], coupling with the aft pylon's fuselage mode, a lateral bending/torsion mode. The instability was described by the pilots as a "shuffle" about the lateral, roll and yaw axes. After closer observation of the instability, it appeared that the AFCS system, introduced in CH-46D for flight path stabilization, and the amount of lag damping significantly affected the behaviour of the 3.2 Hz lag mode with airspeed. "Lateral cyclic characteristics were also affecting the instability as the pilot's hand mass or effective grip force on the cyclic could create pilot feedback instability at 3.2 Hz. [...] The critical results of the flight test program indicated that the source of the instability was insufficient lag damping in the rotor system with the significant initiation and/or sustaining mechanism being airframe structural flexibility, pilot lateral cyclic inputs and the yaw SAS. The results of the 3.2 Hz shuffle mode flight test program recommended the following changes to the CH-46D aircraft: a significant increase in blade lag damping and the implementation of a 3.2 Hz notch filter to suppress excitation of the mode. Due to the ongoing Vietnam War, none of these options were exercised and instead procedural mitigation was imposed recommending a reduction in airspeed, collective setting and/or manoeuvre severity to relieve the oscillation" [29]. The same regressive inplane mode air resonance instability was reported also by [33], this time at 1.8 Hz for the EC-135 helicopter. The air resonance mode of the EC-135 is described as a low-frequency mode characterizing the coupling between regressing lead-lag mode and body roll motion. Such a mode is characterizing especially the soft-inplane hingeless/bearingless rotor heli-

Table 2: Characterisation of 'Rigid Body' and 'Aeroelastic' A/RPC

	Low frequency A/RPCs	**High frequency A/RPCs**	
		Rigid body aircraft	**Elastic body aircraft**
Frequencies	Below 1.5 Hz APC frequencies are usually within 0.5–1.6 Hz (3–10 radian/sec)	Between 1.5–2 Hz (APC) Below 3.5 Hz (RPC) APCs frequencies usually exceed 2 Hz; examples: roll ratchet, bob-weight.	Between 2–8 Hz
Causes	1) Inadequate vehicle dynamic characteristics (aircraft + control system): · high order of the system, large phase delay, low damping, and others; · control system delay; · actuator or control surface rate limit. 2) High control sensitivity (command gain), low force-displacement gradient.	1) Biodynamic interaction: the biodynamic interaction in the "pilot + manipulator + aircraft" system arises due to high-frequency aircraft response to pilot activity caused by inadequate aircraft characteristics (high natural frequencies, low roll mode time constant, high control sensitivity, relative pilot location far away from the centre of gravity)	1) Biodynamic interaction: The biodynamic interaction in the pilot-aircraft system arises due to aircraft structural elasticity and leads to involuntary manipulator deflections transferred to control system.
Characteristics	Pilot closes the loop according to the information received through visual or acceleration perception channels.	The pilot closes the control loop due to aircraft accelerations acting on the body and the arm cause involuntary manipulator deflections which go to the control system and lead to high-frequency A/RPC.	The pilot closes the control loop due to structural oscillations and inertial forces acting on the body and the arm cause involuntary manipulator deflections which go to the control system and provoke high-frequency A/RPC.
Critical components		Flight control system	Airframe modes
Pilot modeling	'Active' pilot concentrating on a task	'Active' pilot concentrating on a task	'Passive' pilot subjected to vibrations
Vehicle dynamics modeling	Flight mechanics	Flight mechanics	Structural dynamics

copters (BO105, RAH-66 Comanche or EC-135) where the lead-lag motion is weakly damped. According to [33], it appeared that, in the basic helicopter operation condition, air resonance was not an issue for the pilots operating the EC-135, the air resonance instability manifesting as a body roll oscillation which was existent but below the pilot perception level. However, when the helicopter was enhanced with an Attitude command/Attitude Hold (ACAH) control system for flying attitude command or flight path following tasks, it became apparent that, increasing too much the roll rate feedback gain, drove the air resonance mode unstable. This time the oscillation was perceived by the pilot as an oscillatory ringing in the helicopter roll response at a frequency of about 1.8 Hz in the case of the EC-135. It was demonstrated that in this case the helicopter was PIO prone when applying the ADS-33 bandwidth criterion. The instability was eliminated by developing an air resonance controller which damped the coupled body-roll air resonance mode when rate feedback was used, independently from the main flight control system. Dryfoos et al., discussing the same air resonance problems, concluded that "slow, stiff rotors would clearly be the most susceptible to the destabilizing effect of roll attitude to lateral cyclic feedback gain" [34].

The physics of this mechanism of instability is explained in [35, 36] by defining so-called 'paths of energy' through which 'vicious' circles of energy transfer are formed between the flap-lag-roll degrees of freedom. It was demonstrated that, when using a controller, the roll motion pumps energy into the flap but there is no energy being pumped back from flap to roll (no instability problem). Especially in a soft-flapwise stiff-inplane rotor, the flap-lag motions pump energy into each other. Although the roll-lag motions are not directly related (the roll can pump energy into the lag but not the other way around), it is possible for the fuselage-lag mode to be driven unstable for a certain value of roll rate gain.

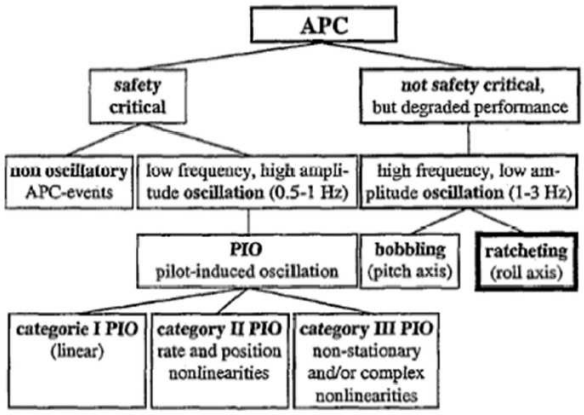


Figure 6: Classification of aircraft/pilot coupling phenomena (for fixed-wing aircraft) [37].

6 Four Categories of A/RPCs

McRuer divided A/RPCs into three categories (Cat. I, II and III) according to the degree of non-linearity of the oscillation of the Pilot-Vehicle System (PVS) [1]. Many researchers have adopted this method since the above mentioned classification; for example, Figure 6, from [37], presents the classification of APC according to McRuer's three main A/RPCs categories. This classification is also illustrative for rotorcraft. Recently, [38] suggested the introduction of a fourth A/RPC category for events that are caused by, or receive a major contribution from structural modes and their interaction with the pilot. These events, also referred to as PAOs, are of special interest for rotorcraft [29]. The four categories are explained in the following sections.

6.1 Category I A/RPC: Essentially Linear PVS Oscillations

A/RPCs in this category are essentially linear and are directly caused by excessive time delays or phase lags in the vehicle dynamics. These are typically a consequence of digital filtering, an improper aircraft or rotorcraft gain (too sensitive or too sluggish), resulting in overall poor handling qualities. Triggers usually occur during high gain tasks. These tasks require frequent small corrections from the pilot and thus increase the pilot's workload. Examples of high

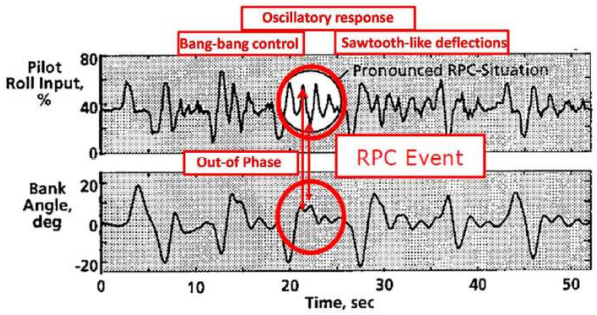


Figure 7: BO105 ATTHeS roll attitude tracking tasks with 160 ms added time delay (taken from [16]).

gain task are the slope landing for rotorcraft, and aerial refueling. Typical frequencies of Category I A/RPC are between 0.3 Hz and 1.5 Hz [1]. A/RPCs in this category are relatively simple to model and best understood. Almost all existing criteria with respect to A/RPC focus on Category I. These types of A/RPCs are least common during operational flying [1,39,40]. An example of an RPC in this category was induced during flight testing with Bölkow BO105 ATTHes helicopter (see [16] and 16). The RPC took place during a slalom task and was caused by an added time delay of 160 ms in the pilot input. The time history is shown in Figure 7. Looking at this figure, one should observe the typical signatures in the time histories of A/RPC events:

1. oscillatory characteristics;
2. the pilot roll inputs and the bank angle are out-of-phase;
3. after the RPC is triggered, the pilot stick input exhibits bang-bang control (max-min or on-off control), increasing the closed loop gain and destabilizing the system even more. It can be said that the pilot is behaving synchronous with the response;
4. the saw-tooth like deflections, indicating control rate limiting.

To demonstrate that the time delay caused the RPC, [26] plotted the time histories of the lateral position tracking task without and with a

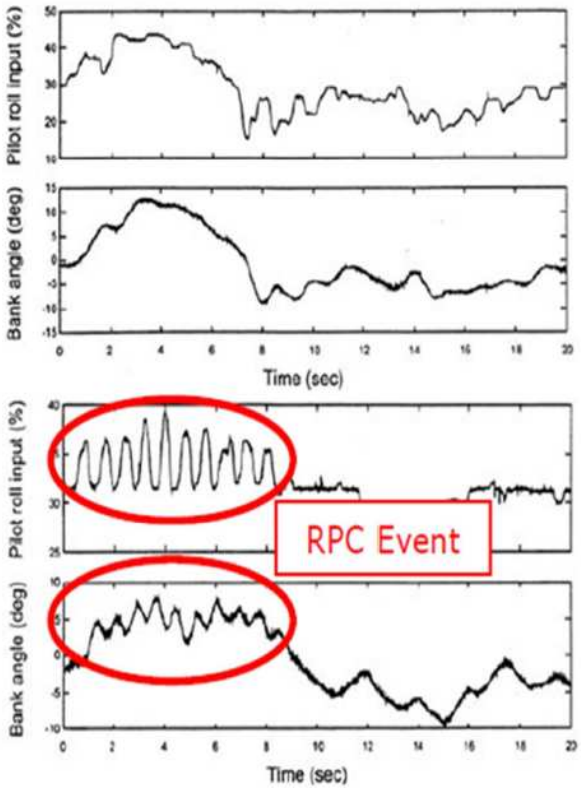
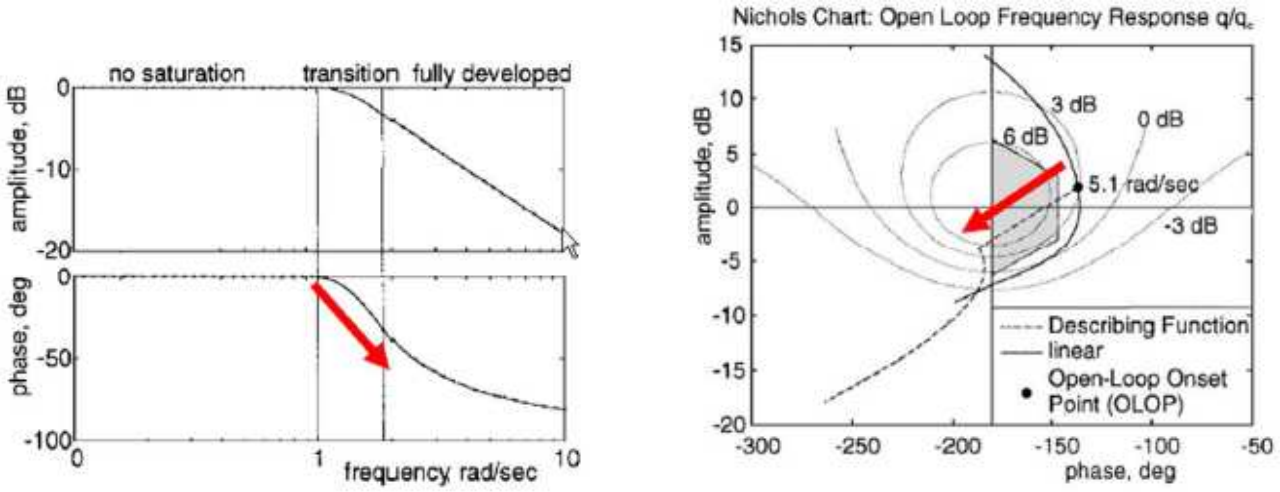


Figure 8: BO105 ATTHeS lateral position tracking tasks without and with 100 ms added time delay [26].

100 ms added time delay (see Figure 8). During the task, the pilot had to track the relative position with respect to a moving vehicle, while flying sideways. It was demonstrated that this RPC (1.2 Hz) was caused by a combination of excess time delay and biodynamic coupling between the pilot's arm and the lateral accelerations of the rotorcraft. During A/RPCs events, pilots mentioned feeling "disconnected from the stick" or "suspecting" aircraft failures [11]. This confirms the suspicion that the proposed word in the definition relating to mental mismatch is key for triggering and sustaining A/RPC events.

6.2 Category II A/RPC: Quasi-Linear PVS Oscillations

A/RPCs in this category are quasi-linear events and are triggered by the nonlinear Rate and/or Position Limiting Elements (RLEs and/or PLEs). Vehicle dynamics are linear until onset, hence the term quasilinear. Typical



(a) Bode plot

(b) Nichols chart

Figure 10: Phase jump after onset (taken from [41]).

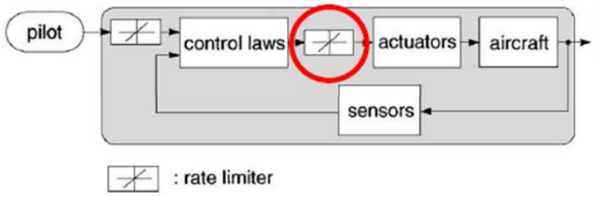


Figure 9: Typical locations of rate limiting elements (taken from [41]).

RLEs can be found in digital FCS or in actuator dynamics as shown in Figure 9 from [41]. After onset of an RLE (trigger) which is usually caused by a large pilot input, time delays build-up fast, causing the discrepancy between the pilot's input and the intended response to develop quickly. The term "cliff-like" behaviour is frequently used [1, 40]. After onset, the phase lag exhibits a jump. This is sometimes referred to as the "jump phenomenon" [41, 42]. This jump is clearly visible in the Bode and Nichols plots in Figure 10. In the time domain, this build-up is visualized in Figure 11. The saw tooth shape is the signature of the rate limiter being active. Although frequencies of the oscillation typically vary for each aircraft and RLEs, many ARC occurrences have a frequency of around 0.5 Hz [1,9,43]. For rotorcraft, such RPC frequency is higher, around 1.8–3 Hz. The relatively new criteria for this category are based on, for example, the use of a describing function for the nonlinear element [44–46] or the Open Loop Onset Point (OLOP) criterion [47]. The effects of rate limiting have either caused or sustained most APC events in the past, like in the YF-22 example [1].

6.3 Category III A/RPC: Essentially Nonlinear PVS Oscillations with Transitions

A/RPC events in this category are triggered by mode or task switching or changes in the aerodynamic configurations (for example flaps, gear, etc.) or propulsion system. This switching is non-linear. For example, shifts or transitions in command type of the FCS cause the mental mismatch to develop. In helicopters with FBW and digital control, there have been RPC occurrences when the command type switched from attitude command to rate command in a Weight-on-Wheels situation [48, 49]. The same situation happened for the fixed wing F-8 DFBW (Digital Fly-By-Wire) test aircraft [1]. Due to the nonlinearities and the fact that dynamics or tasks change, A/RPC occurrences in this category are most difficult to analyze offline [1]. Criteria specifically designed for this category are practically non-existent. The YF-22 APC case and the XV-15 and V-22 divergent

lateral oscillations on the landing gear during ground taxi operations (the first one predicted only on paper, the other encountered during the flight test program) can be included in this category [29].

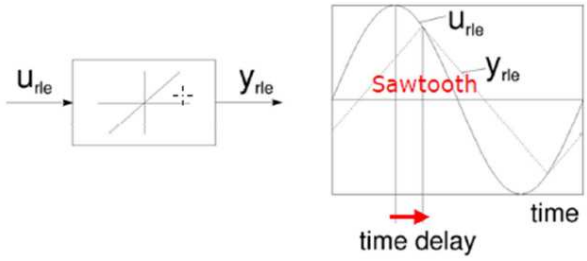


Figure 11: Time delay build-up due to rate limiting (taken from [10]).

6.4 Category IV A/RPC: Oscillations Due to Elastic Structural Modes or Biodynamical Couplings

A/RPCs in this category are due to the coupling of elastic structural modes (aeroelastic) and the pilot or due to biodynamical couplings. They are a "faster" type of A/RPC events with frequencies of at least 1 Hz [9]. This category includes oscillations with a fully attentive or passive pilot in the loop, as they are caused by an involuntarily or passive interaction between the pilot, typically his limbs, and the vibratory motion of the vehicle. The fourth category also corresponds to the so-called "biodynamic couplings", involving structural or aeroelastic modes of the aircraft [39, 40]. In case of large transport aircraft, the pilot might excite the aircraft's structural modes and possibly regresses into an A/RPC event. Common in rotorcraft are the couplings between the pilot and the vehicle dynamics with an external slung load [1, 29]. Other examples for aircraft and rotorcraft can be found in [50–53]. In case of vibration feedthrough to the cockpit and biodynamical coupling, the pilot's limbs are shaken, causing passive and involuntary control inputs. A/RPC events of this kind are called Pilot-Assisted Oscillations (PAOs). Rotorcraft are especially prone to these types of RPCs, due to relatively high-amplitude vibratory environment. In [29] an overview of these events regarding US Navy rotorcraft operations was presented. In [50], a situation is presented where the dynamics of the pilot's arm and the collective handle couples with the rotorcraft vertical response. The example of the BO105 RPC event that was shown in Figures 7 and 8 belongs to this category. In conclusion, there are many different kinds of A/RPCs. Thus, when discussing about an A/RPC event, there is not only one kind of A/RPC that can happen or not. On the contrary, a whole range of A/RPC types needs to be considered, starting from minor but annoying A/RPCs to dangerous ones. "To paint all PIOs with a single brush is to run the risk of panicking and rushing to judgment on the basis of a benign, common event, or doing the opposite: trying to whitewash a serious and potentially deadly design flaw" [15]. Therefore, A/RPCs can be generally considered either safety critical or non-safety critical to aircraft operations.

7 Future Rotorcraft-Pilot Couplings

After an extensive investigation and review, the ARISTOTEL partners were able to identify some critical items for present and future trends in A/RPCs:

1. newer design requirements,
2. evolution of possible technical solutions,
3. evolution of certification requirements.

7.1 Newer Design Requirements

Stability vs. manoeuvrability: increasing the manoeuvrability requirements for aircraft and rotorcraft could be observed in the mission requirements evolution of civil and military customers. For example, new requirements of rotorcraft civil customers demand longer flight time periods over urban and mountainous regions. Because of possible obstacles and low flight profile missions (in case of law enforcement, passengers, medical transport and underslung load missions), an increase in helicopter manoeuvrability is recommended. This concern is especially relevant for missions performed providing Emergency Medical Services (EMS), where

there is a requirement for the possibility to land on short fields with limited area and high obstacles. For military customers, good helicopter handling qualities for Nap-of-the-Earth (NoE) missions are required in order to increase survivability on the battlefield. Also, attack and reconnaissance helicopter operations require dynamic manoeuvring capabilities. An increase in helicopter's manoeuvrability results in a decrease in its static stability margins and therefore a decrease in the reaction time available to the pilot. This may affect RPC occurrence as well.

Flight Envelope Expansion: another trend observed for both aircraft and rotorcraft is the extension of the flight envelope, i.e. the increase of speed (as well as the decrease of approach speed for fixed wing aircraft), altitude, ambient temperature and range. Several areas can be problematic for A/RPCs:

1. increase or maintain the never exceed speed (VNE) and flight altitude combined with the actual trend to decrease the main and tail rotor tip speed (for environmental, 'greening' reasons) could cause additional A/RPC problems not met in earlier designs as the rotors are closer to stall conditions;

2. improvement of capabilities for "Category A" rotorcraft such as reduction of necessary airfield size through specific control strategy closer to the vortex ring state can trigger future A/RPC problems.

Increase of AFCS autonomy: currently, there is a tendency to increase the capabilities of the AFCS and to include larger possible classes of manoeuvres performed with such systems. This is to develop towards the end goal of an optionally fully autonomous vehicle. Cooperation between the pilot with more sophisticated control system and vehicle equipped with these types of AFCS could be the source of possible A/RPC problems. Implementation of reliable pilot mathematical models combined with proper flight mechanics models in the future should investigate AFCS failure modes. This is especially true during emergency situations when the pilot needs to react properly to hardware faults/unexpected flight situations [54]. Ref. [2] is a good example of the trade-offs that must be accepted in rotorcraft FCS design.

Reduction of noise and vibration levels: new rotorcraft, designed to meet 'green' requirements, could lead to more A/RPC problems. The future design of new aerial vehicles — such as heavy lift rotorcraft or large transport aircraft — relates to the development of more flexible conventional structures, or new structural design paradigms. The overall FCS must include the more pronounced flexibility effect in its design. The reason for this is that the natural frequencies of the fuselage and wing/rotor blade structural modes decrease as their size increase. As a consequence, the lower frequency structural modes have a greater influence on the dynamic response of the vehicle. Additionally, weight reduction through the use of composite materials contributes to the development of more flexible structures. The structural flexibility also affects the aeroelastic stability of the vehicle where the pilot biodynamic feedback and FCS feedback can interact with the structure of the vehicle, leading to pilot/control system assisted excitation of the structural modes.

7.2 Evolution of Possible Technical Solutions

Advanced main and tail rotor schemes: designs for more controllable helicopters has resulted in stiffer main rotor hub structures. The rotor design evolution from fully articulated to hingeless and bearingless solutions results in an increase of the number of rotor modes affected by the response of the pilot. The implementation of new rotor control techniques such as Higher Harmonic Control HHC [55] or Individual Blade Control IBC [56] also increase the need for RPC analysis. Solutions like swashplateless rotors [56] may introduce new types of instabilities which can trigger RPCs. Tailrotor replacement by Fenestron, NOTAR or other solutions may also cause problems, and require RPC sensitivity analysis to discover specific critical operating conditions.

Increasing the role of electronics in cockpit design: commercial aircraft manufac-

turers are veering towards FBW control technologies. It is well known that high automation in the cabin reduces situational awareness [57]. While FBW can significantly enhance the aircraft manoeuvrability, it also increases controller bandwidth. This may result in adverse interactions between the human pilot-flight control system-aircraft dynamics. These interactions become more critical in the case of structural spill-over instabilities, due to poor control laws designs or FCS updates incompatible with the airframe. This highlights the need for robust control design techniques and effective analysis methods. Design of more autonomous AFCS with larger authority margin may lead to future A/RPCs. Design analysis should answer several questions such as:

1. is the pilot capable of maintaining control with partial/full AFCS out of order?

2. What is the pilot time delay for overruling the AFCS when needed and what are the appropriate parameters to be used for tuning the AFCS for safe operation?

3. What are the most critical flight states when malfunctions of AFCS occur?

Answering these questions requires appropriate pilot mathematical models and to complement extensive simulations with real human pilot participation.

Smart structures and smart materials incorporation into design: new types of adaptive structures used on aircraft and helicopters as well as additional controls could add new degrees of freedoms into A/RPC analyses

7.3 Evolution of certification requirements

Manoeuvrability requirements evolution: there is a tendency to introduce new requirements in certification documents for performing specific manoeuvres. MIL-H-8501A and later MIL-F-83300 standards, which were applicable in the past, defined limits for helicopter responses based on pilot control input. ADS-33E [20] and, more recently, Handling Qualities of Rotorcraft with External Slung Loads [58], defined handling qualities per specific manoeuvres. Presently, very little requirements are given with respect to RPCs.

Civil design requirements evolution: the evolution of civil helicopter design regulations — FAR 23/25, JAR 27/29, CS 27/29 for aircraft and FAR 27/29, JAR 27/29, CS 27/29 for rotorcraft — shows that the stability requirements have strengthened in time (such as damping requirements, pilot response delay time with AFCS). Fulfilling both the stability specs for the civilian market and manoeuvrability specs for military operators requires the use of trade-offs. Future A/RPC analyses could help to find optimal control strategies for the physical capabilities of pilots.

Environmental requirements evolution: the 'green' environmental aspects, such as restrictive noise, vibration, pollution requirements, new structures, materials, systems, control strategies etc., which are presently not included in the design but might become compulsory in the future, could affect the A/RPC level. The ACARE agenda of the European Community [59] together with the JTI "Clean Sky" initiatives (`http://www.cleansky.eu/`) for implementing new technologies, smart structures, new materials, new control strategies for take-off and approach flight paths in both fixed and rotary wing aircraft (increased manoeuvrability, higher climb and descend ratios near airfields and helipads) may affect the tendency for A/RPCs.

8 Conclusions

This paper presents a review of the current status on A/RPC problem and how it was tackled in the ARISTOTEL project. This European Community project, started in October 2010 and running into 2013, intends to give design and simulator guidelines for investigating rigid-body and aeroelastic A/RPCs of future aircraft and rotorcraft. The present paper presented a new extended database of A/RPC events and proposed a new definition for these phenomena. An extensive discussion on future A/RPCs shows that, modern designs can be more A/RPC prone than their predecessors; therefore, design-

ers should be acquaint with specific knowledge for understanding these problems.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007–2013] under grant agreement nº ACPO-GA-2010-266073.

References

[1] D. T. McRuer. *Aviation Safety and Pilot Control: Understanding and Preventing Unfavourable Pilot-Vehicle Interactions*. Washington DC: National Research Council, National Academy Press, 1997.

[2] Mark B. Tischler, Christina M. Ivler, M. Hossein Mansur, Kenny K. Cheung, Tom Berger, and Marcos Berrios. Handling-qualities optimization and trade-offs in rotorcraft flight control design. In *RAeS Rotorcraft Handling-Qualities Conference*, Liverpool, UK, November 4–6 2008.

[3] GARTEUR FM AG-12. Pilot-in-the-loop oscillations — analysis and test techniques for their prevention, phase I. In *Final Workshop*, EADS Manching, Germany, October 11 2001.

[4] GARTEUR FM AG-15. Pilot-in-the-loop oscillations analysis and test techniques for their prevention, phase II. In *Workshop*, Saab AB, Sweden, May 10 2007.

[5] O. Dieterich, J. Götz, B. DangVu, H. Haverdings, P. Masarati, M. Pavel, M. Jump, and M. Gennaretti. Adverse rotorcraft-pilot coupling: Recent research activities in europe. In *34th European Rotorcraft Forum*, Liverpool, UK, September 16–19 2008.

[6] M. D. Pavel, B. DangVu, J. Götz, M. Jump, O. Dieterich, and H. Haverdings. Adverse rotorcraft-pilot coupling: Prediction and suppression of rigid body RPC. In *34th European Rotorcraft Forum*, Liverpool, UK, September 16–19 2008.

[7] M. Gennaretti, J. Serafini, P. Masarati, G. Quaranta, and O. Dieterich. Aeroelastic and biodynamic modeling for stability analysis of rotorcraft-pilot coupling phenomena. In *34th European Rotorcraft Forum*, Liverpool, UK, September 16–19 2008.

[8] M. Jump, S. Hodge, B. DangVu, P. Masarati, G. Quaranta, M. Mattaboni, M. Pavel, and O. Dieterich. Adverse rotorcraft-pilot coupling: The construction of the test campaigns at the university of Liverpool. In *34th European Rotorcraft Forum*, Liverpool, UK, September 16–19 2008.

[9] AGARD. Handling qualities of unstable highly augmented aircraft. AR 279, AGARD, January 1991.

[10] AGARD. PIO workshop following active control technology: Applications and lessons learned. CP 560, AGARD, 1995.

[11] AGARD. FVP workshop pilot induced oscillations. AR 335, AGARD, February 1995.

[12] D. T. McRuer. Pilot-induced oscillations and human dynamic behavior. CR 4683, NASA, 1995.

[13] M. Shafe and P. Steinmetz. Pilot induced oscillations: Status at the end of the century. CD 2001-210389/VOL 1–3, NASA, Edwards, California, 2001.

[14] S. F. Ali. A low cost simulation system to demonstrate pilot induced oscillation phenomenon. NAG2 4006, NASA Dryden Flight Research Center, March 31 1995.

[15] David G. Mitchell and David H. Klyde. Identifying a PIO signature — new techniques applied to an old problem. In *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Keystone, Colorado, USA, August 21–24 2006. AIAA-2006-6495.

[16] P. G. Hamel. Rotorcraft-pilot coupling — a critical issue for highly augmented helicopters? In *AGARD-FVP Symposium "Advances in Rotorcraft Technology"*, Ottawa, Canada, May 27–30 1996.

[17] D. T. McRuer and H. R. Jex. A review of quasi-linear pilot models. *Human Factors in Electronics, IEEE Transactions on*, HFE-8(3):231–249, September 1967. doi:10.1109/THFE.1967.234304.

[18] Mark B. Tischler. System identification requirements for high-bandwidth rotorcraft flight control system design. *J. of Guidance, Control, and Dynamics*, 13(5):835–841, 1990. doi:10.2514/3.25409.

[19] Mark B. Tischler, Jay W. Fletcher, Patrick M. Morris, and George E. Tucker. Flying quality analysis and flight evaluation of a highly

augmented combat rotorcraft. *J. of Guidance, Control, and Dynamics*, 14(5):954–963, 1991. doi:10.2514/3.20736.

[20] Performance specification, handling qualifies requirements for military rotorcraft. ADS 33-E-PRF, US Army AMCOM, Redstone, Alabama, 2000.

[21] Gloria K. Yamauchi and Larry A. Young. A status of NASA rotorcraft research. TP 2009-215369, NASA, September 2009.

[22] John W. Smith and Terry Montgomery. Biomechanically induced and controller coupled oscillations experienced on the F-16XL aircraft during rolling maneuvers. TM 4752, NASA, July 1996.

[23] Michael Parrag. Use of in-flight simulators for PIO susceptibility testing and for flight test training. CP 2001-210389, NASA, Dryden FRC, Edwards, CA, USA, April 1999. PIO Workshop Pilot-Induced Oscillation Research: Status at the End of the Century.

[24] Ben Lawrence and Gareth Padfield. Handling qualities analysis of the Wright brothers' 1902 glider. *Journal of Aircraft*, 42(1):224–236, 2005. doi:10.2514/1.6091.

[25] Ben Lawrence and Gareth Padfield. Flight handling qualities of the Wright brothers' 1905 Flyer 3. *Journal of Aircraft*, 43(5):1307–1316, 2006. doi:10.2514/1.19607.

[26] C. J. Ockier. Pilot induced oscillations in helicopters — three case studies. Technical Report IB 111-96/12, German Aerospace Center (DLR), Braunschweig, Germany, 1996.

[27] Gerd Bouwer and Heinz-Jürgen Pausder. *In-Flight Simulation with Helicopters*, volume 147, pages 617–625. Springer, 2004.

[28] Mark B. Tischler, Jay W. Fletcher, Patrick M. Morris, and George T. Tucker. Applications of flight control system methods to an advanced combat rotorcraft. TM 101054, NASA, July 1989.

[29] R. Barry Walden. A retrospective survey of pilot-structural coupling instabilities in naval rotorcraft. In *63rd Annual Forum of the American Helicopter Society*, Virginia Beach, VA, May 1–3 2007.

[30] Aircraft Accident Investigation Commission. Japan airlines flight 706 McDonnel Douglas MD-11, JA8580, over Shima peninsula, Japan, June 8, 1997. Technical Report 99-8, Ministry of Transport, Japan, December 17 1999.

[31] H. C. Curtiss, Jr. Stability and control modelling. In *12th European Rotorcraft Forum*, Garmisch-Partenkirchen, Germany, September 22–25 1986. Paper no. 41.

[32] David G. Miller and Fred White. A treatment of the impact of rotor-fuselage coupling on helicopter handling qualities. In *43rd AHS Annual Forum*, St. Louis, MO, USA, May 18–20 1987.

[33] Robin Lantzsch, Jens Wolfram, and Mario Hamers. Increasing handling qualities and flight control performance using an air resonance controller. In *64th AHS Annual Forum*, Montreal, Canada, April 29–May 1 2008.

[34] James B. Dryfoos, Bruce D. Kothmann, and John Mayo. An approach to reducing rotor-body coupled roll oscillations on the RAH-66 Comanche using modified roll rate feedback. In *55th AHS Annual Forum*, pages 1127–1140, Montreal, Canada, May 25–27 1999.

[35] Marilena D. Pavel and Gareth D. Padfield. Understanding the peculiarities of rotorcraftpilot couplings. In *64th AHS Annual Forum*, Montreal, Canada, May 2008.

[36] Marilena D. Pavel. Modeling lead-lag dynamics for rotorcraft-pilot-couplings investigation. In *66th AHS Annual Forum*, Phoenix, Arizona, USA, May 11–13 2010.

[37] Gordon Höhne. A biomechanical pilot model for prediction of roll ratcheting. In *AIAA Atmospheric Flight Mechanics Conference*, Portland, OR, USA, August 9–11 1999. AIAA-1999-4092.

[38] David H. Klyde and David G. Mitchell. A PIO case study – lessons learned through analysis. In *AIAA Atmospheric Flight Mechanics Conference*, San Francisco, California, USA, August 15–18 2005. AIAA-2005-5813.

[39] David H. Klyde and David G. Mitchell. Investigating the role of rate limiting in pilot-induced oscillations. *J. of Guidance, Control, and Dynamics*, 27(5):804–813, 2004. doi:10.2514/1.3215.

[40] David G. Mitchell and David H. Klyde. Recommended practices for exposing pilot-induced oscillations or tendencies in the development process. In *USAF Developmental Test and Evaluation Summit*, Woodland Hills, California, USA, November 16–18 2004. AIAA-2004-6810.

[41] Holger Duda. Prediction of pilot-in-the-loop oscillations due to rate saturation. *J. of Guidance, Control, and Dynamics*, 20(3):581–587, 1997. doi:10.2514/2.4080.

[42] Holger Duda. Effects of rate limiting elements in flight control systems — a new PIO-criterion. In *AIAA Guidance, Navigation and Control Conference*, Baltimore, MD, August 7–10 1995. AIAA-1995-3204.

[43] Flight control design — best practices. TR 029, RTO, 2000.

[44] David H. Klyde, Duane T. McRuer, and Thomas T. Myers. Pilot-induced oscillation analysis and prediction with actuator rate limiting. *J. of Guidance, Control, and Dynamics*, 20(1):81–89, 1997. doi:10.2514/2.3998.

[45] David H. Klyde, Duane T. McRuer, and Thomas T. Myers. PIO analysis with actuator rate limiting. In *AIAA Atmospheric Flight Mechanics Conference*, San Diego, CA, USA, July 29–31 1996. AIAA-1996-3432.

[46] Francesco Amato, Raffaele Iervolino, Stefano Scala, and Leopoldo Verde. Category II pilot in-the-loop oscillations analysis from robust stability methods. *J. of Guidance, Control, and Dynamics*, 24(3):531–538, 2001. doi:10.2514/2.4743.

[47] Holger Duda and G. Duus. New handling qualities database on PIO due to rate saturation. FB 97-53, DLR, 1997.

[48] Carl J. Bauer. A landing and takeoff control law for unique-trim, fly-by-wire rotorcraft flight control systems. In *19th European Rotorcraft Forum*, pages H10–1–9, Como, Italy, September 14-16 1993.

[49] L. R. Stiles, J. Mayo, A. L. Freisner, K. H. Landis, and B. D. Kothmann. "Impossible to Resist" the development of rotorcraft fly-by-wire technology. In *60th AHS Annual Forum*, Baltimore, MD, June 8-10 2004.

[50] John R. Mayo. The involuntary participation of a human pilot in a helicopter collective control loop. In *15th European Rotorcraft Forum*, pages 81.1–12, Amsterdam, The Netherlands, 12–15 September 1989.

[51] T. Parham, Jr. and David Popelka. V–22 pilot-in-the-loop aeroelastic stability analysis. In *47th Annual Forum of the American Helicopter Society*, Phoenix, Arizona (USA), May 6–8 1991.

[52] Gordon Höhne. *Roll Ratcheting: Cause and Analysis*. DLR, 2001.

[53] David L. Raney, E. Bruce Jackson, William M. Adams, and Carey S. Buttrill. The impact of structural vibration on flying qualities of a supersonic transport. In *AIAA Atmospheric Flight Mechanics Conference*, Montreal, Canada, August 6–9 2001. AIAA-2001-4006.

[54] Neil Cameron and Gareth D. Padfield. Handling qualities degradation in tilt-rotor aircraft following flight control system failures. In *30th European Rotorcraft Forum*, Marseilles, France, September 14-16 2004.

[55] Ch. Kessler. Active rotor control for helicopters: motivation and survey on higher harmonic control. *CEAS Aeronaut. J.*, available online, 2011. doi:10.1007/s13272-011-0005-9.

[56] Ch. Kessler. Active rotor control for helicopters: individual blade control and swashplateless rotor designs. *CEAS Aeronaut. J.*, available online, 2011. doi:10.1007/s13272-011-0001-0.

[57] Richard W. Pew. More than 50 years of history and accomplishments in human performance model development. *Human Factors*, 50(3):489–496, June 2008. doi:10.1518/001872008X288493.

[58] Roger H. Hoh, Robert K. Heffley, and David G. Mitchell. Development of handling qualities criteria for rotorcraft with externally slung loads. CR 2006-213488, NASA, October 2006.

[59] ACARE — Report of the Group of Personalities. European aeronautics: a vision for 2020, January 2001.

[60] Arthur Pearcy. *Flying the frontiers: NACA and NASA experimental aircraft*. Naval Institute Press, Annapolis, MD, 1993.

[61] Thomas W. Finch and Gene J. Matranga. Launch, low-speed, and landing characteristics determined from the first flight of the North American X-15 research airplane. TM X-195, NASA, 1959.

[62] Gene J. Matranga. Analysis of X-15 landing approach and flare characteristics determined from the first 30 flights. TN D-1057, NASA, 1961. H-221.

[63] Irving L. Ashkenas, Henry R. Jex, and Duane T. McRuer. Pilot-induced oscillations: Their cause and analysis. NCR 64-143, Norair Report, June 20 1964.

[64] M. J. Abzug and H. B. Dietrick. Interim report on elimination of pilot-lnduced oscillations from the Douglas model A4D-2 airplane. Report ES-26613, Douglas Aircraft Company, March 20 1957.

[65] W. H. Terrill, J. G. Wong, and L. R. Springer. Investigation of pilot-induced longitudinal oscillation in the Douglas model A4D-2 airplane. Report LB-25452, Douglas Aircraft Company, May 15 1957.

[66] Charles C. Crawford and Jones P. Seigler. KC-135A stability and control test. TR 58-13, AFFTC, May 1958.

[67] Carl D. Simmons and Donald M. Sorlie. F-101B air force stability and control evaluation. TR 58-11, AFFTC, May 1958.

[68] J. W. Smith and D. T. Berry. Analysis of longitudinal pilot-induced oscillation tendencies of YF-12 aircraft. TN D-7900, NASA, 1975. H-805.

[69] H. R. Jex. Summary of T-38A PIO analysis. TR 239-1, Systems Technology, Inc., January 1963.

[70] O. A. Levi and W. E. Nelson. An analytical and flight-test approach to the reduction of pilot-induced oscillation susceptibility. *Journal of Aircraft*, 1(4):178–184, 1964. doi:10.2514/3.43579.

[71] D. L. Hirsch and R. L. McCormick. Experimental investigation of pilot dynamics in a pilot-induced oscillation situation. *Journal of Aircraft*, 3(6):567–573, 1966. doi:10.2514/3.43778.

[72] C. R. Chalk. Another study of the T-38A PIO incident. In *AFFDL Flying Qualities Symposium*, Dayton, OH, USA, September 12–15 1978.

[73] W. Hewitt Phillips. Journey in aeronautical research: A career at NASA Langley research center. In *Monographs in Aerospace History*. NASA, 1998.

[74] R. W. Kempel. Analysis of a coupled roll spiral mode, pilot induced oscillation experienced with the M-2F2 lifting body. TN D-6496, NASA, 1971. H-633.

[75] Ralph H. Smith. Notes on lateral-directional pilot-induced oscillations. TR 81-3090, AFWAL, March 1982.

[76] I. L. Ashkenas, R. H. Hoh, and G. L. Teper. Analyses of shuttle orbiter approach and landing. *J. of Guidance, Control, and Dynamics*, 6(6):448–455, 1983. doi:10.2514/3.8524.

[77] R. H. Smith. A theory for longitudinal short-period pilot-induced oscillations. TR 77-57, AFFDL, June 1977.

[78] Bruce G. Powers. An adaptive stick-gain to reduce pilot-induced oscillation tendencies. *J. of Guidance, Control, and Dynamics*, 5(2):138–142, 1982. doi:10.2514/3.56150.

[79] Donald T. Berry, Bruce G. Powers, Kenneth J. Szalai, and R. J. Wilson. In-flight evaluation of control system pure time delays. *Journal of Aircraft*, 19(4):318–323, 1982. doi:10.2514/3.57395.

[80] W. J. Norton. Aeroelastic pilot-in-the-loop oscillations. In *PIO Workshop following Active Control Technology: Applications and Lessons Learned*, Turin, Italy, May 1994. AGARD.

[81] Russian Journal. Aviation science and technology, 1991.

[82] Inadvertent in-flight slat deployment china eastern airlines flight 583 McDonnell Douglas MD-11, B-2171 950 nautical miles south of shemya, alaska april 6, 1993. Aircraft Accident Report PB93-910408, National Transportation Safety Board, October 1993. NTSB/AAR-93/07.

[83] Michael A. Dornheim. Report pinpoints factors leading to YF-22 crash. *Aviation Week and Space Technology*, pages 53–54, November 9 1992.

[84] E. R. Kendall. The design and development of flying qualities for the C-17 military transport airplane. In M. B. Tischler, editor, *Advances in Aircraft Flight Control*, chapter 7, pages 189–210. Taylor & Francis, Philadelphia, PA, 1996.

[85] David H. Klyde, Thomas T. Myers, Raymond E. Magdaleno, and James G. Reinsberg. Identification of the dominant ground handling characteristics of a Navy jet trainer. *J. of Guidance, Control, and Dynamics*, 25(3):546–552, 2002. doi:10.2514/2.4915.

[86] Henry R. Jex. Problems in modeling man-machine control behavior in biodynamic environments. In *Seventh Annual Conference on Manual Control*, pages 3–13, 1972. NASA SP-281.

[87] Bimal L. Aponso, Donald E. Johnston, Walter A. Johnson, and Raymond E. Magdaleno. Identification of higher order helicopter dynamics using linear modeling methods. *Journal of the American Helicopter Society*, 39(3):3–11, 1994. doi:10.4050/JAHS.39.3.

[88] Donald E. Johnson and Raymond E. Magdaleno. Independent assessment of C/MH-53E technical evaluation program (TEP). TR 1251-1 R, Systems Technology, Inc., September 1990.

[89] E. Buchacker. Experience with SIFT flight-test-techniques at the german air force flight test center. In *AGARD-CP-333, Paper 24*, June 1982.

[90] Anonymous. Über den Unfall des Hubschraubers Agusta Bell AB 204 B, HB-XKL. Technical Report 1980/19–1012, Eidgenössisches Verkehrs- und Energiewirtschaftsdepartement, 1980.

[91] Heinz-Jürgen Pausder and Chrisl Blanken. Investigation of the effects of bandwidth and time delay on helicopter roll-axis handling qualities. In *Piloting Vertical Flight Aircraft: A Conference on Flying Qualities and Human Factors*, pages 91–110, NASA Ames Research Center, 1993.

A Database of APC events

Table A.1: Database of APC events.

A/C type*	Acc. year	Acc. date	Aircraft model	Event description	APC type	Ref.**
F	1941	06/27	XB-19	longitudinal PIO during first takeoff and landing	PIO	
F	1947	10/24	XS-1	PIO during gliding approach and landing	PIO	[60]
F	1949	-	XF-89A	PIO during level off dive recovery	PIO	[12]
F	-	-	F-86D	PIO during formation flying- pulling G's	PIO	[12]
F	-	-	F-100	PIO during tight maneuvering	PIO	[12]
F	-	-	F-101	aft CG		[12]
F	1952	03/31	X-15		-	[61,62]
F	-	-	A4D-1 Skyhawk	low altitude; near sonic Mach - Cat. I: arm mass increases feel system inertia; leads via bobweight feedback to unstable coupling with short period dynamics if pilot merely hangs loosely onto the stick after large input	PIO	[63]
F	-	-	C-97 Stratofreighter	approach; landing - Cat. I: lag from radar-detected error to voice command led to unstable closed-loop phugoid mode; critical subsystems included display and vehicle airframe dynamics	PIO	[63]
F	-	-	Curtiss SB2C-1	cruise - Cat. II: porpoising (unable to maintain flight path), hysteresis in stick versus elevator deflection resulted in low frequency speed and climb oscillations	PIO	[63]
F	1957	01/19	A4D-2	high speed category III PIO, during routine flight testing	PIO	[64,65]
F	1958	-	KC-135A	mild lateral-directional PIO associated with w_ϕ/w_d effects	PIO	[66]
F	-	-	B-52	roll PIO while refueling	PIO	[12]
F	1958	-	F-101B	lateral PIO at high q subsonic	PIO	[67]
F	1959	06/08	X-15	Gliding flight approach, Cat. II PIO	PIO	[62,68]
F	-	-	XY2FY-1	post-take off destructive PIO	PIO	[12]
F	1960	01/26	T-38	high speed Cat. III PIO, distributed bobweight and primary control system involved	PIO	[63,68–72]
F	-	-	X-15 T-33VSA F-101B F-106A KC-135A B-58	cruise - Cat. I: zeros of roll/aileron transfer function are higher than dutch roll frequency, leading to closed-loop instability at conditions with low dutch roll damping; due to airframe dynamics	PIO	[63]
F	-	-	XF-10 F-101B F-102A	cruise - Cat. II: pitch up due to unstable kink in the alpha-moment curve; led to moderate period oscillations of varying amplitudes during manoeuvres at the critical angle of attack	PIO	[63]

Continued on next page

Table A.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Aircraft model	Event description	APC type	Ref.**
F	1960	-	B-58	lateral Dutch roll oscillations lead to lateral PIO due to actuator limiting of AFCS	PIO	[73]
F	1961	-	X-15	lateral PIO, w_ϕ/w_d research study	PIO	[12]
F	1961	05/18	F-4	low altitude record run second pass, destructive PIO	PIO	[12]
G	1962	-	Parasev	lateral rocking PIO during ground tow	PIO	[60]
F	1962	09/14	B-58	lateral-directional control-associated crash	PIO	[12]
F	-	-	T-38A	low altitude; near sonic Mach - Cat. I: arm mass increases feel system inertia; leads via bobweight feedback to unstable coupling with short period dynamics if pilot merely hangs loosely onto the stick after large input	PIO	[12]
F	1967	05/10	M2-F2	lifting body lateral-directional Cat. II PIO	PIO	[74,75]
S	1967	Oct.	Shuttle	ALT-5 lateral PIO, just prior to longitudinal PIO	PIO	[68,76]
F	1968	-	Transport A/C	PIO due to the insufficient stability margin + aeroelasticity	PIO/ PAO	Russia
F	-	-	Transport A/C	PIO during landing	PIO	Russia
F	-	-	Transport A/C	high-frequency PIO (3 Hz); interaction between pilot, wheel characteristics, aeroelasticity	PIO/ PAO	Russia
F	-	-	F-18	MTE (A2A Refueling) - PIO on A2A refueling exercises on early FCS version	PIO	[1]
F	-	-	Fighter A/C	PIO during pitch stabilization; interaction between pilot, manipulator, feel system characteristics	PIO/ PAO	Russia
F	1974	01/20	YF-16	on-ground (Taxi) - unplanned first flight during high-speed taxi test	PIO	[1,68,75]
F	-	-	YF-12	Cat. III PIO	PIO	[68,77]
F	1975	-	Tornado	short take off - heavy landing		[12]
F	-	-	YF-12	high frequency flexible mode involvement, Cat. I PIO	PIO	[68,77]
F	-	-	A-6	lateral effective bobweight effects		[12]
F	-	-	Fighter A/C	PIO during roll stabilization task (0.4 Hz); interaction between yaw and roll	PIO	Russia
F	1976	01/26	Tornado	landing - landing accident during flight test of prototype no. 5		[1]
S	1977	10/26	Shuttle	ALT-5 during landing approach, Cat. III PIO	PIO	[76,78]
F	1978	04/18	DFBW F-8	PIO during touch and go operation		[1,79]
F	-	-	Transport A/C	high-frequency roll ratchet (2.5 Hz); interaction between pilot, manipulator, aeroelasticity	PIO/ PAO	Russia
F	1978	-	Fighter A/C	PIO (0.8 Hz) due to the too high control sensitivity during flight tests	PIO	Russia

Continued on next page

Table A.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Aircraft model	Event description	APC type	Ref.**
F	-	-	Airbus A-320	several undocumented PIOs that reportedly occurred during development	PIO	[1]
F	-	-	F-5A	cruise - Cat. I: spiral mode driven unstable if roll information is degraded during gunnery; driven by display and vehicle	PIO	[63]
F	1986	-	Voyager	pilot coupling with symmetric wing bending	PIO	[80]
F	1986	-	Fighter A/C	PIO (0.8 Hz) due to the too high control sensitivity	PIO	Russia
F	-	-	F-86D, F-100C	low altitude; near sonic Mach - Cat. II: valve friction plus compliant cabling resulted in large oscillations at abort period	PIO	[63]
F	-	-	A-3D	low altitude; near sonic Mach; cruise - Cat. II: transonic snaking; separation over rudder causes control reversal for small deflections, leading to limit cycle if rudder used to damp yaw oscillations; due to airframe dynamics and feel system	PIO	[63]
F	1988	Dec.	JAS-39	approach - heavy turbulence; pilot had little experience with aircraft; pilot was controlling flight path angle and met insufficient control power, inducing rate limiting; resulted in crash and destruction of test aircraft	PIO	[12]
F	1990	-	JAS-39	PIO during approach Cat. II–III PIOs	PIO	[12]
F	-	-	F-111	pilot lateral control coupling with sustained under wing heavy store limit cycle oscillation	PIO	[80]
F	-	-	F-14	high angle of attack, with some sideslip angle		[12]
F	-	-	AD-1	oblique wing		[12]
F	1991	-	Tu-154M	roll PIO during landing due to high control sensitivity between pilot and lateral side-stick characteristics	PIO	[81]
F	1992	03/22	Antonov An-30	longitudinal and lateral PIO	PIO	Russia (CAA)
F	1993	04/06	MD-11	PIO triggered by inadvertent slat deployment	PIO	[82] NTSB: DCA93MA037, A-93-143/152
F	1992	04/25	YF-22	PIO after touch down and wave off in afterburner, Cat. III PIO	PIO	[83]
F	1993	-	JAS-39	MTE (low altitude flight) - APC event during low altitude flight demonstration		[1]
F	1994	-	B-2	approach; landing; A2A refueling		[1]
F	1995	04/28	Airbus Industries A320	lateral oscillation	PIO	NTSB: CHI95IA138

Continued on next page

Table A.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Aircraft model	Event description	APC type	Ref.**
F	1995	-	Boeing 777	several PIOs during development flight test; pitch oscillation at touchdown triggered by deployment of spoilers, pilot's use of a pulsing technique to control a 3 Hz bending mode, oscillations after take-off triggered by mistrimmed stabiliser	PIO	[1]
F	1996	02/01	Beech Baron	PIO during landing; jury awarded victim $3.95 mil for "negligence and pilot-induced oscillation in the landing phase"	PIO	[15]
F	1996	06/26	F-16 DBTC	display induced PIO during terrain following	PIO	[15]
F	1996	-	C-17	roll PIO during landing	PIO	[84]
F	1997	06/08	MD-11	longitudinal PIO; captain initially indicted, then cleared of negligence	PIO	[15,30]
F	1997	07/31	MD-11	longitudinal PIO during landing	PIO	J-AIB: JAL 706-1997
F	1998	-	Boeing 757	lateral oscillations during approach and landing	PIO/ PAO	[15]
F	1999	01/15	Boeing 767	buckled fuselage during landing		[15]
F	1999	01/15	Boeing 767-300	longitudinal PIO	PIO	AAIB: EW/C991301
F	1999	03/08	F-18F	FCS modifications made due to landing on carrier	PIO	[15]
F	1999	09/14	Falcon 900	during descent pilot attempted to level with autopilot engaged, resulting in PIO; pilots convicted of manslaughter		Romania (CAA): 1999091401, [15]
F	1999	10/09	Falcon-900	PIO during descent	PIO	[15]
F	2001	-	C-17A	numerous PIO events during approach & landing, some with damage to wings, flaps, engine nacelles, landing gear	PIO	[15]
F	2001	02/08	A321	PIO during landing, damaged wing tip, fence and ailerons	PIO	[15]
F	2001	03/17	A320	PIO during take off	PIO	NTSB: CHI01FA104
F	2001	06/07	Learjet 24	Lateral oscillation		NTSB: LAX01TA204
F	2001	07/01	X-35B JSF	Cat. I PIO during hover	PIO	[15]
F	2002	12/07	A321	PIO during landing	PIO	[15]
F	2002	-	T-45	directional PIO on landing rollout	PIO	[85]
F	2004	10/28	F/A-22	PIO during air-to-air tracking	PIO	[15]
F	2008	01/10	A319	lateral oscillation		EASA: A08W007
F	2008	07/31	Aero Commander 200D	inadvertent pilot-induced oscillation during landing flare/touchdown	PIO	NTSB: CHI08CA225
F	2008	09/27	Sport Flight Intl. Astra	yaw oscillations during landing		NTSB: SEA08CA212

Continued on next page

Table A.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Aircraft model	Event description	APC type	Ref.**
G	2009	01/16	Glaser-Dirks DG-400	pitch oscillation during cruise		NTSB: WPR09FA089
G	2009	06/29	Schempp-Hirth Ventus	released from tow plane and encountered a longitudinal PIO	PIO	NTSB: WPR09LA317
F	2009	08/31	Cessna 152	student pilot initiated a longitudinal PIO during flare	PIO	NTSB: WPR09CA430
F	2009	11/06	Cirrus SR20	the pilot's improper flare initiated a longitudinal PIO	PIO	NTSB: WPR10CA054

*F: Fixed Wing, S: Shuttle, G: Glider

B Database of RPC events

Table B.1: Database of RPC events.

A/C type*	Acc. year	Acc. date	Rotorcraft model	Event description	RPC type	Ref.**
H	1964	-	Bo-46	rotor control/gyro system coupling		[29]
H	1967	-	CH-46D	flexible 'shuffle mode' air resonance		[29]
H	1967	-	CH-46D Sea Knight	3.2 Hz 'shuffle' oscillation; out of phase coupling of rotors w/ aft pylon fuselage mode; changes made to the aircraft and operations	PAO	[29]
H	1968	-	CH-47	rotor/sling load bounce		[29]
H	1970	-	AH-56	flexible control actuation system		[86]
H	1978--1985	-	CH-53E	APC with flexible modes, several major instances in precision hover and with heavy sling loads, including heavy landings, dropped loads; extreme category I to category II PIOs	PIO	[87, 88]
H	1978	-	CH-53E (USN)	flexible modes/sling loads		[87]
H	1980	-	CH-53G (GAF)	flexible modes/sling loads	PAO	[89]
H	1980	-	CH-46E	flexible mode-air resonance 'shuffle mode'		[29]
H	1980	06/13	AB-204B	vertical oscillations following the sudden release of a slung load		[90]
H	1981	-	SH-60	flexible mode ground resonance		[29]
H	1988	-	UH-60 ADOCS	excessive time delays		[19]
T	1989	-	V-22	3.0 Hz roll mode; coupling with roll and main rotor system's regressive lag mode; PAO from large aft rotor flapping; procedural centering of control stick, reducing rotor flapping and increased rotor lead-lag damping	PAO	[51]
T	1990	-	V-22A Osprey (FSD)	3.2 Hz asymmetric wing chord mode due to aerodynamic phenomena; coupling with lateral cyclic inputs; addition of a notch filter at 3.2 Hz	PAO	[29]

Continued on next page

Table B.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Rotorcraft model	Event description	RPC type	Ref.**
T	1991	-	V-22A Osprey (FSD)	3.8 symmetric wing chord bending mode w/ 4000 lb load; pilot coupling through longitudinal cyclic; notch filters introduced at frequency	PAO	[29]
T	1991	-	V-22A Osprey (FSD)	4.2 Hz symmetric wing chord mode coupled with the pilot Thrust Control Lever (TCL, commanding rotor collective); minor coupling at 5.3 Hz with symmetric wing torsion mode; asymmetric notch filters added	PAO	[29]
H	1992	-	S-76B	flight control mode shifting	PIO	[29]
H	1993	-	BO105 ATTHeS	time delay/attitude command		[91]
H	1994	06/02	47D-1	pilot inducted lateral oscillation due to heavy cyclic control forces in hover	PIO	NTSB: LAX94LA235
H	1995	-	BO105 ATTHeS	biomechanical/airframe coupling	PAO	[16, 26]
T	1997	-	V-22B Osprey (EMD)	1.4 Hz high focal roll mode oscillation due to change in mass balance weight; relaxation of pilot grip on cyclic	PAO	[29]
H	1998	12/03	EC135-P1	helicopter encountered wake turbulence of a MD-80 airplane and PIOs occurred during recovery	PIO	NTSB: NYC99FA032
T	1999	02/01	V-22 Osprey	hover over ship; caused suspension of sea trials	PAO	[15]
H	2000	08/08	OH-58C	PIO during a practice autorotation	PIO	NTSB: ATL00TA080
H	2000	12/18	SA365-N1	longitudinal and lateral PIO during landing		NTSB: NYC01LA059
G/C	2003	01/01	Air Command Commander Elite	inadvertent phugoid PIO due to wind gust	PIO	NTSB: CHI03LA048
G/C	2003	04/23	Denzer RAF 2000	abrupt lift-off caused longitudinal PIO during take off		NTSB: ANC02FA064
H	2003	06/28	Schweizer 269C	lateral oscillation		NTSB: DEN03LA115
G/C	2003	11/16	Northam RAF 2000	longitudinal oscillations during level flight		NTSB: NYC04LA035
H	2004	05/08	Robinson R44	longitudinal PIO due to experiencing low cyclic force while initiating a hover after take off	PIO	AAIB: G-CBXX
H	2005	08/13	Robinson R44	the inadequate remedial action during landing by the pilot caused pitch oscillations	PIO	NTSB: CHI05LA235
H	2006	01/10	AS350BA	yaw initiated PIO caused helicopter to crash	PAO/ PIO	NTSB: LAX06LA072
H	2006	10/16	Robinson R22 Beta	PIO in yaw axis started during cruise flight		NTSB: DEN07CA013
H	2007	12/05	UH-1B	pilot caused vertical oscillations due to collective bounce	PAO/ PIO	NTSB: SEA08LA043
H	2008	05/01	Robinson R22 Beta II	student pilot started a lateral PIO in hover		NTSB: LAX08CA126

Continued on next page

Table B.1 — continued from previous page

A/C type*	Acc. year	Acc. date	Rotorcraft model	Event description	RPC type	Ref.**
H	2008	06/29	UH-1B	collective bounce leads to vertical oscillations during autorotation	PAO/ PIO	NTSB: ANC08LA083
H	2009	05/12	Robinson R44	initiated yaw oscillations turned into yaw-pitch PIO		NTSB: ANC09GA040
H	2009	11/15	Robinson R44 Astro	inexperienced pilot caused mixed PIO		AAIB: G-WEMS

*H: Helicopter, G/C: GyroCopter, T: Tiltrotor
**NTSB: National Transportation Safety Board, AAIB: Air Accidents Investigation Branch

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Power Absorption Characterisation of Electro-Mechanical Flight Control Actuators via Detailed System Modelling

E. Denti, G. Di Rito, R. Galatolo and F. Schettini
Dipartimento di Ingegneria Aerospaziale – Università di Pisa, Italy

Keywords: *flight actuators, electro-mechanical systems, power analysis, modelling and simulation*

Abstract

The paper deals with the modelling of electro-mechanical flight actuators for power absorption studies. The recent trend to the electrification of all aircraft systems actually arises challenging issues about the management of the on-board power, with a strong need of accurate systems' models. One of the key engineering concerns is to define the modelling detail that is adequate for power characterisation. In this paper, two models of an electro-mechanical actuator for flight controls are developed: a detailed one, created in the AMESim environment, that includes the simulation of all system parts (3-phase brushless motor, mechanical transmission, PWM power electronics and control electronics), and a simplified Simulink model. The transient and the steady-state actuator power absorptions during a sample flight manoeuvre are evaluated, highlighting and discussing the importance of a detailed modelling for a complete characterisation of the electrical energy requests.

1 Introduction

The conversion of on-board systems to the "all-electric" concept is nowadays one of the most relevant challenges in aircraft design [1] [2] [3] [4]. As for any technological innovation, this clearly arises new issues related to performances and reliability, but also implies specific concerns about systems interactions in terms of electrical power management [5]. The complete electrification of all aircraft power systems actually entails the implementation of smart logics for sharing the available energy among the loads, and the design of these logics requires the characterisation of the power absorption of each on-board system as a function of mission phase and aircraft operating point [6]. In particular, a specific attention must be dedicated to safety-critical systems (e.g. the flight control actuators), whose power requests have to be fulfilled continuously as well as completely [7] [8] [9]. Modelling and simulation activities play a key role in this design loop: firstly, because the power management can be tested in operating conditions that are rarely encountered during flight or that are difficult to be reproduced with on-ground rigs, pointing out the design criticalities; secondly, because the electro-

mechanical actuation itself is a novel technology for flight control applications, and the influence on performances of electrical power quality, actuator thermal behaviour and power electronics efficiency needs to be investigated. A strong effort is thus required for system engineers to develop models that are capable of predicting the dynamic performances and the power absorption of electro-mechanical flight controls, with a level of modelling detail that is adequate for the objectives [10].

In this paper, two models of an electro-mechanical actuator (EMA, Fig. 1) are developed and compared in terms of power characterisation capability: a detailed model that includes the simulation of all system parts (3-phase brushless motor, mechanical transmission, PWM power electronics and control electronics); and a simplified model, in which the power electronics is not simulated, the brushless motor is described with reference in the *dqo* frame [11], and the mechanical transmission is assumed to be completely rigid. The work is organised into two parts. Firstly, both EMA models are described in terms of structure and basic working equations; then, once defined the model parameters with reference to the primary control surface of a light military jet trainer, the transient and the steady-state EMA power absorptions during a sample flight manoeuvre are evaluated with both models, highlighting and discussing the importance of the modelling detail for the description of the actuator energy consumption characteristics.

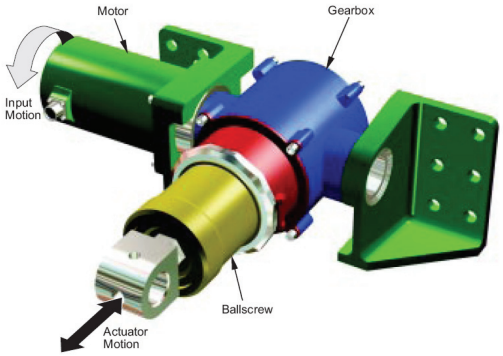


Fig. 1 – Electro-mechanical actuator.

2 EMA system architecture

The EMA system modelled in the work is referred to a primary flight control application [8] [12] [13], and it is basically composed of

- a 3-phase synchronous brushless motor;
- a mechanical transmission with a single-stage gearbox and a low-pitch screw jack (actuating the aerodynamic surface via lever arm);
- a 270 V_{dc} PWM power electronics, equipped with 12 MOSFET switches for the isolation of the motor phases[1], Fig. 2;
- a control electronics implementing three closed-loops: on actuator position, motor shaft speed and motor currents.

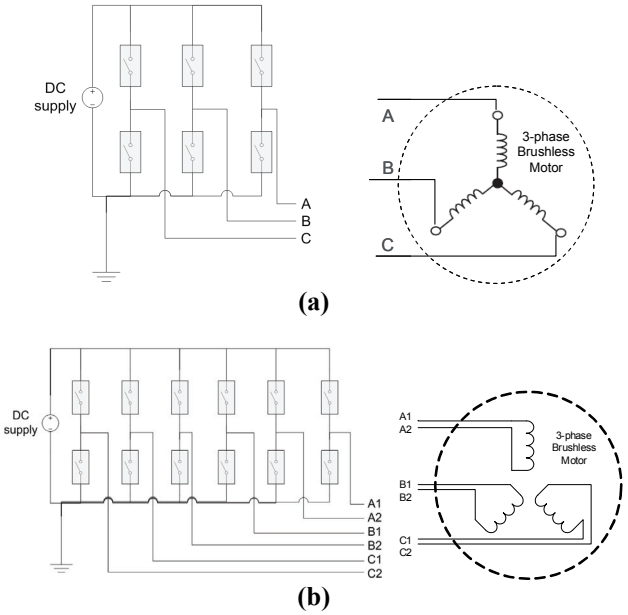


Fig. 2 – Motor electronics solutions: (a) traditional, with 6-switches; (b) phase-isolating, with 12 switches.

Though different EMA solutions can be identified, the proposed architecture allows to point out some relevant specific concerns related to the electro-mechanical actuation of

[1] In this arrangement, the three phases of the motor are not connected via a neutral point, but each phase is linked to a full H-bridge. In case of a motor phase failure, it can be isolated by opening the related power switches, and the remaining two phases can still work.

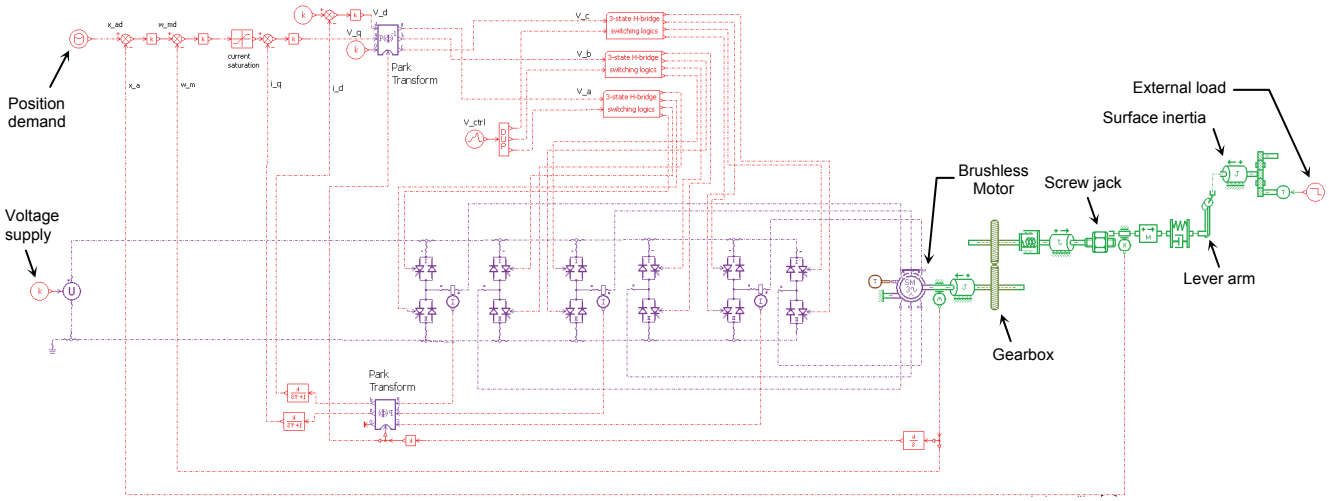


Fig. 3 – Detailed model of the electro-mechanical actuation system (AMESim).

the primary flight controls, such as high power density (brushless motor), fast dynamic response (motor speed control), high efficiency (screw jack), adequate reliability (fault-tolerant power electronics).

3 EMA models

3.1 Detailed model

The detailed model of the EMA system has been developed in the LMS-AMESim environment, Fig. 3. This choice has been essentially driven by the object-oriented approach of the tool, which allows the engineer/analyst to rapidly and efficiently prototype complex multi-physical systems, by essentially reproducing their architecture and the components topology.

As shown in Fig. 3, the model includes the simulation of the whole EMA power electronics, which implements a three-state PWM technique for the phase current drive[2] [14]. It is worth noting that the simulation of a phase-isolating electronics with 12 switches (all having the same characteristics, i.e. $R_s^{(on)}$ and $R_s^{(off)}$, if activated or not, Table 1) led to the necessity of creating a new component model for 3-phase brushless motors (the software library only provides models with neutral connections, i.e. *delta* or *star* type), in order to obtain as outputs all the six terminals of the three phases.

The detailed model also takes into account the compliance of the mechanical transmission elements as well as the friction on the actuator moving parts (motor shaft, gears, screw, nut and the aerodynamic surface itself). Though more precise information would be necessary to evaluate the importance of these two aspects on the EMA dynamics (e.g. the stiffness and damping data of the gear-screw, screw/nut and nut/surface connections), the physical phenomena have been included in the detailed EMA model for obtaining qualitative and "trend" indications.

3.2 Simplified model

The simplified model of the EMA system, created in the Matlab-Simulink environment, has been developed under the following assumptions:

- the 3-phase brushless motor dynamics is described in the *dqo* frame (no Park transforms are used) neglecting the homopolar axis
- the power electronics is *ideal* (no voltage drop on a closed switch, no current

[2] The voltage to each phase is provided as a pulse-wave of pre-defined period (T_{pwm}, Table 1) with three possible amplitudes, obtained by switching the DC voltage supply.

through an opened switch, perfect voltage tracking)
- the mechanical transmission is rigid

The electric equations along the direct and quadrant axes of the motor are thus given by Eqs. (1)-(2),

$$V_d = R\,i_d + L_d \frac{di_d}{dt} - L_q\,p_m\,i_q\,\omega_m \tag{1}$$

$$V_q = R\,i_q + L_q \frac{di_q}{dt} + L_d\,p_m\,i_d\,\omega_m + K_b\,\omega_m \tag{2}$$

where R is the motor phase resistance, p_m is the number of magnet pole pairs, ω_m is the motor shaft speed, K_b is the motor back-electromotive force coefficient, while V_d, V_q, L_d, L_q, i_d and i_q are the voltages, the inductances and the currents along the direct and the quadrant axes respectively.

The motor torque is then provided by Eq. (3),

$$T_m = [(L_d - L_q)\,p_m\,i_d + K_b]\,i_q \tag{3}$$

and the system dynamics, once assumed the mechanical transmission perfectly rigid, can be described with a single momentum equation, referred to the motor shaft, Eq. (4),

$$J_{tot}\,\dot{\omega}_m = T_m - T_{ext} - B_{tot}\,\omega_m - T_{f\,tot}\,\mathrm{sgn}(\omega_m) \tag{4}$$

where

$$J_{tot} = J_m + \frac{J_g}{\tau_g^{\,2}} + \frac{J_s}{\tau_g^{\,2}}\left(\frac{p_s}{2\pi\,l_s}\right)^2 \tag{5}$$

$$T_{ext} = \frac{p_s}{2\pi\,l_s\,\tau_g}\,H_s \tag{6}$$

In Eq. (4), J_{tot} is the system inertia, B_{tot} is overall viscous damping coefficient, $T_{f\,tot}$ accounts for sliding friction effects (Coulomb, Stribeck, etc.), T_{ext} is the external loading torque scaled to the motor shaft; while in Eqs. (5)-(6), p_s is the screw pitch, l_s is surface lever arm, τ_g is the gearbox ratio, H_s is surface hinge moment, and J_m , J_g and J_s are the inertia of the motor shaft, gears and aerodynamic surface, respectively.

Concerning the actuator control, all the four loops (position, motor speed, direct and quadrant currents) are assumed to be simply proportional, Eqs. (7)-(10),

$$V_d = -K_i\,i_d \tag{7}$$

$$V_q = K_i\,(i_{qi} - i_q) \tag{8}$$

$$i_{qi} = K_\omega\,(\omega_{mi} - \omega_m) \tag{9}$$

$$\omega_{mi} = K_x\,(x_{ai} - x_a) \tag{10}$$

in which K_i , K_ω , K_x are the control gains of the current, speed and position loops respectively, and x_a is the actuator position, Eq. (11).

$$\dot{x}_a = \frac{p_s}{2\pi\,\tau_g}\,\omega_m \tag{11}$$

In addition, both voltage and current demands are limited, by saturating the related values to V_{max} and i_{max} respectively.

The power balance is finally obtained by multiplying Eq. (1) and Eq. (2) by i_d and i_q respectively and by summing them, Eq. (12).

$$P_e^{(in)} = P_e^{(s)} + P_e^{(l)} + P_m^{(s)} + P_m^{(l)} + P_m^{(out)} \tag{12}$$

In Eq. (12), $P_e^{(in)}$ is the electrical power input, $P_e^{(s)}$ is the electrical power stored in the system (only related to transients since no capacitors are considered), $P_e^{(l)}$ is the electrical power lost in the circuitry, $P_m^{(s)}$ and $P_m^{(l)}$ are the mechanical powers stored and lost respectively, while $P_m^{(out)}$ is the mechanical power output, Eqs. (13)-(18).

$$P_e^{(in)} = V_d\,i_d + V_q\,i_q \tag{13}$$

$$P_e^{(s)} = L_d \frac{di_d}{dt}\,i_d + L_q \frac{di_q}{dt}\,i_q \tag{14}$$

$$P_e^{(l)} = R\,(i_d^{\,2} + i_q^{\,2}) \tag{15}$$

$$P_m^{(s)} = J_{tot}\,\dot{\omega}_m\,\omega_m \tag{16}$$

$$P_m^{(l)} = B_{tot}\omega_m^{\,2} + T_{f\,tot}\,\mathrm{sgn}(\omega_m)\,\omega_m \tag{17}$$

$$P_m^{(out)} = T_{ext}\,\omega_m \tag{18}$$

3.3 Data and simulation parameters

The basic data of the modelled EMA system are given in Table 1, where m_n is the nut mass, and K_g, B_g, K_s, C_s, K_l and C_l are the stiffness and damping characteristics of the gear/screw, screw/nut and nut/surface connections used in the detailed model.

Table 1 – EMA system data

Parameter	Value	Unit
R	0.74	Ohm
p_m	1	-
K_b	0.5263	V sec / rad
$L_d = L_q$	4.8e-3	H
$J_m = J_g$	8e-4	kg m^2
J_s	0.2	kg m^2
p_s	5	mm
l_s	60	mm
τ_g	4	-
B_{tot}	3.8e-3	N m sec / rad
$T_{f\,tot}$	0.0119	N m
K_i	9	V /A
K_ω	0.3183	A sec / rad
K_x	2.618e5	rad / (sec m)
V_{max}	270	V
i_{max}	30	A
T_{pwm}	1e-4	sec
$R_s^{(on)}$	1e-4	Ohm
$R_s^{(off)}$	1e4	Ohm
K_g	5.73e4	N m / rad
B_g	950	N m sec / rad
$K_s = K_l$	1e9	N / m
$C_s = C_l$	1e5	N sec / m
m_n	2	kg

Concerning the simulation parameters, the two models are characterised by different numerical integration methods: a variable-step integrator[3] has been necessary for simulating the detailed model, while a fixed-step method (Runge-Kutta with 5e-5 sec time step) has been set for the simplified one.

The models exhibit very different computational performances. As an example, on a PC equipped with an Intel Quad Core Q9550 2.83GHz (3.25GB RAM), by recording 10 system states (output torque, output speed, direct current, electrical power input, commands and feedbacks of position, motor speed and quadrant current) at 5e-5 sampling time, the simplified model simulates 10 seconds in 10.8 seconds, while the detailed model takes about 1.5 hours (the performance depends on the frequency contents of the input signals).

These data must be carefully taken into account when comparing the models (section 4), since a dramatic increase of the computational cost derives from using the detailed model.

4 Simulation test

The two EMA models have been tested with reference to the same time history of elevator position command, as obtained by the flight simulator of a light military trainer aircraft. Starting from an initial condition of horizontal trim flight (Table 2), the aircraft is supposed to experience a pull-up manoeuvre at 0.5 seconds.

Table 2 – Initial flight conditions (horizontal trim)

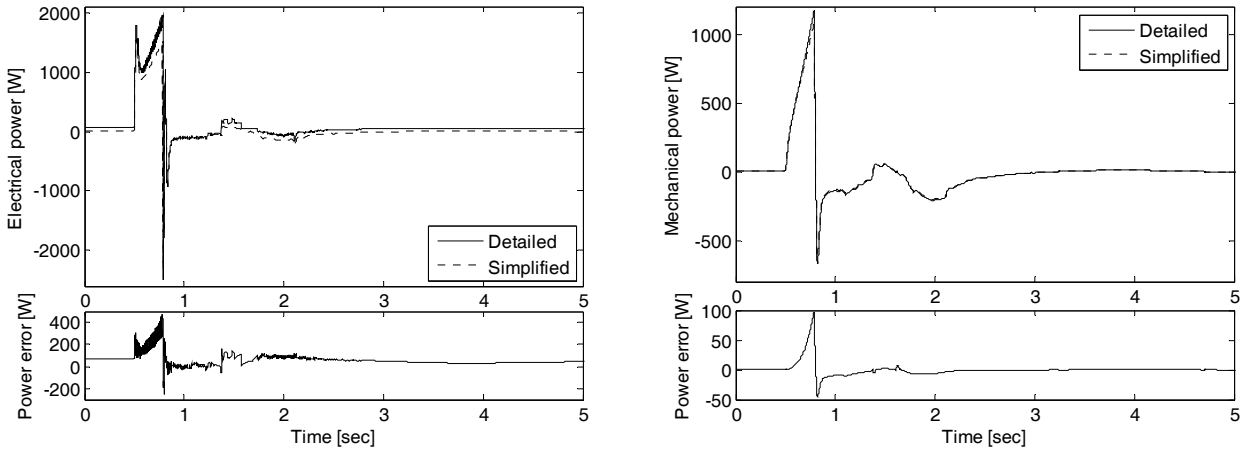
Parameter	Value	Unit
Altitude	0	m
Mach number	0.3	-
Angle of attack	0.62	deg
Elevator deflection	-4.85	deg

The hinge moment loading the EMA system during the flight manoeuvre has been estimated via Eq. (19),

$$H_s = \frac{1}{2}\rho V_\infty^{\,2} S_s c_s C_{H\delta_e}\delta_e \qquad (19)$$

where δ_e is the elevator deflection, $C_{H\delta e}$ is the hinge moment coefficient derivative, ρ is the air density, V_∞ is the airspeed, while S_s and c_s are the surface and the chord of the flapped portion of the horizontal tail respectively. Equation (19) does not take into account the dependence

[3] The AMESim software, when used with a standard setting, automatically selects the most appropriate integrator among a set of available multi-step methods, e.g. Adams-Moulton and backward differentiation type.



(a) (b)

Fig. 4 – EMA power flows: (a) electrical power input; (b) mechanical power output.

of the control hinge moment on the aircraft angle-of-attack. This approximation has been considered acceptable for the purposes of the work: firstly, because a co-simulation running the flight simulator and the EMA model would have been necessary for taking into account the angle-of-attack effects; secondly, because the angle-of-attack influence on the control hinge moment is typically minor [15].

The simulation results of the two models in terms of power flow characteristics and position response are shown in Fig. 4 and Fig. 5. It can be noted that the simplified model essentially matches the position dynamics provided by the detailed model (the maxima errors are lower than 0.02 mm), so it is surely convenient to use the simplified model when the simulation objective is the EMA control design.

On the other hand, significant differences between the models' predictions can be observed on the power flow signals. The error on the mechanical power output, being related to an essentially equivalent position tracking, is limited (lower than 10%), but the errors on the electrical power input are relevant during both steady-state and transient conditions. Before the pull-up command, the electrical power absorption predicted by the detailed model is 71 W, while the data provided by the simplified model is only 1.7 W. During the transient phase (at the actuator velocity inversion), the electrical power input error is more than 20%, Fig. 4. In addition, Fig. 6 shows the efficiency (defined as

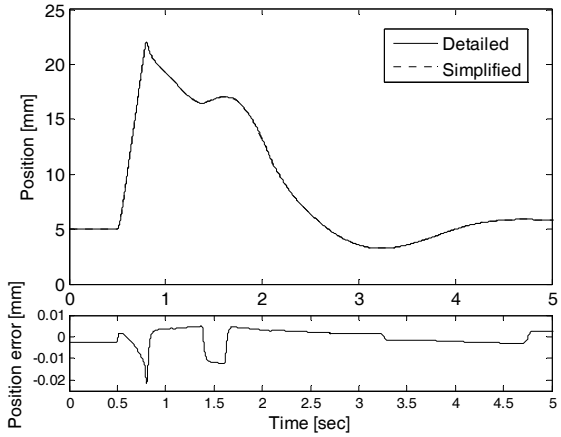


Fig. 5 - EMA position response.

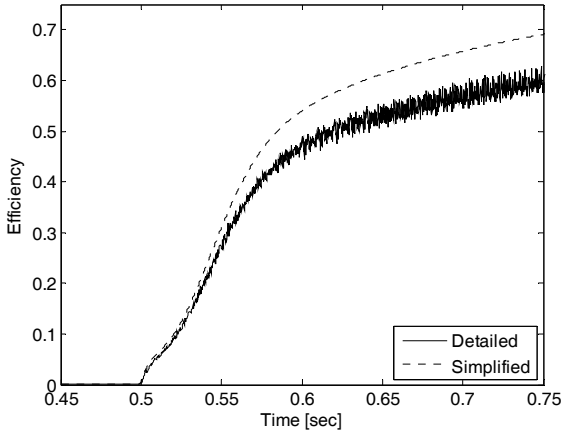


Fig. 6 - EMA efficiency (during the transient phase).

the ratio between the mechanical power output and the electrical power input) during the transient phase: it can be noted that the simplified model overestimates the efficiency up to about 10%.

Simulation results clearly point out that, if the simulation objective is the EMA power absorption characterisation, it is necessary to model the power electronics, since it significantly impacts on the electrical power flow. This can also be highlighted by means of Fig. 7, where the two models are compared in terms of quadrant current: during the steady-state condition, there is a very small current error (the applied torque is essentially the same), while the electrical power input differs of about 70 W. This energy totally derives from the Joule losses in the MOSFET switches of the EMA power electronics, which are completely neglected in the simplified model.

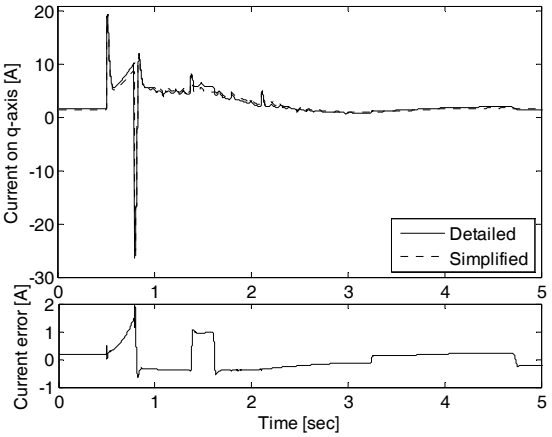


Fig. 7 – Motor current (q-axis).

5 Conclusion

Two models of an electro-mechanical flight control for a light military jet trainer, characterised by different levels of complexity, have been developed. The simulation studies highlight that a simplified model, disregarding the actuator power electronics, can be appropriate for control design purposes. However, if the objective is the power absorption characterisation, it is strictly necessary to model the power electronics, since it significantly impacts on electrical power flows. Nevertheless, the computational cost for simulating the power electronics is relevant, so that numerical optimisation techniques and/or model reductions should be applied in order to increase the model practicality.

References

[1] Moir I., "More-Electric Aircraft - System considerations", *IEE Colloquium on Electrical Machines and Systems for the More Electric Aircraft*, London, November 1999, pp. 10/1-9.

[2] Howse M., "All Electric Aircraft", *IEE Power Engineering Journal*, Vol. 17, N. 4, 2003, pp. 35-37.

[3] Emadi A. and Ehsani M., "Aircraft power systems: technology, state of the art, and future trends", *IEEE Aerospace and Electronic Systems Magazine*, January 2000, pp. 28-32.

[4] Rosero J. A., Ortega J. A., Aldabas E. and Romeral L., "Moving towards a More Electric Aircraft", *IEEE Aerospace and Electronic Systems Magazine*, March 2007, pp. 3-9.

[5] Maldonado M. A. and Korba G. J., "Power management and distribution system for a More Electric Aircraft (MADMEL) - Program status", *IEEE Aerospace and Electronic Systems Magazine*, December 1999, pp. 3-8.

[6] Anastasio V. and Di Donna L.. "Power by wire meets environmental friendly requirements". *Proceedings of the XIX AIDAA National Congress*, Forlì (Italy), 2007.

[7] Weimer J. A., "The role of electric machines and drives in the More Electric Aircraft", *IEEE Conference on Electric Machines and Drives*, Madison (USA), June 2003, Vol. 1, pp. 11-15.

[8] Botten S. L., Whitley C. R., and King A. D., "Flight control actuation technology for next-generation All-Electric Aircraft", *Technology Review Journal*, Fall/Winter 2000, pp. 55-68.

[9] Trentin A., Zanchetta P., Wheeler P. and Clare J., "Power flow analysis in electro-mechanical actuators for civil aircraft", *IET Electric Power Applications*, 2011, Vol. 5, Issue 1, pp. 48-58.

[10] Kuhn M. R., Otter M. and Raulin L., "A multi level approach for aircraft electrical systems design", *Proceedings of the 6th International Modelica Conference*, Bielefeld (Germany), March 2008, pp. 95-102.

[11] Fitzgerald A. E., Kingsley C. Jr and Umans S. D., *Electric Machinery*, 6th edition, McGraw Hill, 2003.

[12] Garcia A., Cusidò I., Rosero J.A., Ortega J.A. and Romeral L., "Reliable electro-mechanical actuators in aircraft", *IEEE Aerospace and Electronic Systems Magazine*, 2008, pp. 19-25.

[13] Davis M. A., "High-performance electro-mechanical servoactuation using brushless DC motors", *Motor-Con Conference*, Atlantic City (USA), April 1984.

[14] Krishnan R., *Electric motor drives – Modelling, analysis, and control*, Prentice Hall, Upper Saddle River, New Jersey, 2001.

[15] Hoerner S. F. and Borst H. V., *Fluid-dynamic lift: practical information on aerodynamic and hydrodynamic lift*, Borst-Hoerner, 1975.

Improving Flight Safety by Pushing the Simulation Envelope: The European Project SUPRA

Eric Groen
TNO, the Netherlands

Lars Fucke
Boeing Research & Technology Europe, Spain

Mikhail Goman
De Montfort University, United Kingdom

Vladimir Biryukov
Gromov Research Flight Institute, Russia

Hafid Smaïli
National Aerospace Laboratory, the Netherlands

Abstract

The European research project SUPRA (Simulation of Upset Recovery in Aviation) develops break-through simulator technologies to improve training of upset recovery techniques. Current flight simulators are considered inadequate for the simulation of many upset conditions as the aerodynamic models apply to the normal flight envelope, whereas aircraft behavior may change significantly outside this envelope. Furthermore, standard hexapod-based motion systems are unable to reproduce the high accelerations, angular rates, and sustained G-forces inherent to upset conditions. SUPRA investigates 1) new engineering methods to extend the aerodynamic model for high angles of attack; and 2) innovative motion cueing solutions, including centrifuge-based simulation. These new technologies will be evaluated by test pilots for a selection of relevant upset scenarios. The project's overall objective is to enable crews to detect and recover from upsets that could lead to Loss of control – in flight (LOC-I) situations, todays major cause of fatal accidents in commercial aviation.

1 Introduction

For several years now, Loss of control - in flight (LOC-I) continues to be the leading cause of fatal accidents in commercial aviation today[1]. Many LOC-I accidents have been attributed to a lack of the crew's awareness and experience in extreme flight conditions. In the course of loss of control events, the aircraft often enters unusual attitudes or other types of upsets. To prevent or timely exit a loss of control situation it is essential that the pilots rapidly recognize the condition, initiate recovery action and follow appropriate recovery procedures. Inadequate recovery may exacerbate the situation and lead to loss of the aircraft.

In-flight upsets are infrequent events in today's operations and many commercial pilots have never experienced such a situation, neither on part 25 certified (large transport category) aircraft nor during training on smaller airplanes or in military aircraft. This fortunate fact has unfavorable implications for the proficiency of aircrews in dealing with such events and calls for specific upset recovery training. Aviation authorities recognize the need to educate pilots on upset recovery techniques. In-flight training

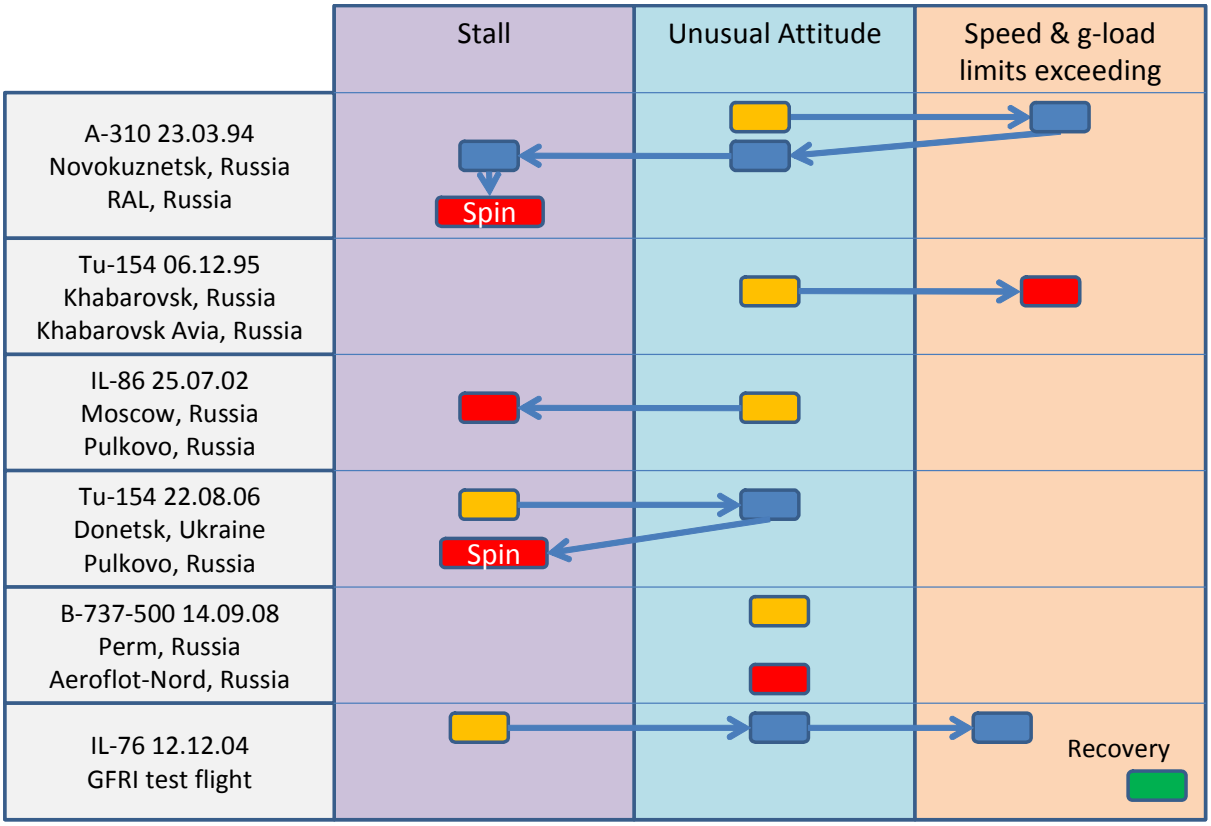


Fig. 1 Example of LOC-I events passing through various stages of upset.

with large aircraft is expensive and unsafe. Therefore, it is generally agreed that a ground-based flight simulator capable of accurately representing extreme flight conditions would significantly improve the effectiveness of upset recovery training programs. Since commercial pilots already receive a large part of their training in flight simulators, this would also be a cost-effective solution.

However, current flight simulators are considered inadequate for the simulation of many upset conditions as the aerodynamic models merely apply to the normal flight envelope. However, upset events can take the aircraft outside the normal envelope where aircraft behavior may change significantly, and pilots may have to adopt unconventional control strategies[2]. Furthermore, standard hexapod-based motion systems are unable to reproduce the high accelerations, angular rates, and sustained G-forces occurring during upsets and recovery from upsets. The European Seventh Framework Program project SUPRA – Simulation of Upset Recovery in Aviation – aims to push both the aerodynamic and the motion envelope of ground-based flight simulators and investigate the feasibility of conducting advanced upset recovery simulation. The research not only involves hexapod-type flight simulators but also experimental centrifuge-based simulators.

2 SUPRA upset scenarios

Consistent with general understanding an upset is defined as an airplane in flight unintentionally exceeding the parameters normally experienced in line operations or training[3]:

- Unusual attitudes (pitch attitude greater than 25 degrees nose up, or greater than 10 degrees nose down; bank angles > 45 degrees);
- Stall;
- Spin;

- Exceeding Mach or G-load limits.

Upset situations are highly dynamic. As shown in Fig. 1 for a selection of LOC-I accidents and a test flight performed by the Gromov Flight Research Institute (GFRI, Zhukovsky, Russia), flight parameters may change rapidly during a single upset event. Detection, understanding and initiation of proper recovery actions are challenging tasks for commercial flight crews; such is analysis and simulation of those scenarios. Breaking the scenarios down into categories of upsets makes them amenable to engineering analysis and enables structured, scenario-based training with the goal to provide flight crews with a readily available category-specific set of recovery actions whenever one of the upset types is detected and rapid crew action is required.

Based on analysis of the publicly available LOC-I accident and incident data, e.g.[4], and discussions with the SUPRA expert advisory group, the SUPRA consortium has identified a number of scenarios that will be used for the final validation. These scenarios have been grouped into three types: Unusual Attitudes, Approach-to-Stall and Stall. Recovery from developed spins was considered beyond the scope of commercial pilot training. The existing industry Upset Recovery Training Aid (URTA) defines a number of Unusual Attitude scenarios that will be used as the basis for the SUPRA Upset Scenarios[3]. Since the URTA is widely accepted by the industry and regulatory authorities it is regarded as an ideal starting point. The expert advisory group suggested to validate the transfer of the URTA scenarios to SUPRA scenarios that extend beyond the currently validated envelope using the extended aerodynamic models and enhanced simulator motion developed under the SUPRA project.

3 Aerodynamic modeling

The mathematical aerodynamic and flight dynamic model, together designated the “aircraft model”, are the heart of modern flight simulators. They are derived using a variety of engineering methods including wind tunnel and in-flight measurements as well as computational fluid dynamics and system identification methods. For Level D certified Full Flight Simulators (FFS) the model output accurately matches aircraft responses measured in-flight. However, this Proof-of-Match is only performed for conditions within the normal flight envelope. Much of the aerodynamic data outside that envelope, although some of it is available through wind-tunnel and flight testing, is currently not integrated into simulator data packages. In other words, simulated aircraft behavior is currently only valid and reliable within the boundaries of the normal envelope. Analysis of LOC-I accident data however shows that transport aircraft can exceed the boundaries of the normal envelope in the course of an upset event (Fig. 2).

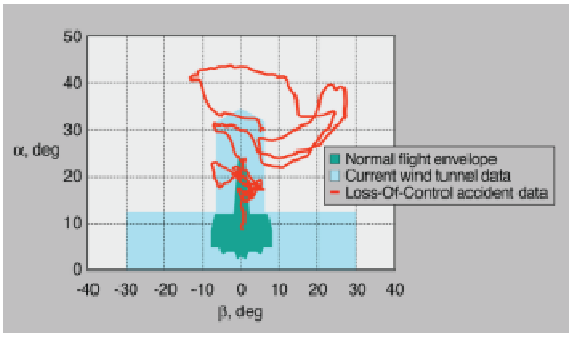


Fig. 2 LOC-I accident data, typical flight envelope and wind tunnel data[5].

The extension of the flight envelope required for the simulation of advanced upset conditions is depicted in Fig. 3 in terms of angle of attack (α), rotation rate (ω), and Mach number (M) typical for take-off/landing and cruise flight. Wind tunnel and flight test data on Mach number are available in the yellow region for a limited range of angles of attack. Although this includes stall conditions, this is not sufficient for simulation of aircraft departures, post-stall gyration and incipient spin modes. Beyond-stall conditions at high angles of attack special wind tunnel tests are conducted only at low speed using static and rotary balance experimental rigs (blue region). At moderate and cruise Mach numbers, stall conditions can occur in the flight envelope region where aerodynamic loads are below structural limits (see the question mark in

Fig. 3) and where wind tunnel data is not available. The aerodynamic modeling within SUPRA is targeted to fill this area.

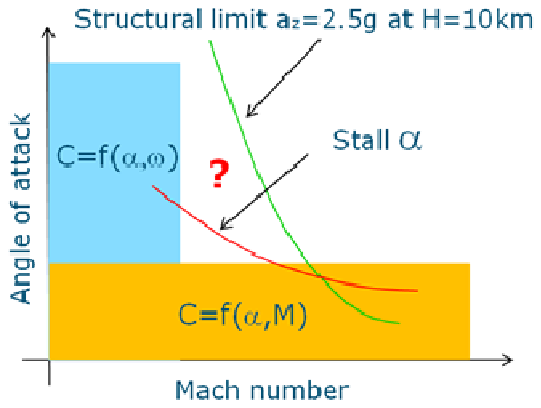


Fig. 3 Illustration of available aerodynamic data in extended flight envelopes of α and M.

As control and stability characteristics can change significantly when leaving the mostly linear aerodynamic envelope, it is important to develop simulation models that are capable of reproducing this behavior in order not to misinform the simulator pilot. A number of important aerodynamic effects should be taken into account at stall and post-stall flight conditions, for example:

- The lift force reaches a maximum value at some angle of attack. Further increase of angle of attack causes flow separation and loss of lift (stall).
- Dynamic aerodynamic hysteresis in the stall region, resulting from time delays in development of flow separation, may lead to dynamic instability in pitch, roll, and yaw.
- For a number of airplane configurations, for example T-tail, significant pitching up moments are generated due to the wing tail interaction leading to the onset of a deep stall.
- Asymmetrical development of flow separation leads to onset of destabilizing aerodynamic rolling/yawing moments and steady autorotation.
- Unsteady aerodynamic coupling between the wing and tail can generate aerodynamic negative damping effects leading to onset of bucking, wing-rock, etc.

3.1 SUPRA engineering methods

Within SUPRA, experts from De Montfort University use system identification methods to capture the unsteady aerodynamic phenomena, representing the major effects by ordinary differential equations[6][7]. This so-called "phenomenological model" for high angles of attack is superimposed on the standard mathematical model, preserving the traditional aerodynamic dependencies in the normal flight regime and corresponding low angles of attack.

Previously, phenomenological models for unsteady aerodynamic dependencies at stalled flow conditions were identified using experimental wind tunnel data obtained in forced oscillation tests. Under SUPRA it is being investigated whether it is possible to produce the non-linear aerodynamic responses required for the extended aircraft model using Computational Fluid Dynamics (CFD) in a cost-effective way. The Dutch National Aerospace Laboratory (NLR) has developed advanced CFD methods and previously validated these methods extensively against wind tunnel data for a wide range of extreme flight conditions of military aircraft; covered phenomena include shock waves, flow separation, and vortices. CFD predictions provided a very good estimate of the conditions encountered in real flight [6][8][9][10]. In SUPRA, CFD predictions are made at harmonic variations of angle of attack, sideslip and roll rates for different amplitudes and frequencies.

The SUPRA engineering approach is illustrated in Fig. 3 for the non-linear variation of the lift coefficient with angle of attack. Other than in static conditions, there is a significant dynamic hysteresis effect when the angle of attack varies periodically in the stalled region (green loop). This typical aerodynamic response can be captured using a phenomenological model based on a simple differential equation. There is a close match between the model and CFD predictions (red loop) [11].

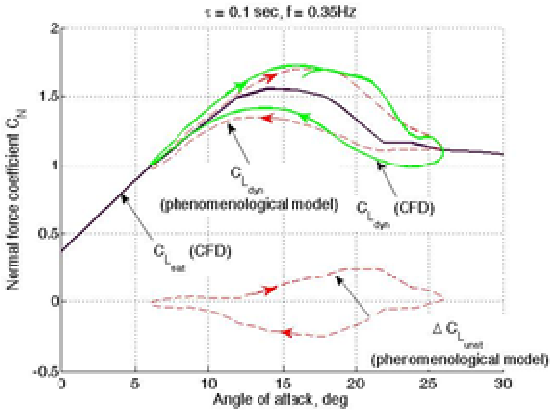


Fig. 4 Comparison of CFD and phenomenological predictions of dynamic hysteresis of the lift coefficient as function of angle of attack. The solid black line represents static conditions.

3.2 Reconfigurable model

An important feature of the SUPRA phenomenological aircraft model is that it will not be matched to a particular aircraft type. Rather it will represent class-specific aircraft behavior. By adjusting the appropriate parameters, the model can be reconfigured to represent different aircraft classes. In Fig. 5 this is illustrated for the difference in pitching moment characteristics of a T-tail with tail-mounted engines, on the one hand, and a low tail with wing-mounted engines on the other hand.

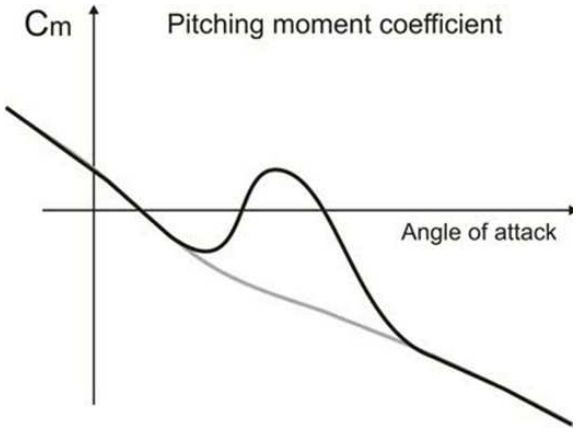


Fig. 5 Impression of pitching moment for low-tail (gray) and T-tail configuration (black).

3.3 Model validation

Recently a first validation of the augmented SUPRA aircraft model was carried out by GFRI and NLR with piloted simulator trials. At first, a generic low-tail, under-wing engines configuration was considered. The aircraft controllability and performance were assessed in nominal flight conditions, requiring small amplitude control, and also in attitude capture tasks, and various upset conditions, which require more aggressive control inputs.

According to Cooper and Harper aircraft handling qualities are: “Those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform the task required in support of an aircraft role.”[13] For the purpose of SUPRA the baseline aircraft should at least show acceptable handling qualities (Cooper-Harper rating 1-3) in normal, stabilized conditions preceding the upset event.

Based on this first model evaluation, the pilots concluded that the controllability of the aircraft model was, in general, satisfactory, with Cooper-Harper ratings below 3. As expected, they identified several dynamic flight parameters that still need further optimization, like a slight but deviant pitch-up tendency. Approach to stall, full stall and even spin behavior of the augmented model was considered a substantial improvement over the baseline model. Still, a stick shaker activation model should be included. Based on these findings the model will be further optimized in the next months before the test pilots from the SUPRA expert group will perform the final simulator validations.

4 Motion cueing

In order to provide the pilot with motion feedback, FFS are traditionally equipped with a hexapod motion platform, also known as “Stewart platform” (Fig. 6). The six linear actuators move synergistically to translate or rotate the simulator cabin. The motion is controlled by motion cueing software, which transforms the output of the mathematical

aircraft model into input signals for the simulator motion platform.



Fig. 6 The Generic Research Aircraft Cockpit Environment (GRACE) at NLR.

As shown in Fig. 7, linear ($F_{a/c}$) and rotational ($\omega_{a/c}$) aircraft motion signals are normally high-pass (HP) filtered, so as to merely reproduce brief motion onsets that stay within the limited simulator motion space. An additional low-pass filter (LP in Fig. 7) is often used to simulate sustained linear aircraft accelerations, such as the longitudinal acceleration during takeoff, by tilting the simulator cabin, a procedure know as "tilt coordination". Classical motion cueing is acceptable for a wide range of commercial pilot training scenarios, but it has clear limitations for upset simulation which may require reproduction of large angular rates, large attitude excursions, un-coordinated flight and sustained accelerations [13]. Within SUPRA, NLR and the Central Aerohydraudynamic Institute (TsAGI) in Russia, will investigate the possibility to amend their classical motion cueing filters to the specific needs of upset recovery simulation, for example by prepositioning the cabin to achieve a larger linear stroke. These experiments are planned towards the end of 2011.

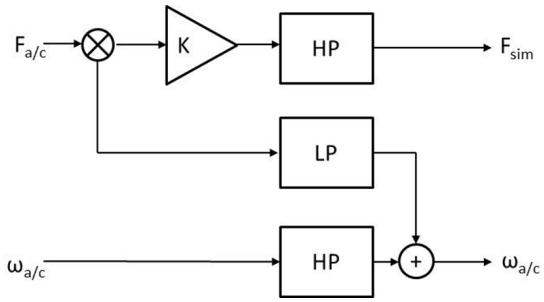


Fig. 7 Scheme of classical motion cueing algorithm, transforming linear ($F_{a/c}$) and angular ($\omega_{a/c}$) aircraft motions into simulator motion by a set of high-pass (HP) and low-pass filters (LP).

4.1 Centrifuge-based simulation

Hexapod motion platforms are inherently limited to a 1g regime. This limitation is overcome by centrifuge-based flight simulators, such as the new-generation research facility DESDEMONA (acronym for "DESorientation DEMONstrator Amst") at TNO (Fig. 8). The facility integrates a single-seat cockpit with a six degrees-of-freedom (DoF) motion platform (see Table 1 for specifications). For the purpose of SUPRA the originally military cockpit has been outfitted with a commercial aircraft flight deck. The simulator cabin is fully gimbaled, i.e. can rotate infinitely about all axes; can move vertically along a heave axis (±1m) and horizontally along a linear arm (±4m). The linear arm can rotate about its central vertical axis to generate sustained centripetal forces in the cabin. With the maximum arm of 4m and a maximum rotational speed of 155°/s

Table 1 Motion specifications of DESDEMONA.

	Centrifuge axis	Radius track	Heave track	Cabin roll	Cabin yaw	Cabin pitch
Position	>360°	+/- 4.0 m	+/- 1.0 m	>360°	>360°	>360°
Velocity	155 °/s	3.2 m/s	2.2 m/s	180 °/s	180 °/s	180 °/s
Acceleration	45 °/s^2	4.9 m/s^2	4.9 m/s^2	90 °/s^2	90 °/s^2	90 °/s^2

DESDEMONA can simulate sustained G-loads of up to 3g. Unique about DESDEMONA's six DoF motion capabilities is that it can combine onset cueing along the x, y and z-axis (like a hexapod simulator) with sustained acceleration cueing. In addition, unusual attitudes and large attitude changes (in excess of 60° bank or pitch) can be simulated one-to-one. Obviously, innovative motion cueing solutions have to be developed within SUPRA to fully employ these motion capabilities.

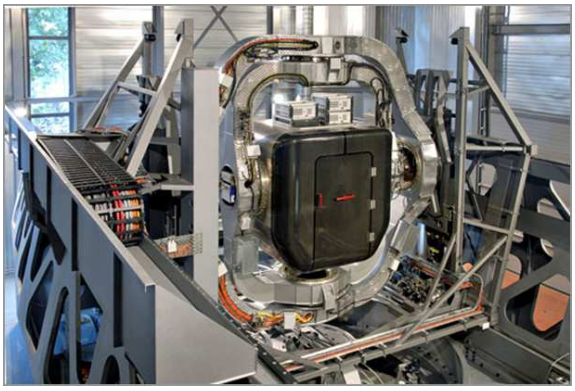


Fig. 8 DESDEMONA facility at TNO, the Netherlands.

The first motion cueing developments on DESDEMONA were focused on new G-cueing strategies which produce less false cues than conventional centrifuge solutions. In conventional centrifuge-based G-cueing the simulator cabin moves in the direction of the rotation (tangentially) and the cabin rolls outward when the G-force increases in order to keep the resultant force coordinated with the pilot's z-axis (in coordinated flight). Hence, when G-load increases the simulator spins up its centrifuge axis and the cabin rotates outward (roll for the pilot); when G-load decreases, the centrifuge axis slows down and the cabin rolls inward (opposite roll for the pilot). This conventional solution, used in all centrifuge-based flight simulators, suffers from a large false cue that gives the pilot a false sensation of rotation (tumbling), especially when decreasing the G-load after a loaded maneuver (e.g. pulling up from a nose-low attitude). With the extra rotational degrees of freedom of the DESDEMONA cabin the false tumbling cue can be largely avoided by applying G-cueing strategies that use less, or no roll motion of the simulator cabin. Instead another axis, such as the pitch and/or yaw, can be used to align the G-load in addition to the roll axis. The motion cueing developments are accompanied by psychophysical experiments to determine the motion settings that are best perceived by test pilots.

Ultimately, the extended simulator motion capabilities achieved within SUPRA will enable test pilots to identify the most important degrees of freedom to recover from particular situations. In real aircraft this is difficult, since motions in different axes are mechanically coupled. In a full-motion flight simulator it is possible to selectively reproduce some motions while ignoring others. We expect that the final evaluation trials will provide this kind of information.

4.2 Recovery performance with G-cueing

Efficient upset recovery in commercial transport aircraft may require the crew to load the aircraft up to the limit load of +2.5g, or up to +3.8g depending on design takeoff weight [11]. In the context of SUPRA we performed an experiment in DESDEMONA to study the effect of G-cueing on recovery performance of commercial pilots having negligible (less than 5h) experience with in-flight aerobatics or G-maneuvering[15].

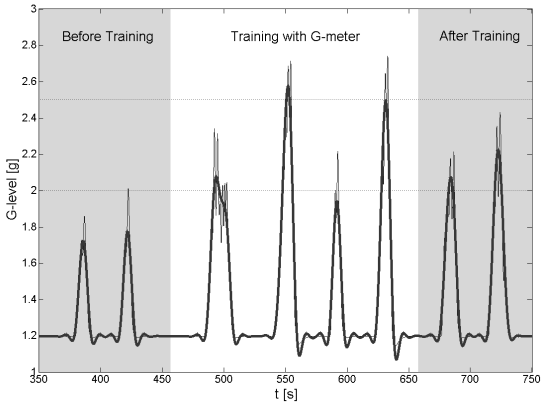


Fig. 9 Example recording of G-load produced by a pilot trying to pull 2g resp. 2.5g (dotted lines), before and after training with G-meter.

Before the experiment, all 13 participating pilots were asked to pull 2g and 2.5g,

respectively, without any feedback (e.g. instruments or out-the-window visual) but from their seat-of-the-pants. Next, they received a training session with feedback from a G-meter, where they were instructed to pull the same two G-loads twice. Finally, their performance was tested again without the G-meter. As the recording in Fig. 9 shows, pilots were hesitant in applying G at first: Before training they pulled on average 1.8g and 2.1g, in response to the target levels of 2.0g and 2.5g, respectively. In the re-test after the training with the G-meter, these values were 2.0g and 2.4g, respectively.

After a health break, the same pilots participated in the actual experiment in which we tested the effect of G-cueing on unusual attitude recovery. At the time of the experiment the advanced aerodynamic model was not yet available, so that a maneuver was chosen that remains within the normal validated flight envelope. The flight model and instrument panel were derived from a modern short- to medium-range generic twin-engine narrow-body jet airliner with wing-mounted engines and a fuselage mounted tail plane. Pilots were asked to recover from a 30° nose-low attitude as a simplified analogue for upset recovery requiring a maximum G-load of 2.5g. Each maneuver was initiated at 180kts, where the pilot pitched the aircraft down to a 30° nose-low attitude. During the recovery airspeed was kept within 320kts.

The pilots' performance was compared between a no-motion condition (fixed-base) and a G-cueing condition (centrifuge-based). The pilots were kept naïve about the motion conditions. The dependent variable was the maximum G-load computed by the aircraft model during the recovery.

As is shown in Fig. 10, with actual G-cueing pilots controlled the pull-up maneuver at the adequate G-level of 2.5g and with little variance. In contrast, in the Fixed-based condition they controlled the maneuver at G-loads significantly higher than 2.5g. In itself it may not be surprising that pilots performed inaccurately without any feedback on the G-load. However, the practical implication of this phenomenon is that flight simulators without G-cueing capabilities, such as the standard FFS, will evoke unrepresentative upset recovery behavior, overstressing the (simulated) aircraft.

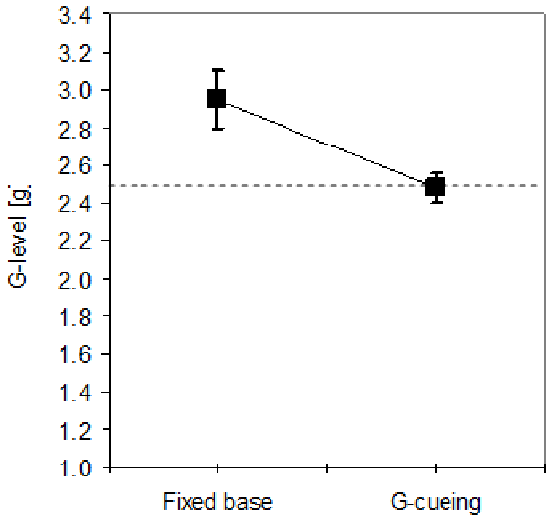


Fig. 10 G-load in Fixed-base and G-cueing condition during pull-up maneuver in Desdemona[15].

5 Conclusion

The SUPRA research project investigates possibilities to push the flight simulator aerodynamic and motion envelopes, thus enabling advanced training of upset recovery techniques. The aerodynamic modeling employs a unique combination of engineering methods, so as to capture the major aerodynamic effects that occur at high angles of attack.

The motion cueing research within SUPRA, on the one hand, aims to push the envelope of standard FFS by optimizing the motion cueing software; and on the other hand, investigates the effectiveness of a new-generation centrifuge-based simulator for the simulation of G-loads inherent of some upset recovery techniques. Preliminary results show that inexperienced pilots are hesitant in applying G, but after training their recovery performance with G-cueing is significantly better than without G-cueing.

Improved flight simulator technologies beyond the current state-of-the-art can be used to train pilots to detect and recover from upset conditions, and hence prevent LOC-I.

References

[1] Boeing Commercial Airplanes, Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations 1959 – 2010.

[2] Burks B., "Upset Recovery Training: A Call for a Higher Standard of Training," *RAeS Flight Simulation Conference: Towards the Edge of the Envelope*, Royal Aeronautical Society, London (UK), June 3-4, 2009.

[3] Flight Safety Foundation, "Industry Airplane Upset Recovery Training Aid, Rev. 2, "November 2008. http://www.flightsafety.org/pdf/AP_UpsetRecovery/AP_UpsetRecovery_Book.pdf.

[4] Lambregts A., Nesemeier G., Newman R., and Wilborn J., "Airplane Upsets: Old Problem, New Issues," *AIAA Modeling and Simulation Technologies Conference*, American Institute of Aeronautics and Astronautics, Honolulu (HI), August 18-21, 2008, AIAA-2008-6867.

[5] Cunningham K., Foster J., Shah G., Stewart E., and Ventura R., "Simulation Study of Flap Effects on a Commercial Transport Airplane in Upset Conditions," *AIAA Atmospheric Flight Dynamics Conference and Exhibit*, American Institute of Aeronautics and Astronautics, San Francisco (CA), August 15-18, 2005, AIAA-2005-5908.

[6] Goman M.G., Khrabrov A.N., "State-Space Representation of Aerodynamic Characteristics of an Aircraft at High Angles of Attack," *Journal of Aircraft* 31, (5), 1994, pp.1109-1115

[7] Abramov, N.B., Goman, M.G., and Khrabrov, A.N. "Aircraft Dynamics at High Incidence Flight with Account of Unsteady Aerodynamic Effects," *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, American Institute of Aeronautics and Astronautics, Providence (RI), August 16-19, 2004, AIAA-2004-5274.

[8] Prananta B.B., Veul R.P.G., Houwink R., Hounjet M.H.L. , and van Muijden J., "A generic flexible aircraft loads database system for fatigue analysis," *International Forum on Aeroelasticity and Structural Dynamics*, Stockholm, June 18-20, 2007.

[9] Soemarwoto B.I., Boelens O.J., Allan M. Arthur M.T., Butefisch K., Ceresola N., and Fritz W. "Towards the simulation of unsteady manoeuvre dominated by vortical flow," *33rd Fluid Dynamics Conference*, American Institute of Aeronautics and Astronautics, AIAA-2003-3528.

[10] Boelens O.J, Spekreijse S.P., Sytsma H.A., and de Cock K.M.J., "Comparison of measured and simulated flow features for the full-scale F-16XL aircraft," *45th AIAA Aerospace Sciences Meeting and Exhibit*, American Institute of Aeronautics and Astronautics, Reno (NV), January 8-11, 2007, AIAA-2007-489.

[11] Goman M., Abramov N., Khrabrov A., Soemarwoto B., "Aerodynamic Model Development for Simulation of Upset Recovery of Transport Airplane," *Applied Aerodynamics: Capabilities and Future Requirements*, Royal Aeronautical Society, London (UK), July 27-28, 2010.

[12] European Aviation Safety Agency (2010) Certification Specification CS-25.337, Amendment 10 (December 23, 2010). Retrieved from http://easa.europa.eu/agency-measures/certification-specifications.php#CS-25.

[13] Cooper G. E., Harper R. P., Jr., "The use of pilot rating in the evaluation of aircraft handling qualities," Technical Report NASA Ames Research Center, NASA-TN-D-5153, 1969.

[14] Fucke L., Biryukov V., Grigoriev M., Groen E., Wentink M., Field J., and Soemarwoto B. "Developing Scenarios for Research into Upset Recovery Simulation,", *AIAA Modeling and Simulation Technologies Conference*, American Institute of Aeronautics and Astronautics AIAA, August 2-5, 2010, Toronto (CA), AIAA-2010-7794.

[15] Ledegang W., Groen E., Wentink M., "Effect of G-cueing on Pilot Performance in Centrifuge-Based Simulation of Unusual Attitude Recovery," *AIAA Modeling and Simulation Technologies Conference*, American Institute of Aeronautics and Astronautics Portland (OR), August 8-11, 2011, AIAA-2011-6701.

CEAS 2011

Space Re-entry vehicles technologies

Preliminary Design of Minimum Weight Configurations for a Reentry Vehicle

A. Viviani, L. Iuspa, G. Lanzillo
Seconda Università di Napoli (SUN), Dipartimento di Ingegneria Aerospaziale e Meccanica (DIAM)
via Roma 29, 81031 Aversa, Italy – E-mail: Antonio.viviani@unina2.it

G. Pezzella
Centro Italiano Ricerche Aerospaziali (CIRA)
via Maiorise, 81043 Capua, Italy – E-mail: g.pezzella@cira.it

Keywords: *Winged Reentry Vehicle, Aerothermodynamics, Coons Surface parametric model, Optimization, Genetic Algorithm.*

Abstract

The current paper reports a multi-disciplinary study on the preliminary design of a Winged Re-entry Vehicle (RV-W). The main objective of this study is the searching for the minimum weight configuration taking into account both the thermal and the operational constraints. The structural stresses are considered by the mean of the asymptotic dynamic pressure and the normal load factors peaks, and crew liveability conditions are accounted for by controlling that the inner temperature remains locally below a duly imposed value. An optimization procedure, using Genetic Algorithm (GA), is performed to find the optimal configuration. The RV-W shape is modelled by a suitable parametric model, based on Coons surfaces, and a parametric insulating material distribution is foreseen. The analysis is performed over the hypersonic flight regime range because it is characterised by the larger thermal and mechanical stresses. The re-entry trajectory is calculated by the three-degree-of-freedom model with no bank angle. The shape aerodynamics is performed via a hypersonic panel method and the thermal state of the external surface is determined with the radiative equilibrium hypothesis and boundary layer model. The material for modelling the Thermal Protection System (TPS) is Li-900.

Nomenclature

C_{he}	Stanton number
c_p	specific heat at constant pressure
D	drag
g	gravity acceleration
g_j	constraint function j
H_j	penality exponent of constraint function j
L	lift
k	thermal conductivity
m	RV-W mass
n_z	normal load factor
q_{gw}	heat flux in the gas at the wall
$q_{rad,w}$	thermal radiation heat flux at the wall
q_∞	free flow dynamic pressure
R	RV-W position vector magnitude
r	altitude
r_n	nose radius
T	temperature
T_{aw}	adiabatic wall temperature
T_w	wall temperature
t	time
th	TPS thickness
V	flight velocity
x	length coordinate
$\underline{x}$	design variable array
x_i	design variable
W	weight
γ	flight path angle
ρ	density
ρ_e	free stream density
Φ	local body angle

1 Winged Reentry Vehicle

A reentry vehicle has to perform an unpowered gliding flight, transferring safely a crew from a Low Earth Orbit (LEO) to the Earth's surface. At the entry into the Earth's atmosphere, the RV-W owns a high energy, both potential and kinetic, and basically performs a braking mission intended to reduce its own speed during the descent. The shape design and the reentry trajectory parameters have to increment the vehicle deceleration, mainly due to wave drag, through the higher and less dense atmosphere layers in order to reduce the thermal and mechanical loads. As consequence, the vehicle is characterised by a blunt shape, a flat windward side and a large Angle of Attack (AoA). These features magnify the wave drag, which is the main dissipation mean to convert the RV-W kinetic energy in thermal energy owned by the fluid behind the bow shock ahead the RV-W. Finally, a Thermal Protection System is necessary to protect both the vehicle structure and the crew during the descent.

2 Parametric Model

The outer shape parametric model is built by *Coons*-surface patches having as boundaries proper pieces of the upper, the lower, the wing planform and the cross sections mainlines [1]. Some possible parametric shapes are shown in Fig. 1.



Fig. 1 Examples of parametric outer shapes

The outer shape is built to enclose a fixed volume accommodating a four-people crew. A volume of 1500x800x1500 mm is assigned to each crew member and a 20% increase is foreseen for the global ground dimension. Four potential cabin layouts are considered, as shown in Fig. 2.

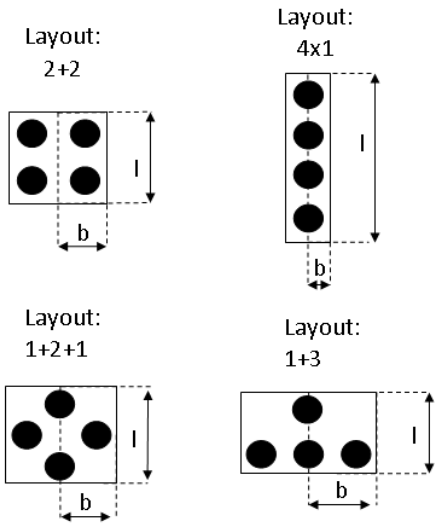


Fig. 2 Crew arrangements

As shown in Fig. 3, the symmetry plane forward outline is controlled by the points F_1 and F_2 and they are connected to the cabin corners by the upper and the lower line. The F_2 abscissa is one half of the l_1 design variable whereas the ordinate is controlled by a non-dimensional parameter defining h_2 as a fraction of h.

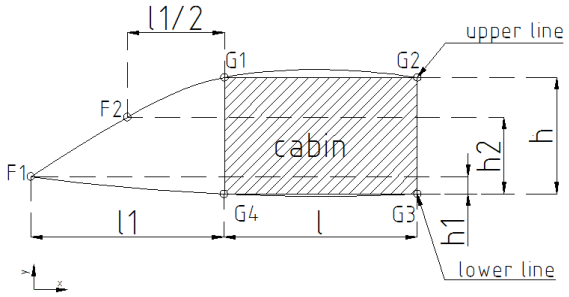


Fig. 3 RV-W symmetry plane forward outline

The nose radius lying in the symmetry plane is defined by the dimensional parameter *rnose*, which control the fillet between the upper and the lower line. The backward symmetry plane shape is modelled by two B-splines that connect the bottom of the crew box to the F_3 point, as shown in Fig. 4.

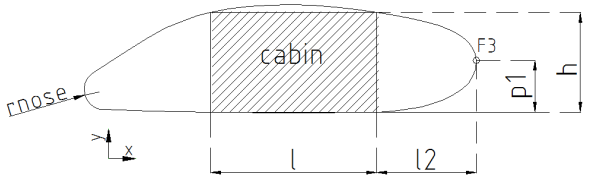


Fig. 4 RV-W symmetry plane outline

The RV-W wing planform, shown in Fig. 5, is designed exploiting a segmented line passing through the control points E_0, E_1, E_2, E_3, E_4.
The E_i *(i=1,2,3,4)* points ordinates are controlled by a set of non-dimensional parameters which allocate the E_i-points along the RV-W chord, whereas the E_0 ordinate is coincident with the cabin back side. The E_i *(i=0,1,2,3,4)* points abscissas are mainly influenced by the semi-wingspan, that is provided explicitly by the dimensional variable k_1 whereas the k_i *(i=0,1,2,3,4)* quantities are expressed as a fraction of k_1.

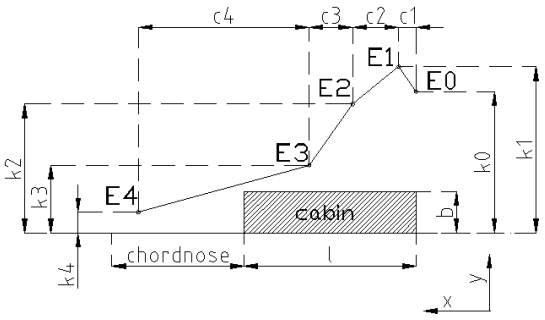


Fig. 5 Segmented wing planform

The wing planform smoothness depends on the fillet between the pairs of connected segments. A convenient routine has been realised to control all the wing fillets by a single parameters in order to not introduce less sensitive variables. As shown in Fig. 6, the smoothed wing outline is linked to the nose and the bottom of the model by two B-splines and the model wireframe is completed by five transversal sections. They intersect the model chord in five static locations partly expressed as function ratio of the vehicle and partly depending on the cabin and semi-wingspan position. The most forward and the most backward cross sections are obtained via a B-splines connection, whereas the central ones are developed from a segmented line connecting the control points, as shown in Fig. 7.

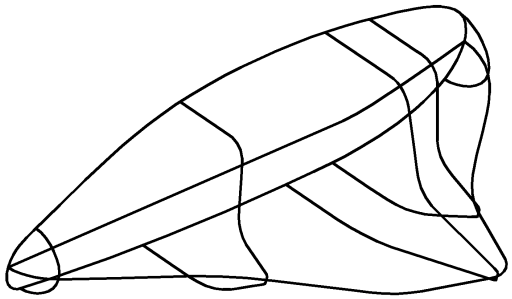


Fig. 6 Half-model wireframe

A set of non-dimensional parameters control the a_i design variables, which are expressed as a function of the cabin height and local wingspan. The cross-section smoothness is regulated by a unique parameter acting on all the section fillets.

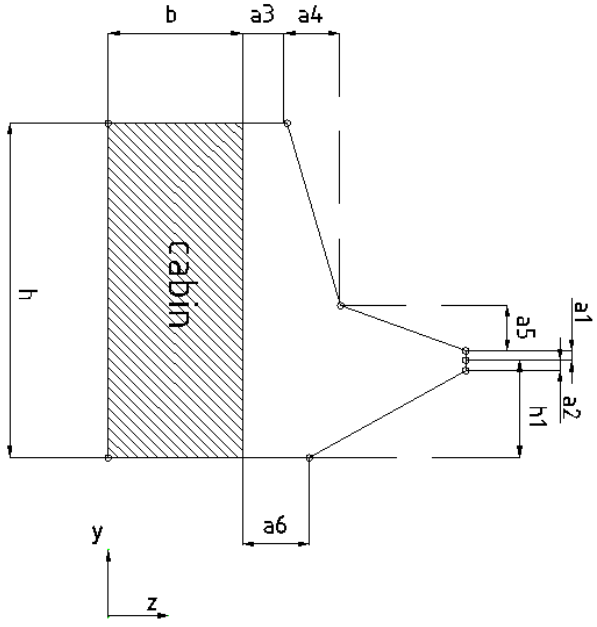


Fig. 7 Cross-section segmented line

Finally, the outer shape discretization is made by four-vertex panels, which are exploited as

both thermal and aerodynamic computational elements.

3 Reentry trajectory

The reentry trajectory model relates the vehicle velocity with the altitude so it defines the RV-W mission profile. The motion is completely described by the General Equation for Planetary Flight [2]. The vehicle is considered as being a mass point, so the descent is modelled by the three-degree-of-freedom model. The mission profile contemplates a gliding flight, which develops completely into the descent plane as no bank angle is contemplated. The re-entry flight is characterised by strong deceleration, so the centrifugal and Coriolis terms involving the planet angular velocity are neglected. The equations describing the RV-W descent motion into the planetocentric frame – considered as an inertial frame — reduce to:

$$\begin{aligned} \frac{dV}{dt} &= -\frac{D}{m} - g\sin\gamma, \\ V\frac{d\gamma}{dt} &= \frac{L}{m} - g\cos\gamma + \frac{V^2}{R}\cos\gamma \end{aligned} \tag{1}$$

The altitude rate of change is provided by the kinematic equation:

$$\frac{dr}{dt} = V\sin\gamma. \tag{2}$$

The range of speed is extended until $M_\infty \geq 5$ and the aerodynamics coefficients, C_L and C_D, are calculated by the mean of the hypersonic panel method HYPER [3] [4]. The RV-W mass is valued via statistical formulas, considering each vehicle sub-system separately [5]. The TPS mass is calculated from its distribution. The velocity and the altitude initial condition are set assuming a de-orbit manoeuvre starting from a circular orbit from 120 km, so they read:

$$\begin{aligned} r_{t=0} &= 120km, \\ V_{t=0} &= 7835ms^{-1}. \end{aligned} \tag{3}$$

An initial negative flight-path angle value, $\gamma(t_0)$, is necessary to perform a descent flight.

4 Thermal Protection System

The Thermal Protection System has to protect efficiently the re-entry vehicle interior from the aerodynamic heating during the descent. It is modelled with LI-900 material, which is widely employed in Space Shuttle Orbiter TPS. Its physical properties of interest are summarised in Table 1 [6].

Table 1. LI-900 material properties

k	4,76E-2W/mK
ρ	1.44E2kg/m^3
c_p	6.28E2J/kgK
ε	0.88

The TPS thermal state is evaluated by a transient, one-dimensional analysis for each outer shape four-vertex element. A parametric law operating on the body natural direction allocates the insulating material. It contemplates a major amount of insulating coating for the highest thermal stressed area, *i.e.* the windward side and especially the stagnation point area. A two-segment, linear law allocates the insulating coating to the elements standing the lower symmetry line. Its distribution is governed by four parameters acting on: the maximum TPS thickness, the thickness rates of decrease, and the law shift point. The insulating material coating decreases from the lower to the upper side in a parametric linear way because the farther from the symmetry line the element the less the thermal stress. Therefore, the whole thickness distribution is managed by five parameters. The TPS is considered as a set of volumes of insulating material in regard to the thermal state computation. Each volume is modelled as independent from the surrounding ones, *i.e.* adiabatic sidewalls are assumed. The transient, one-dimensional heat conduction model through the thickness reads:

$$c_p \frac{\partial T}{\partial t} \rho th_{ele} = k \frac{\partial^2 T}{\partial x^2} . \quad (4)$$

The TPS is assumed to be in thermal equilibrium with the atmosphere at the entry. The inner surface is modelled as adiabatic, and it is a conservative hypothesis as being the internal TPS temperature as constraint. The wall temperature is calculated according to the radiative equilibrium hypothesis [7]. Thus, the thermal problem formulates:

$$\begin{aligned} c_p \frac{\partial T}{\partial t} \rho th_{ele} &= k \frac{\partial^2 T}{\partial x^2} , \\ T_{|t=0} &= T_{atm|r=120km} = 285K , \\ T(t)_{|x=0} &= T_{re}(t), \\ \frac{\partial T}{\partial x}(t)_{x=th_{ele}} &= 0. \end{aligned} \quad (5)$$

Further hypothesis are formulated to predict the outer surface thermal state: laminar boundary layer, the gas temperature at the surface is equal to the wall temperature and the one-dimensional heat transfer mechanism normal to surface is the only one considered. The heat flux balance at the walls considers only the radiative heat flux from the gas to the wall, q_{gw}, and the thermal radiation heat flux, $q_{rad,w}$, cooling the surface, and it reads:

$$\begin{aligned} q_{gw} - q_{rad,w} &= 0, \\ q_{gw} &= \varepsilon \sigma T_w^4 . \end{aligned} \quad (6)$$

The q_{gw} is predicted by using the relation proposed in [8]:

$$q_{gw} = C \rho_e^N V^M . \quad (7)$$

The quantity C, and the constants N and M depend on the considered part of the vehicle. Next Eqs. (8) refer to the stagnation point:

$$\begin{aligned} M &= 3, \; N = 0.5, \\ C &= \left(1.83E-8\right) r_n^{-1/2} \left(1 - \frac{c_{pw} T_w}{0.5V^2}\right) . \end{aligned} \quad (8)$$

Furthermore, the surface is split in windward and leeward side via the local body angle with respect to the free stream flow, Φ. The heat flux in the gas at the four-vertex element belonging to the windward side, Φ sloped and whose centre is x away from the stagnation point, is modelled by the large-AoA flat-plate heat-transfer model, summarized in Eqs. (9):

$$\begin{aligned} C_1 &= k_C(\phi)(x^{-1/2})(1-g_w), \\ k_C &= (2.53E-9)(\cos\phi)^{1/2}(\sin\phi), \\ g_w &\approx \frac{c_{pw} T_w}{0.5V^2} . \end{aligned} \quad (9)$$

The q_{gw} at four-vertex elements belonging to the leeward surface is modelled via the laminar compressible boundary layer model,

$$q_{gw} = C_{he} \rho_e U_e \left(T_{aw} - T_w\right) . \quad (10)$$

5 Operational Constraints

The RV-W is stressed by both mechanical and thermal loads during the descent. The quantities conventionally adopted as measure of the mechanical constraints, in a preliminary design stage, are the asymptotic dynamic pressure and the normal load factor peaks experienced during the re-entry trajectory [9]. The free flow dynamic pressure, q_∞, is a structural load indicator, as being proportionally to aerodynamic forces. Its admissible peak value is assumed to be 14kPa, the same considered for the Space Shuttle Orbiter. The normal load factor, n_z, concerns both the crewmembers g-loads and the re-entry vehicle structure stresses. The maximum n_z peak value is assumed as being 2.5. The thermal constraint is considered to ensure both liveable condition to the crew and a acceptable structural thermal load. The interior temperature is calculated locally and if the TPS interior side overtakes locally the temperature of 422K during the descent [10], the thermal constraint is considered to be violated. The thermal constraint is naturally vectorial. A proper manipulation is necessary to reduce it to a scalar function.

6 Minimum mass configuration: the optimization procedure

The RV-W weight minimization is the optimization target. The structural and thermal constraint functions are exploited to evaluate the configuration feasibility. As consequence, the optimization procedure is single-objective and constrained, and it formulates:

$$\begin{aligned} &\min! \, W\left(\underline{x}\right) \\ &\underline{x}_i \le x_i \le \bar{x}_i \quad i = 1, 2, ..., I \\ &g_j^{MIN} \le g_j\left(\underline{x}\right) \le g_j^{MAX} \quad j = 1, 2, ..., J \end{aligned} \qquad (11)$$

$W(\underline{x})$ states for the re-entry vehicle weight, $\underline{x}$ is the configuration array filled by $I=30$ x_i scalar parameters governing the crew layout, the outer shape design, the angle of attack, the initial flight-path-angle and the TPS distribution law. The constraint functions, $g_j(\underline{x})$, specify as follows:

$$\begin{aligned} &0 \le \frac{q_{\infty,MAX}}{q_{\infty,LIM}} \le 1, \\ &0 \le \frac{n_{z,MAX}}{n_{z,LIM}} \le 1, \\ &T_{constraint} = \sum_{i=1}^{NUMELE} \left(T_{i,MAX} - T_{LIM}\right) \delta_i \le 0 \\ &\delta_i = \begin{cases} 1 & T_{i,MAX} > T_{LIM} \\ 0 & otherwise \end{cases}. \end{aligned} \qquad (12)$$

The optimization procedure is performed via a Genetic Algorithm method implemented in *ProGenie* [11]. The execution flow is shown in Fig. 8. The computational cost of a single configuration has reduced as much as possible to obtain an optimization procedure computationally affordable and so a massive use of the array programming technique is made. The thermal problem is evaluated simultaneously for each computational volume at each time step, and the re-entry vehicle geometric properties are calculated by a vectorial procedure. The single configuration computational cost is about thirty seconds on an up-to-date, entry level computer.

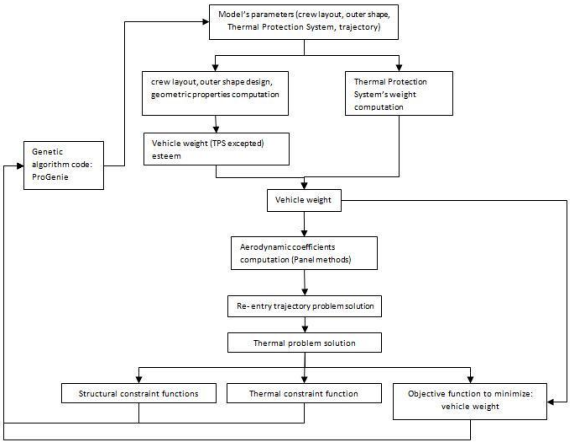


Fig. 8 Optimization procedure execution flow

The fitness function reads:

$$fitness\left(\underline{x}\right) = \frac{W_{\max} - W\left(\underline{x}\right)}{\prod_{j=1}^{J} k_j\left(\underline{x}\right)}$$

$$k_j\left(\underline{x}\right) = \begin{cases} \left(\dfrac{g_j^{\max} - g\left(\underline{x}\right)}{g_j^{\max} - g_j^{\min}}\right)^{H_j} & g_j\left(\underline{x}\right) < g_j^{\min} \\ 1 & g_j^{\min} < g_j\left(\underline{x}\right) < g_j^{\max} \\ \left(\dfrac{g_j\left(\underline{x}\right) - g_j^{\min}}{g_j^{\max} - g_j^{\min}}\right)^{H_j} & g_j\left(\underline{x}\right) > g_j^{\max} \end{cases}. \qquad (13)$$

7 Results

The most performing RV-W configuration and the related optimization procedure are presented. Preliminary optimization runs are necessary to set GA general parameters, the design variables ranges and bit resolutions. The GA settings are summarised in Table 2.

Table 2. GA general parameters setup

Number of generations	50
Individuals per standard population	50
New individuals at each generation	50
Selection operator	Roulette Wheel
Crossover operator	Crossover Single Cut
Mutation operator	Yes, dynamic
Starting population	Extended
Fitness scaling	Cosine law

For the sake of brevity, the results show only the time history of the more physical meaningful design variables for the most performing individual at each generation. The ranges of some of the shown parameters are reported in Table 3.

Table 3. Design variables range

l_1	2500 mm	5000 mm
l_2	600 mm	1600 mm
k_1	1200 mm	2000 mm
r_n	10 mm	300 mm
α	35 deg	42 deg
th_1	10 mm	400 mm
H_1	0.05	0.3
P_1	0.2	0.7

In Fig. 9 and Fig. 10, the l_1 and the l_2 time history outline the longitudinal dimensions of the most performing individuals.
The wing span trend is shown in Fig. 11

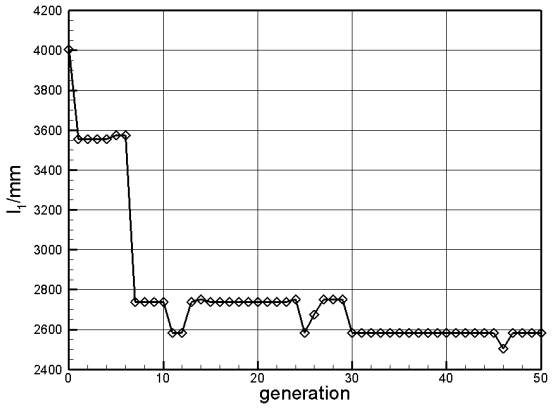


Fig. 9 l_1 time history

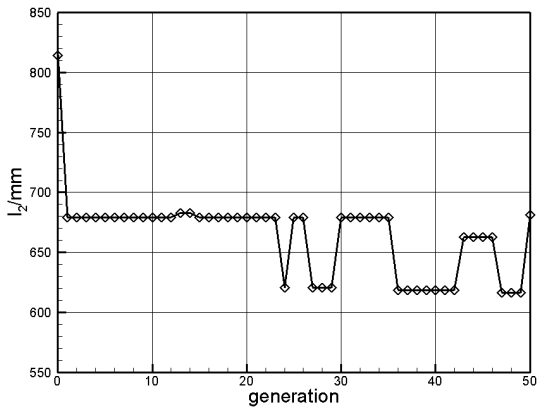


Fig. 10 l_2 time history

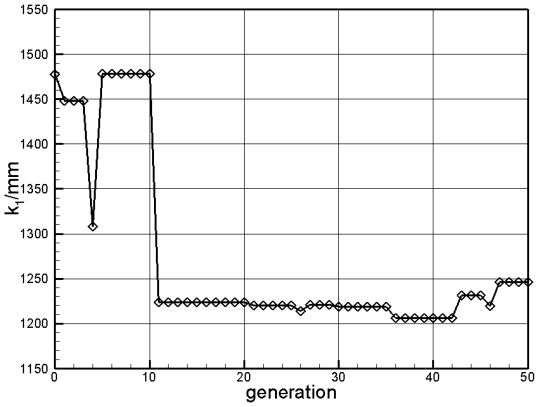


Fig. 11 k_1 design variable

The wing planform height is controlled by the design variables H_1 and P_1, expressed as a fraction of the cabin height and represented in Fig. 12. The crew layout is 2+2-type. The nose radius time history is shown in Fig. 13.

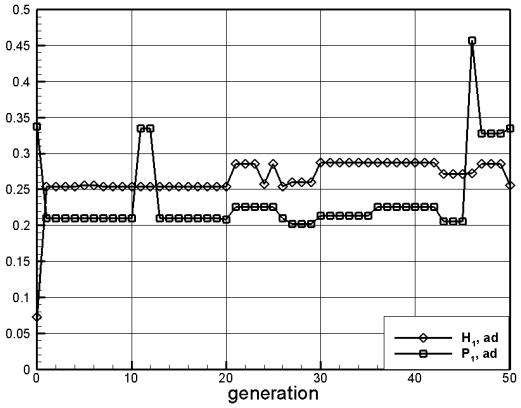


Fig. 12 H_1 and P_1 time history

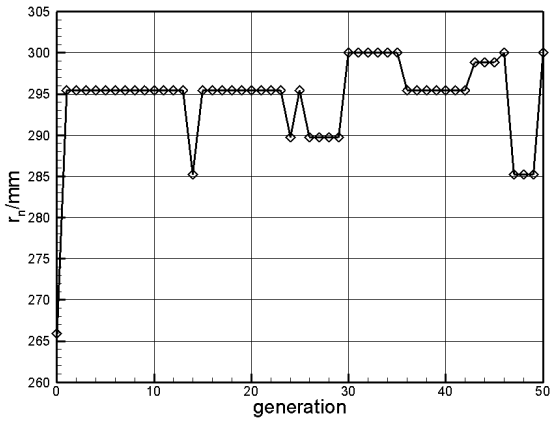


Fig. 13 r_n time history

In Fig. 14 and in Fig. 15 the trends of the parameters governing the wing planform shape are reported.

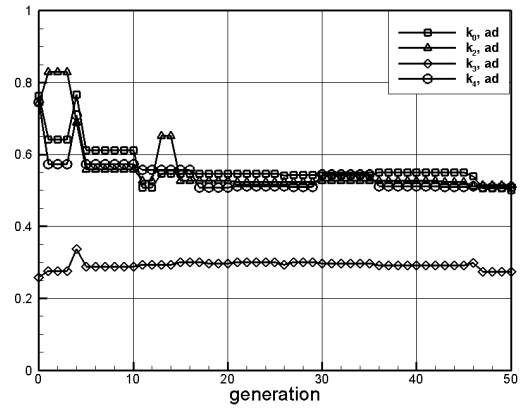


Fig. 14 k_0, k_2, k_3 and k_4 time history

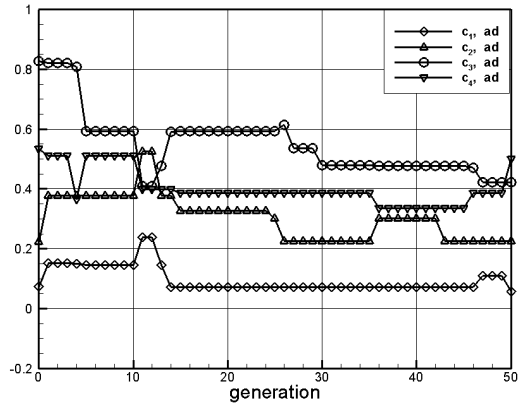


Fig. 15 c_1, c_2, c_3 and c_4 time histories

The insulating material distribution is managed by th_1, in Fig. 16 , and the parametric distribution law design variables. Some examples of TPS coating are shown in Fig. 17, Fig. 18, Fig. 19.

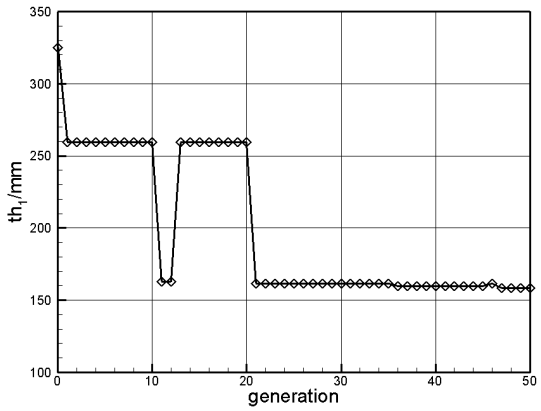


Fig. 16 th_1 time history

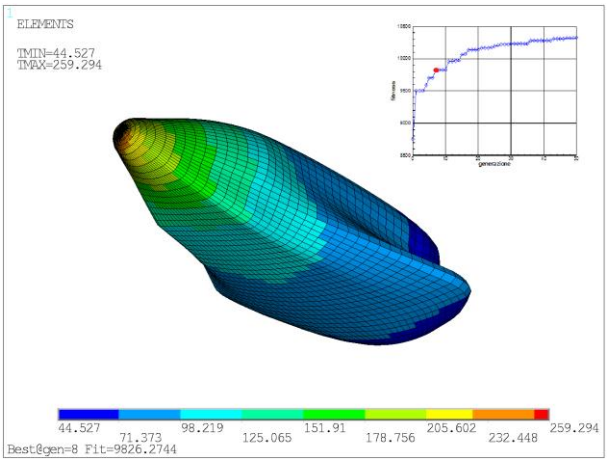


Fig. 17 Shape and TPS distribution at 8th generation

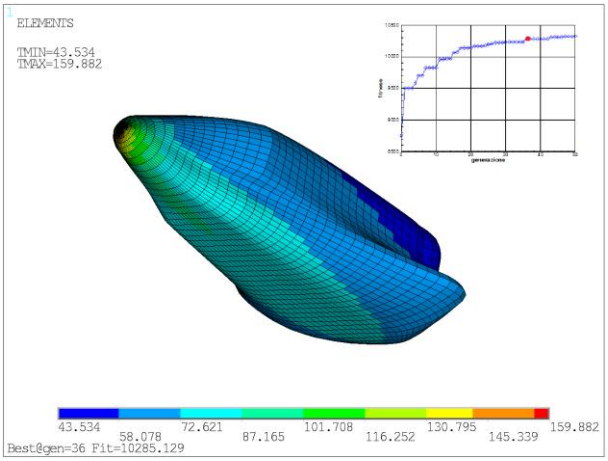


Fig. 18 Shape and TPS distribution at 36th generation

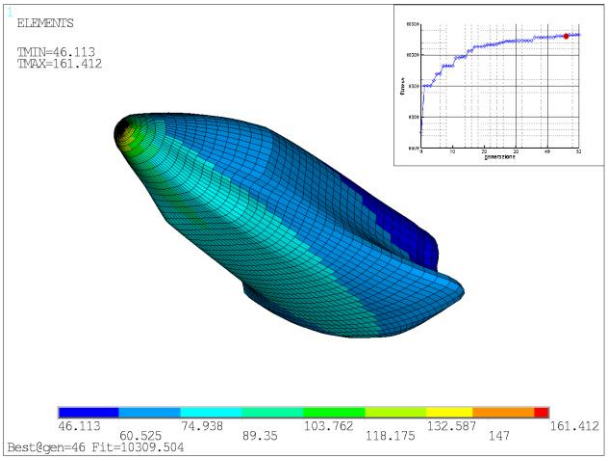


Fig. 19 Shape and TPS distribution at 46th generation

The aerodynamics is governed by the angle of attack shown in Fig. 20, and the related aerodynamic coefficient in Fig. 21.

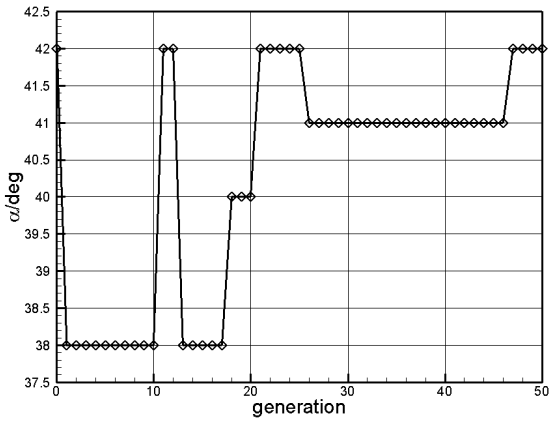


Fig. 20 α time history

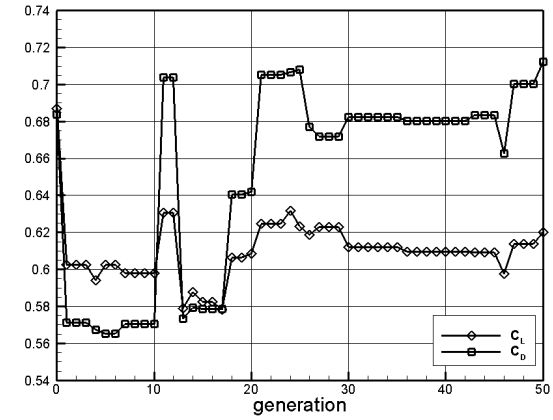


Fig. 21 C_L and C_D time histories

Furthermore, the initial flight-path angle γ_0 is equal to -1°. The TPS weight and the RV-W weight time histories are shown in Fig. 22 and Fig. 23; the fitness trend is reported in Fig. 24.

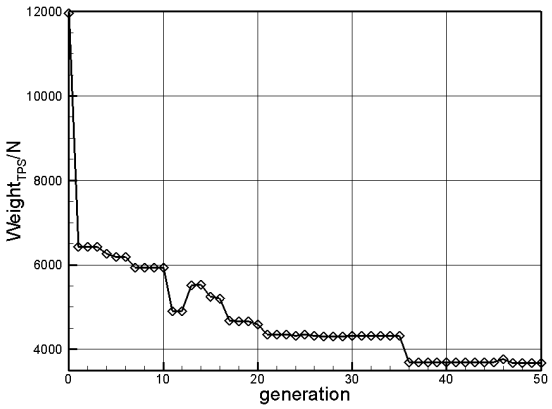


Fig. 22 TPS weight time history

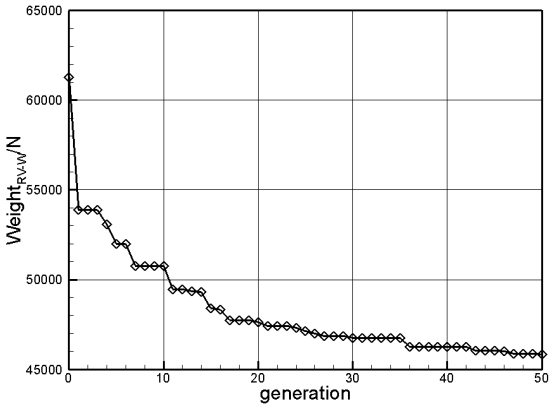


Fig. 23 RV-W weight time history

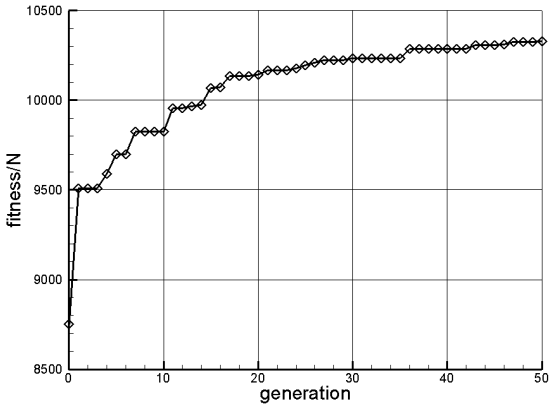


Fig. 24 fitness time history

The optimum is found at the 50th generation, its outer shape and the TPS thickness distribution are shown in Fig. 25, and its main characteristics are provided in Table 4.

Table 4. Optimum main design parameters

Layout	2+2-type
l_1	2583.088mm
l_2	681.133mm
k_1	1246.613mm
r_n	299.165mm
α	42deg
γ_0	-1deg
th_1	158mm
C_L	0.62
C_D	0.712
W_{TPS}	3670.077N
$W_{RV\text{-}W}$	45835.076N
$q_{\infty,peak}/q_{\infty,lim}$	0.825
$n_{z,peak}/n_{z,lim}$	0.852
$T_{constraint}$	0

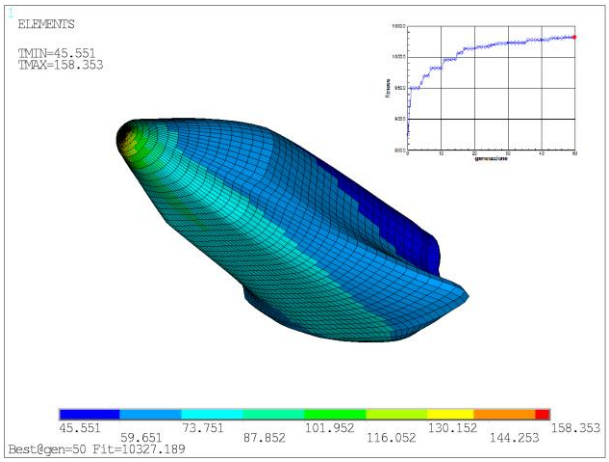


Fig. 25 Optimum outer shape and TPS distribution

8 Conclusions

The RV-W outer shape is characterised by a trade-off between the need for a compact and lighter configuration and the need for an adequate aerodynamic drag. As consequence, the longitudinal dimensions and the wingspan do not approach their minimum values due to the necessity of a satisfactory drag area. The most performing wing layout is delta-type and backward. The wing planform height and the roundness of the windward side give a blunt shape to the vehicle and, as expected, the nose radius value is large in order to reduce the local thermal load. The TPS thickness at the nose reaches its optimum value after some generations; further reduction is not admissible because of the thermal constraint. As showed, the TPS mass saving is obtained by a reduction of the insulating material standing both the leeward side and the less thermal stressed windward side area. The optimum angle of attack, 42°deg, and the hypersonic flight regime efficiency, about 0.8, are in agreement with the data available in the literature for the RV-W class of vehicle. As expected, the initial flight-path angle is -1°. As the γ_0 absolute value increases, the descent speed increases, and also mechanical and thermal stresses increase.

As the optimization procedure advance, the TPS weight and the RV-W weight largely reduce and the fitness increases because more performing and feasible individuals are generated. In that regards, further enhancements are foreseen. The study of the RV-W shape model behaviour in supersonic and subsonic regimes will be also considered. Several TPS materials will be taken into account to achieve both mass saving and reusability, and the downrange and cross-range performance analysis will be implemented.

References

[1] Piegel L., Tiller W., *The NURBS Book*, 2nd ed., Springer, 1996.
[2] Weiland C., *Computational Space Flight Mechanics*, Springer, 2010.
[3] Public Domain Aeronautical Software. *Hypersonic Arbitrary Body Program*, http://www.pdas.com/hyper.html.
[4] Gentry A. E., Smith D. N., Oliver W. R., " The Mark IV Supersonic-Hypersonic Arbitrary Body Program User's Manual", Technical Report, Air Force Flight Dynamics Laboratory, 1973.
[5] Rohrschneider R. R., "Development of a Mass Estimating Relationship Database for Launch Vehicle Conceptual Design", Georgia Institute of Technology, 2002.
[6] National Aeronautics and Space Administration. *TPSX Material Properties Database*, www.tpsx.arc.nasa.gov.
[7] Hirschel E.H., *Basics of Aerothermodynamics*, Springer, 2004.
[8] Schneider S. P., " Methods for Analysis of Preliminary Spacecraft Designs", Purdue University, 2003.
[9] Hirschel E.H., Weiland C., *Selected Aerothermodynamic Design Problem of Hypersonic Flight Vehicles*, Springer, 2009.
[10] Myers D.E., Martin C. J., Blosser M.L., "Parametric Weight Comparison of Advanced Metallic, Ceramic Tile and Ceramic Blanket Thermal Protection System", National Aeronautics and Space Administration, 2000.
[11] Iuspa L., Scaramuzzino F., "A bit-masking oriented data structure for evolutionary operators implementations in genetic algorithms", *Soft Computing*, Vol. 5, No. 1, 2001, pp. 58-68.

New Manufacturing Method of Carbon Carbon Production for Re-entry systems

Marta Albano, Plinio Coluzzi, Carlo Vassalli, Andrea Grilli, Mario Marchetti
Sapienza Università di Roma (DIAEE) , Italy

Keywords: *Carbon/Carbon, genetic algorithm, optimization, pyrolysis*

Abstract

Re-entry systems has been one of the most critical and unavoidable conditions. In fact, the re-entry environment imposes the most severe aerodynamic heating in addition to shock and acoustic loads. Like the launch environment, the re-entry environment also includes natural conditions such as wind, rain, hail, sand, and dust.
This problem can be handled only with a right choice of materials, as the re-entry corridor is narrow. Advanced composite are the actual solution to this problem and Carbon Carbon is the widely used. The problem in their manufacturing is nowadays not solved. The chemical process has not been understood yet and the process parameter are not optimized.
In this study a mathematical model of the pyrolysis is analyzed and it is optimized in order to obtain the best process parameter combination.

1 Introduction

Advanced composites are "engineered" materials with high performance properties and characteristics. They are usually light in weight, strong and stiff.

Among advanced composites, Carbon-Carbon (C/C) composites are widespread. C/C composites are largely used in space and mechanic high temperature applications.
C/C composites combine the many desirable properties of fiber-reinforced composites (such as high specific strength, stiffness and toughness) with the refractory properties of carbon at high temperatures. Their characteristics are suitable for space applications like aircraft brakes, rocket nozzles, missile nose tips and leading edges of the Space Shuttle. [1]
The production of C/C is expensive and time-consuming. The difficulties begin from the carbon fibers production to the pyrolysis process. The pyrolysis process is so a critical aspect that nowadays it is still not well understood. Actually it is still missing a universal recognized mathematical model and an optimization of the process parameters and this is the object of this study.
A chemical model is implemented in a genetic algorithm in order to optimize the process parameters, in fact genetic algorithms are widely used in the scientific world as an optimization instrument.

2 The liquid phase impregnation method

The manufacturing method chosen is the Liquid phase impregnation. This process consist in:

- Preparation of a preform using a carbon/phenolic prepreg provided by Advanced Composite Group.
- Pyrolysis of the phenolic resin at suitable temperatures and time.
- Infiltration of the pyrolized plate by using liquid phenolic resin. Different infiltration methods has been used
- Infiltration – pyrolysis cycle is repeated several times (3-9) until the desired density is achieved.

The goal of this study is to optimize the pyrolysis process by the implementation of a mathematical pyrolysis model into a genetic algorithm.

3 The choice of phenolic resin

Phenolic resins have prominent features like:

- Excellent thermal behaviour
- High strength level
- Long term thermal and mechanical stability
- Excellent fire, smoke, and low toxicity characteristics
- Excellent electrical and thermal insulating capabilities
- Excellent cost performance characteristics

Moreover very few resins have carbon yield over 40% on pyrolysis. Almost all phenolics resins are good carbon forming materials.[2]
The carbon yield of this resin is the highest among thermosetting resin, as the table below shows.

Precursor	*Carbon yield (%)*
Phenolic resins	50–55
Furan resins	50–60
Oxidized polystyrene	55
Polyvinyl alcohol	50
Polyacrylonitrile	44
Polyvinylidene chloride	25
Cellulose	20
Epoxy resins	5
Polystyrene	5
Polymethylmethacrylate	5

Figure 1 Carbon yield of Thermoset resin.[2]

The phenolic ring bonds are less stable because the phenolic nuclei are in opposite sense and this lead to an increase in the bond length between aromatic and aliphatic carbon, with a concomitant decrease of bond energy. This phenomena results in an increase of carbon yield.[3]
But phenolic resin also have a T_g greater than the onset of carbonization temperature (this is due to the high primary covalent and secondary N bond cross-link density), and so the material will not carbonize in a rubbery state and the pyrolysis gases will explode the porosity.
Although the density of phenolic resin composite is lower than pitch ones, as the diagram below shows, the range of process temperature are much lower and in the range of 500-1000°C the densities are comparable. The carbon yield is always higher although the basal spacing is higher too. The pressure does not increase the carbon yield as for the pitch.

4 Pyrolysis process

The pyrolysis reactions convert the phenolic matrix to amorphous carbon. Pyrolysis is basically a heating cycle. The precursor is heated slowly in a reducing or inert environment, over a range of temperature that varies with the nature of the particular precursor. The organic material is decomposed into a carbon residue and volatile compounds diffuse out to the atmosphere. The process is complex and several reactions may take place at the same time such as dehydrogenation, condensation and isomerization. The carbon content of the residue is a function of the nature of the precursor and the pyrolysis temperature.
After carbonization, the residual material is essentially all carbon. However its structure has little graphitic order and consists of an

aggregate of small crystallites, each formed of a few graphite layer planes with some degree of parallelism and usually with many imperfections. These crystallites are generally randomly oriented. The carbonized material is often called “amorphous” or “ baked carbon”. It is without long-range crystalline order and the deviation of the interatomic distances of the carbon atoms is greater than 5%. Amorphous carbon is hard, abrasion resistant, brittle, and has low thermal and electrical conductivities. In most instances amorphous carbon is only the intermediate stage in the production processes[4].

The potential for damage during the pyrolysis is great, due to thermo-mechanical stresses that develop in the composite and pressures from evolving gas products. In order to keep stresses below the critical level that would cause damage to the part, the rate of gas production is controlled by controlling the heating rate of the composite. Because the pyrolysis reactions are not well understood, a large factor of safety must be used when determining heating rates. This leads to excessively long cycle times. Also the limited understanding of the reactions requires that process cycles be developed by trial and error methods. Long cycle times and trial and error process cycle development leads to high processing costs. An increased understanding of the pyrolysis reactions is essential for improvement of the processing of carbon/carbon.
The structure of cured phenolic matrix in a carbon/phenolic composite is primarily methylene bridged phenolic units. Through pyrolysis it is desired to eliminate all non-carbon species and produce a char of coalesced carbon rings.

The general thought [6][5]about the kinetic scheme of the chemical reaction of the pyrolysis is that it is composed by three regions:

- In the first phenolic resin breaks its crosslinks and decomposes into Cresol and Phenol.
- In the second region new crosslinks are formed and various gases are developed.
- In the third region the ring structure changes and final char is formed.

All these regions are characterized by the temperature and activation energy. in fact, while the reactions of the first region begin at 300°C, the ones of the second take place in a range of 400°C to 600°C and the ones of the third region only above 600°C.

Not only the temperature influences the final pyrolyzed composite structure, in fact also the carbonization rate is very important. From the rheological study of Kuo et al.[9][10], it is clear that the mechanical properties, the average wear and obviously the internal microstructure (think about porosity) are strictly influenced by the carbonization rate.
Moreover, this bring the correlation between the time and temperature, in fact, from the mathematical model which we study, it is possible to understand their mutual influence.

5 Mathematical model and implementation

The mathematical model of the pyrolysis imply the comprehension of the chemical model of the pyrolysis as aforementioned. A model inspired by the studies of Dimitrienko [7][8]has been used.
It consist of four equations which describes the variation of the main chemical species which develop during the process.
The four species are the carbon fiber, which are supposed to no react during the process as there is an inert environment, the phenolic resin, the char, which is supposed mainly formed by the phenanthrene and pyrene, and at the end the gases. The parameters for this last equations are which one of the hydrogen as this is the main gas product, with a presence of the 42% (percentage identified by[5])

These equations are then implemented into a genetic algorithm which has been developed on purpose. Different objective functions has been

used. The best ones are the char optimization and the char maximization together with the minimization of the process time and temperature.

From this theoretical study of pyrolysis [11] it was found that:

- There is a minimum temperature at which all the char is formed.
- Many temperature and time combinations allow to reach the final optimum parameters. The optimum solutions are more than one and this allows to be free to adapt the process conditions to the requirements of each case (such as the maximum temperature reachable for a specific oven, or the minimum time of production) . However these parameters have to be chosen carefully because also final mechanical properties change with the temperature and duration conditions.
- After few hours it is possible to obtain char only.

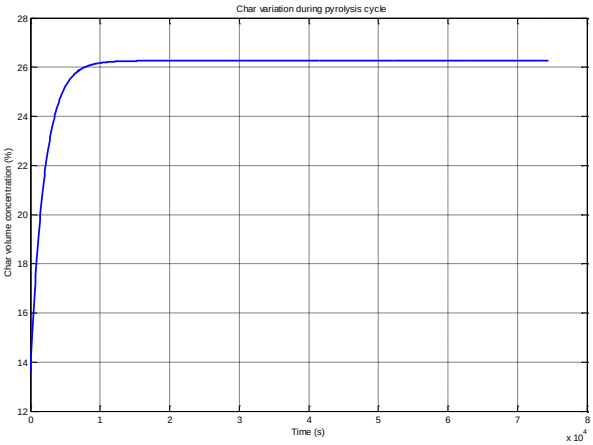


Figure 2 Example of char variation during pyrolysis cycle.

When all the resin is transformed into char, the final char percentage depends on the fiber concentration: the more fibers, the less char and gas. If fiber percentage is from 30 to 60% , the final char percentage ranges respectively from 40 to 20% and the gas percentage from 25 to 15% .

6 Experimental analysis

As explained in 2, some experiments has been performed in order to verify the quality of the mathematical model and the quality of the optimization.
At first only the first pyrolysis cycle was studied. The parameters of the pyrolysis cycle are respectively:

- 6 hours and 34 minutes; 800°C.
- 7 hours and 16 minutes 794°C.

A chemical investigation with a thermo gravimetric analysis (TGA) was done on the samples. TGA is a technique that measures the mass of a sample while it is heated in this case. At the end it shows the mass loss of volatile components (moisture, monomers, solvents), polymer decomposition, combustion of carbon black and final residues (ash, filler, fibers). The method allows to study the decomposition of products and materials and to draw conclusions about their individual constituents. In particular, it is possible to know if some resin content remained after the pyrolysis cycles.
The two plates and a third one non pyrolyzed were studied and compared.

6.1 TGA results

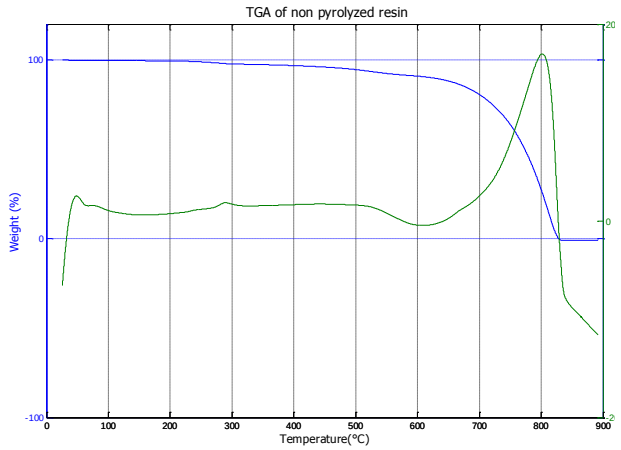


Figure 3 TGA results of non pyrolyzed resin.

Analyzing the TGA diagram (green line, Fig. 3) we can see that there is a first endothermic zone from about 80°C to 278°C. The heat flow begins from 1.547 W/g and decreases until 0.6039 W/g before increasing again. In this zone there is an evaporation of water.
A low peak appear at 288°C, at this temperature there should be a crystallization of the resin and not a decomposition because the peak is not high as the second peak. From theory we know that the first chemical bonds re-arrangements of phenolic resin begins at 300°C, so the results are consistent with the theory.
The second peak is at 442°C, here the decomposition is taking place.
At 600°C there is a flex which correspond to a minimum of the DSC line (blue line). This means that there is certainly a decomposition here.
From the diagram it is possible to identify the three pyrolysis region of the resin:

- From 300°C to 442°C bonds break themselves.
- From 442°C to 500°C where bonds continue to break
- Above 500°C there is a decomposition and new bonds are formed, so there is a release of heat.

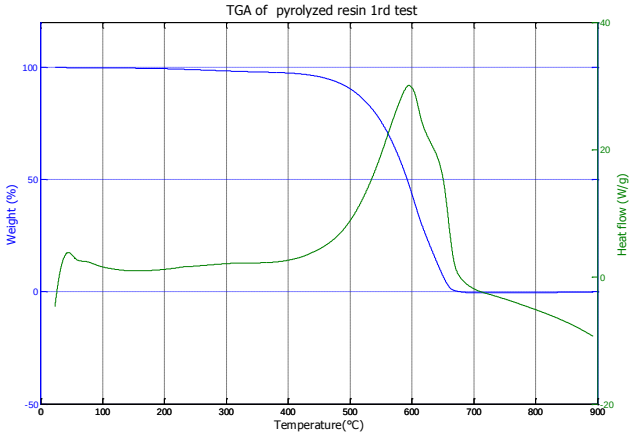


Figure 4 TGA of first experiment.

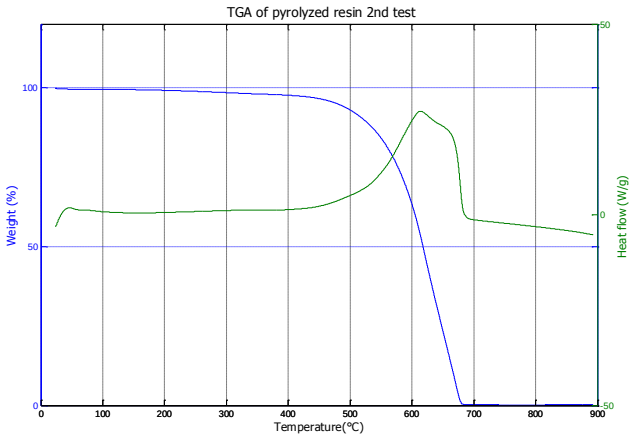


Figure 5 TGA of second experiment.

The diagram trends of the pyrolyzed resins are quite similar. They are very different from the non pyrolyzed resin. So surely there is no resin residue in these two samples.
The highest peak shows two peaks inside, this is because there are two different carbonaceous structures which are decomposing. One of the structures is the carbon fiber, in fact it decomposes at temperatures below 650°C in air.

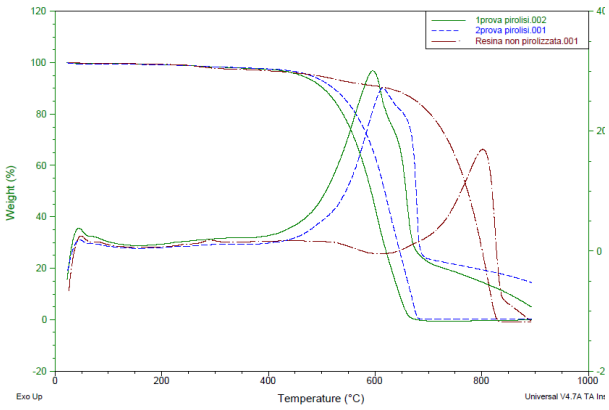


Figure 6 Comparison between the two experiments and the non pyrolyzed plate (red line).

The TGA analysis shows that there is a clear different behavior of the pyrolyzed samples from the non pyrolyzed resin, so it is possible to say that the mathematical model can represent the physical phenomena.

7 Conclusions

A mathematical model of the pyrolysis phase of the manufacturing process has been framed. A genetic algorithm has been developed in order to study and optimize the pyrolysis process model. The algorithm calculates the optimal process parameters as temperature and time.
A strict dependence of these two parameters has been observed by the results.
In order to verify the mathematical model, two experiments have been performed and a TGA analysis was done on the samples.
From the TGA Analysis it is possible to have the confirmation of the mathematical model used for the pyrolysis process. In fact there is no a resin residue after the pyrolysis cycle. Furthermore the pyrolysis temperature used in the two experiments is not high. There is an effective optimization of the process.

Acknowledgements

Thanks to Momentive Specialty Chemicals Italia S.p.A. and to Advanced Composites Group for the material supply.

A special thanks to Prof. Maria Laura Santarelli who executed TGA tests and helped in the reading of the diagrams.

References

[1] J.Cleland, F. Iannetti, "Thermal protection system of the space shuttle", NASA Contract Report 4227, June 1989
[2] G. Savage, "*Carbon-carbon composites*", Chapman & Hall, London, 1993
[3] G.S.Learmonth, D.P.Searle, "*Thermal degradation of phenolic resin. V*", Journal of applied science, Volume 18, Pages 437-443, 1969
[4] Hugh O. Pierson, "*Handbook of carbon, graphite, diamond and fullerenes*", Noyes publications, New Jersey, 1993
[5] K. A. Trick and E.Saliba ," Mechanisms of the pyrolysis of phenolic resin in a carbon/phenolic composite", *Carbon*, Volume 33, Issue 11, Pages 1509-1515, 1995
[6] K.A. Trick, T.E. Saliba and S.S. Sandhu,"A kinetic model of the pyrolysis of phenolic resin in a carbon/phenolic composite", *Carbon*, Volume 35, Issue 3, Pages 393-401, 1997
[7] YU.I.Dimitrienko," Modelling of carbon carbon composite manufacturing processes", Composites, Part A, edited by Elsevier, pages 221-230, 1999
[8] Jurij I.Dimitrienko," *Thermomechanics of composites under high temperatures*", Kluwer Academic Publishers, Dordrecht (Holland), 1999
[9] H.H. Kuo J.H. Chern Lin C.P. Ju, "Effect of carbonization rate on the properties of a PAN/phenolic-based carbon/carbon composite", *Carbon*, Volume 43, pages 229-239, 2005
[10] H.H. Kuo J.H. Chern Lin C.P. Ju, "Tribological behavior of fast-carbonized PAN/phenolic-based carbon/carbon composite and method for improving same", *Wear* edited by Elsevier, Issue 258, Pages 1555-1561, 2005
[11] Marta Albano, Plinio Coluzzi, Davide Micheli, Mario Marchetti, " Carbon-Carbon pyrolysis mathematical modelization and process optimization", Compositi e Nanotecnologie, Volume 6, n°1, February 2011

Hot Structure Sandwich panel for advanced structure assembly (ASA) project

P. Coluzzi, F. Agostinelli, C. Vassalli, M. Albano, M. Marchetti
Dipartimento di Ingegneria Astronautica, Elettrica ed Energetica
Sapienza Università di Roma,
Via Salaria 851, 00138 Roma, Italy

e-mail: plinio.coluzzi@uniroma1.it, web page: http://www.saslab.eu

Keywords: *Thermal Protection System, Re-entry system, IXV, RLV, CMC*

Abstract

The space transportation segment is now involved in the research and development of new technologies for the next generation RLVs (Reusable Launch Vehicles) capable to guarantee a safer, cheaper and continuosly space transportation services.

One of the most critical part on such vehicles is represented by Thermal Protection System (TPS).

The goal of this work is to present a new design for the IXV (Intermediate eXperimental Vehicle) TPS subsystem that is able to provide both the structural and the insulating function in the same component.

This was possible thanks to the specific configuration of the sandwich structure that is composed by three layers completely customizable to the specific mission profile.

The core of the sandwich is an high temperature resistant carbon foam as heat absorber, while the outer surface sheets are made on Ceramic Matrix Composite (CMC).

1 Introduction

Ceramic Matrix Composites (CMCs) composite are considered the primary materials for hot structures of the present and future reentry vehicles.

Carbon laminate composites based on the impregnation and/or infiltration process have proven their suitability under extreme thermo-mechanical environments in different structural parts like nose caps, nozzle jet vanes and engine flaps.

A step forward has been done within the last years to improve the manufacturing technology and to tailor the properties of the materials.

The use of high conductivity carbon fibers on CMCs guarantees high stiffness but the thermal conductivity still remains high.

The scope of this work is to exploit the thermomechanical properties of CMCs combined with the hig thermal insulation of the carbon foam.

The other goal of this paper is the validation of the sandwich configuration and the demonstration that CMCs are efficient cost materials for the sandwich panel with carbon foam core.

The TPS sandwich obtained will be a part of the primary windward surface structure of the IXV reentry vehicle [1].

1.1 TPS Concept

The main objective of the thermal protection system is to keep as much as possible low the internal temperature of the structure in order to guarantee the structural integrity of the vehicle without compromising the aerodynamic performances.

These duties can be fulfilled customizing the sandwich configuration to the specific reentry profile.

The sandwich structure solution is composed by a carbon foam core with two external skins made on CMCs. This configuration was chosen with the needs to combine the structural and thermal functions of the TPS.

The thermal functionalities has suggested to positioning the carbon foam between CMCs laminates to increase the thermal shielding, while the external skins provide the structural support to the entire configuration.

A coating was performed on the outer sheet in order to increase the thermal emissivity and to provide the oxidation protection from the environment. A classical SiC coating was chosen because of its well proven behaviour on such conditions.

1.2 Core and Face sheet Materials

The core of the TPS is a carbon foam arranged in a reticulated open cells micrustructure.

Carbon foam is recognized as having great potential as a component within hybrid thermal protection systems for low angles reentry vehicles. In this concept, the carbon foam supports CMC surfaces by providing selectable insulating properties. Nevertheless, the thickness of the foam can be set according to the reentry trajectory.

The final solution selected is a good compromise between high mechanical properties, low weight and relatively low conductivity. The foam core was selected in order to guarantee the lowest conductivity over the structural performances.

A test campaign was performed in order to characterize the physical and mechanical properties of CMCs. In this paper only the test results are showed, while the detailed description of the tests is argument of a dedicated paper.

The thermal properties investigations were conducted on a 2mm thick sample using a Linseis Laser Flash thermal constants analyzer model TC-700 (ASTM C 1470) . The values are referred to the through-plane direction. The test results are reported in Table 1.

The thermal conductivity on through-plane direction is much lower than that on-plane direction. This is due to the fact that for the two dimensional CMC composites, many of the physical properties are dominated by the carbon fiber properties, by the fraction volume and architecture.

One of the main property for a good design of the sandwich structure is the Coefficient of Thermal Expansion (CTE). This measurement was performed using a Linseis L75 dilatometer (ASTM E 228). Graphs for in plane and crosswise measurements are respectively given in Fig. 1 and Fig. 2. The same method was used for both the CMC laminate and the carbon foam. Thermal conductivity and specific heat are reported in Table 1, while the CTE graph is given in Fig. 3.

Furthermore, emissivity measurements up to 1600°C of the SiC coating were taken by means of a SpectroPyrometer. The average emissivity measured was 0,85. This value was implemented in the thermal simulations.

The mechanical characterization of CMC was performed according to the ASTM standards for ceramic composite materials. The tests includes tensile strength, flexural strength, elastic modulus and Poisson ratio. A summary for the CMC is reported in Table 2.

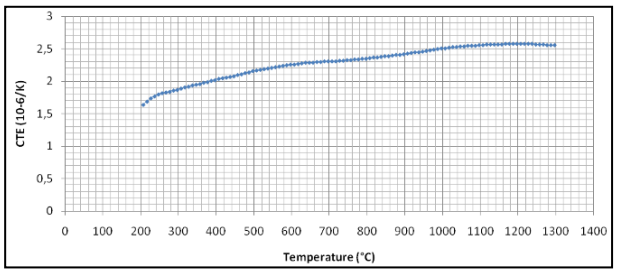


Fig. 1: CMC in plane CTE

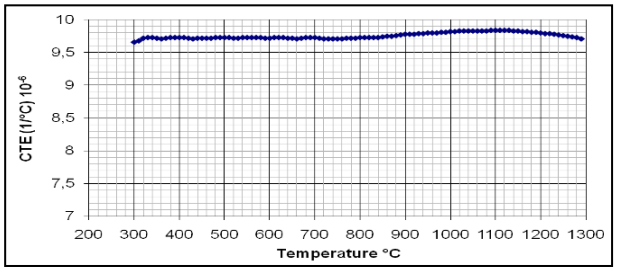


Fig. 2: CMC crosswise CTE

Temperature (°C)	Specific Heat ($Jkg^{-1}°C^{-1}$)	Thermal Conductivity ($Wm^{-1}°C^{-1}$)
20	677	0.022
300	1147	0.064
600	1415	0.13
900	1491	0.221
1200	1590	0.343

Table 1: Carbon Foam Thermal properties

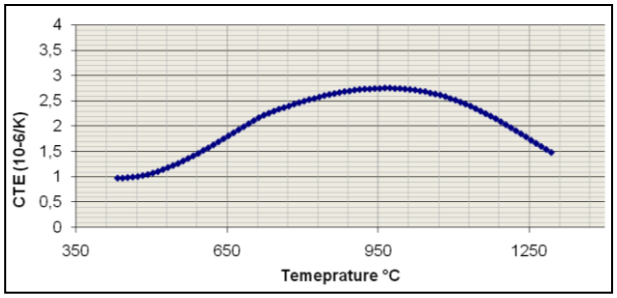


Fig. 3: CTE of Carbon foam

Tensile strength (MPa)	Elastic Modulus (GPa)	Flexural (MPa)	Flexural Modulus (GPa)	Shear (MPa)	Poisson Ratio
125	100	152	50	4.7	0.36

Table 2: CMC mechanical properties

2 Design and FEM analysis

For a given material, the heat-shielding and thermo-structural overall performance of the sandwich TPS panel depend on the thickness of the face sheets and on the thickness and density of the foam core.

A screening of several sandwich configurations has been investigated starting from a single CMC laminate. Increasing the foam core thickness, replacing that of the CMC, gave rise to several possible sandwich configurations.

A series of FEM analysis on the sandwich configuration were performed using the heat loads profiles of the mission.

In Fig. 4 are reported the heat loads profiles of the mission for four different area of the TPS panel.

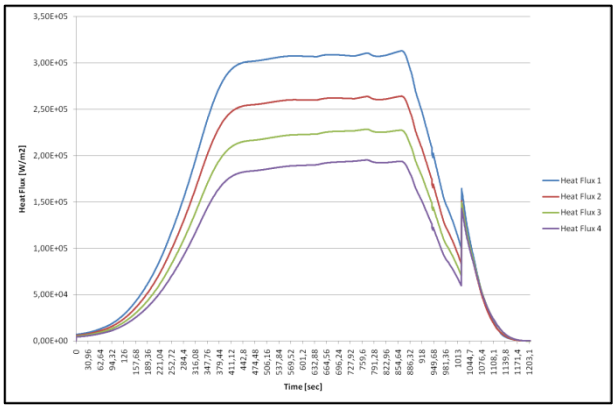


Fig. 4: TPS Heat Flux

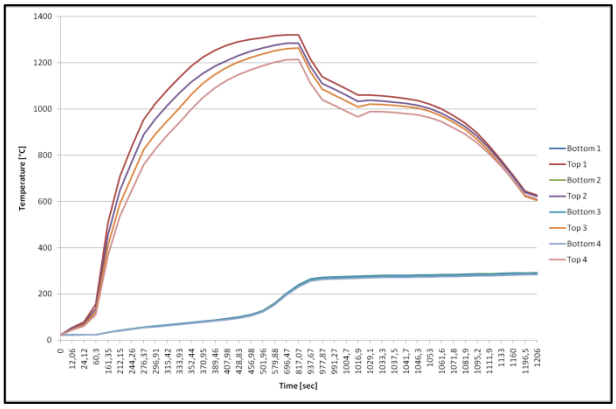


Fig. 5: Time history of the TPS temperature

The time-temperature trend is showed in Fig. 5.

Fig. 6 shows an overview of the TPS at the maximum temperature reached (804 sec).

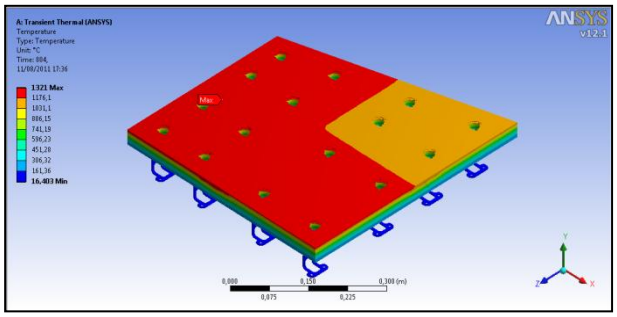


Fig. 6: TPS Temperature

As expected, the highest temperature is located in the top face sheet in correspondence of the maximum heat flux (1346°C) exhibited at end of the "plateau". On the other hand, the lowest is located in the bottom face sheet in correspondence of the minimum heat flux (291°C) reached at the end of the reentry mission.

This behaviour is mainly due to the higher heat capacity of the foam and to the direction of the heat flux towards both the face sheets during the cooling phase. Moreover, the top face sheet cooled more quickly than the foam thanks to the higher conductivity and to the higher emissivity due to the presence of the coating.

So, the TPS sandwich is exposed to an high thermal gradients between the skins with relatively lower temperatures on the inner surface allowing the protection of the substructures. The higher is the thermal gradient through the thickness of the sandwich the better is the protection of the substructure and the efficiency of the TPS.

The thermal gradient through the thickness will induce bending effects on the structure, forcing the TPS panel to bulge outward influencing the aerodynamic profile around the shape.

Therefore, a thermo-mechanical analysis [5] has been performed at the highest thermal gradient exhibited (t=753 sec). The temperature nodal values has been imported into the structural model and the nonlinear structural analysis was run.

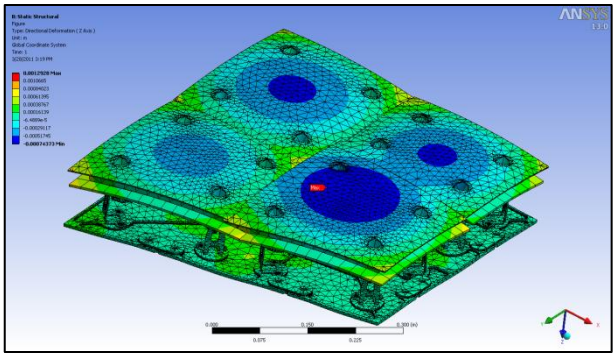


Fig. 7: TPS Directional Deformation

Fig. 7 shows the displacements results. The values showed goes from 0,7 to 1,2 mm out of the plane with no significant effects on the aerodynamic profile.

3 Conclucion

A procedure for the optimization process has been performed in order to design an high efficiency TPS sandwich and a low weight structure.

The geometrical configuration of the sandwich panel is in accord with the test article requirements and a preliminary design is obtained.

A non-linear and thermo-mechanical analysis were performed using a FEM code.

The results shows an high thermal gradients through the thickness on the TPS structure during re-entry phase.

The thermo-mechanical analysis was performed at the maximum gradient temperature obtained by the thermal mission profile.

The maximun geometrical variations recorded is in the order of 1 mm out of the plane with no relevance for the aerodynamic profile distribution.

Acknowledgments

This research project has been sponsored by the Italian Space Agency (ASI) inside the

Advanced Structure Assembly (ASA) program. Prime contractor of the program is Thales Alenia Space Italy, Turin.

References

[1] Shideler, J. L., G. L. Webb, and C. M. Pittman, "Verification Tests of Durable Thermal Protection System Concepts," *J. Spacecraft and Rockets*, vol. 22, no. 6, 1985, pp. 598–604.

[2] Ko, William L., Robert D. Quinn, and Leslie Gong, Finite-Element Reentry Heat Transfer Analysis of Space Shuttle Orbiter, NASA TP-2657, 1986.

[3] John D. Buckley and Dan D. Edie, *"Carbon-Carbon Materials and Composites"*, ISBN: 0815513240, NOYES PUBLICATIONS 1993, pp. 211-212, pp. 251-251

[4] Ko, William L., *"Thermostructural Behavior of a Hypersonic Aircraft Sandwich Panel Subjected to Heating on One Side"*, NASA TM-4769, 1997.

[5] Noor, Ahmed K., W. Scott Burton, and Charles W. Bert, *"Computational Models for Sandwich Panels and Shells*", Applied Mechanics Reviews, vol. 49, no. 3, Mar. 1996, available as ASME Reprint no. 186.

[6] Steven L. Donaldson, Daniel B. Miracle, ASM Handbook Composites Volume 21, Chap. Post-Processing and assembly of Ceramic Matrix Composites, ASM Intl., ISBN 0871707039

[7] William L. Ko, Heat Shielding Characteristics and Thermostructural Performance of a Superalloy Honeycomb Sandwich Thermal Protection System (TPS), NASA/TP-2004-212024

[8] David E. Glass, *"Ceramic Matrix Composite (CMC) Thermal Protection Systems (TPS) and Hot Structures for Hypersonic Vehicles"* 15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference, 2008

[9] B. Laub, M. J. Wright, E. Venkatapathy *"Thermal Protection System (TPS) Design"* 6th International Planetary Workshop, Atlanta Georgia, 21-22 June 2008

CEAS 2011

Testing and Certification

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Academic Flight Testing Capabilities Enabling Ultralight Aircraft Certification

Lorenzo Trainelli, Alberto Rolando, Cesare Cardani
Dipartimento di Ingegneria Aerospaziale, Politecnico di Milano, Italy

Keywords: *ULM, FLIGHT TESTING, CERTIFICATION, VLA, LSA*

Abstract

We describe a ground-breaking activity, where the educational needs of the Department of Aerospace Engineering at the Politecnico di Milano (DIA-PoliMi) have been successfully matched with the industrial requirements of Ing. Nando Groppo Srl, a leading Italian Ultra-Light Machine (ULM) manufacturer. The latter was seeking support in the process of type certification of its "Trial" model, a new three-axis control ULM. DIA-PoliMi contributed its capability in performing comprehensive, reliable and cost-effective flight testing campaigns dedicated to small aircraft, built upon the experience accumulated by the extensive use of its in-house designed "Mnemosine" FTI system through over 100 flight missions in support to the "Flight Testing" graduate course. The "Trial" certification procedure in compliance to the demanding German airworthiness regulations LTF-UL has been successfully finalized in January, 2011, demonstrating the availability of sustainable, simple-to-operate, dependable, and fully safe flight testing tools for the global market of small aircraft.

1 Introduction

In Italy, as well as in many other countries, airworthiness regulations for very small aircraft below the CS-VLA [1] weight specifications have been very limited, when not nearly absent, in the past years. This state of affairs significantly impacts ULM (UltraLight Machine) – i.e. airplanes with a maximum allowable take-off weight of 450 kg – operations, since manufacturers at large typically do not undertake well-established, rigorous flight testing programs with all the related investments, let alone a complete certification procedure.

Inevitably, when flight testing is discarded or loosely performed, crucial safety issues arise. Furthermore, unqualified ULMs may encounter serious problems in approaching national markets where airworthiness standards for these machines are enforced. Of course, given that flight testing tools and teams dedicated to ULMs are lacking, almost no manufacturer is the position to afford the high costs typical of flight testing activities as normally performed for higher category aircraft.

Reckoning with this situation, the Department of Aerospace Engineering at the Politecnico di Milano (hereafter referred to as 'DIA-PoliMi') resolved to capitalize its long-term involvement in flight testing for manned and unmanned aircraft, including ULMs. In recent years DIA-PoliMi has developed a specific line of education and applied research focused on this particular topic. In particular, from 2004 onwards, the graduate course of "Sperimentazione in Volo" ("Flight Testing"),

has been regularly held in the second semester of the final year of the Master Degree in Aeronautical Engineering [2,3]. Teaching duties have been entrusted to top experts from world-class aeronautical industry, in co-operation with DIA-PoliMi's faculty members. Naturally, this circumstance generates a deeper understanding of the industrial viewpoint on flight testing, and a considerable transfer of specialized knowledge to the academic counterpart.

Parallel to curricular education, DIA-PoliMi is engaged in the ongoing development of the FTI (Flight Test Instrumentation) system "Mnemosine". This started as the core of a PhD project in 2005 [4] and has resulted in a proprietary, integrated FTI suite tailored on light aircraft, such as ULMs and General Aviation fixed- and rotary-wing vehicles, capable of supporting a wide range of flight testing activities [5-7]. The system was fully designed, developed and implemented in-house and has now reached a considerable degree of maturity, having been regularly employed in the laboratories of the "Flight Testing" course totaling over 100 actual flight missions.

In the past three years, DIA-PoliMi enjoyed the collaboration with the Club Astra flying school and Ing. Nando Groppo Srl, a leading Italian ULM manufacturer. These companies provided specialized support to didactic activities, such as the flight missions of the "Flight Testing" course. Furthermore, Ing. Nando Groppo Srl was seeking assistance in the process of type certification of its recent "Trial" ULM model according to German airworthiness regulations, in an effort to expand its market on a European scale. As a result, DIA-PoliMi provided the "Mnemosine" FTI system, assuring the related management and maintenance support, and contributed to flight test mission planning, flight test operations, data processing and analysis, eventually leading to the successful completion of the certification procedure in January, 2011.

The present paper describes the complex certification project outlined above and its fundamental constitutive elements, illustrating the accomplishment of a solid, reliable expertise in flight testing for small aircraft at fully sustainable global costs and extremely mild impact on the airframe, its performances, and Human-Machine Interface issues. Further details can be found on Refs. [8,9].

2 DIA-PoliMi's FTI system

The FTI system "Mnemosine" is a low cost, simple-to-operate onboard data acquisition system designed to support extensive flight testing campaigns not only for didactic and research purposes, but also in view of possible applications in the small aircraft global market. "Mnemosine" (the goddess of memory, mother of the Muses) is based on a federated architecture in which the system is divided in a number of autonomous nodes (named after the Muses, the goddesses presiding over arts and sciences), detailed in Table 1.

Table. 1 FTI node functions.

Node name	Node task
Eutherpe	Flight control position acquisition
Klios	Data logging
Melete	System power management and distribution
Polimnia	GPS data acquisition and system time generation and distribution
Talia	Engine data acquisition
Terpsicore	Inertial measurement acquisition
Urania	Air data acquisition
Melpomene	Command and display unit
Calliope	Flight control force acquisition

Every single node can operate independently from the others, and is specialized for specific task: it has processing power, memory, power supply and all the signal conditioning/interface resources required to manage the particular sensor(s)/device(s) it manages. All data generated by the modules is then shared over a common communication line: a digital data bus.

This architecture permits to distribute the units across the aircraft placing every module as close as possible to the sensor it manages, thus avoiding to lay down long, noise-sensible analog signal lines. Information is converted to a digital format at node level, processed and transmitted over a robust medium. Software partitioning, fault confinement, as well as a high

degrees of flexibility and modularity are further considerable advantages of the chosen architecture in contrast to centralized solutions.

Table. 2 FTI parameter list.

Parameter	Node	Rate (Hz)
X acceleration	Terpsicore	60
Y acceleration	Terpsicore	60
Z acceleration	Terpsicore	60
Pitch angle	Terpsicore	60
Roll angle	Terpsicore	60
Yaw angle	Terpsicore	60
Pitch rate	Terpsicore	60
Roll rate	Terpsicore	60
Yaw rate	Terpsicore	60
Elevator control position	Eutherpe	10
Aileron control position	Eutherpe	10
Left pedal position	Eutherpe	10
Flap position	Eutherpe	10
Rudder force	Calliope	10
Aileron force	Calliope	10
GPS time of week	Polimnia	4
GPS week no.	Polimnia	4
GPS X position (ECEF)	Polimnia	4
GPS Y position (ECEF)	Polimnia	4
GPS Z position (ECEF)	Polimnia	4
GPS X velocity (ECEF)	Polimnia	4
GPS Y velocity (ECEF)	Polimnia	4
GPS Z velocity (ECEF)	Polimnia	4
GPS fix quality	Polimnia	4
GPS tracked satellites	Polimnia	4
GPS DOP	Polimnia	4
Propeller RPM	Talia	2
Total pressure	Urania	10
Static pressure	Urania	10
Angle of attack	Urania	10
Angle of sideslip	Urania	10
Event (“top”) counter	Melpomene	

All data undergo a twofold process before they finally become available to the user. The first phase is performed on-board, in real time and at node level, when an acquired parameter signal is subjected to suitable conditioning and filtering, transformed into engineering units, time stamped in order to assure correct time synchronization across the various time histories, and stored on non-volatile memory. The second phase is performed off-line, as a post-processing involving all the operations necessary to support data analysis, such as computation of derived parameters or information fusion from different sensors, such as trajectory reconstruction by GPS and inertial data fusion in tight coupling.

The crucial need for reliable and accurate data time-referencing feature is met by application of CAFFE (CAN for Flight test Equipment), a DIA-PoliMI developed variation of CANAerospace, which is a widely used protocol targeted to avionic systems [10]. In addition to the CANAerospace characteristics of robustness, efficiency, lightness, good performances, and large diffusion, CAFFE brings a built-in data time-stamping capability.

Table 2 shows the full array of “Mnemosine” acquired parameters together with the respective node and their acquisition rate, with respect to the configuration implemented for the present activity.

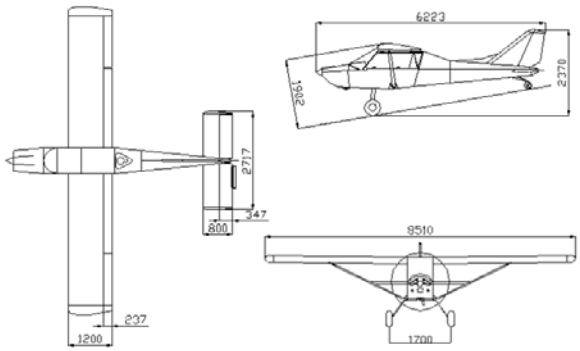


Fig. 1 Nando Groppo’s “Trial”

3 Nando Groppo’s “Trial” model

The “Trial” is the last product designed and manufactured by Ing. Nando Groppo Srl (Fig. 1). It represents a rustic, sturdy albeit very light, reliable airplane suited for training and recreational purposes, such as cross-country flight. It is characterized by a high wing and a taildragger, fixed landing gear. The straight, rectangular wing is endowed with flaps and conventional ailerons, and has been fitted with vortex generators on the leading edge for the full wing span. The tail features a classical

reverse-T configuration, with conventional elevators and rudder. The fuselage is particularly slim, due to the tandem seat configuration. The power plant can be either a Rotax or a Jabiru engine in the 80 HP class

Table. 3 Principal dimensions of the Ing. Nando Groppo Srl "Trial" model.

Wing span	8,51 m
Overall length	6,22 m
Overall height	1,90 m
Wing surface	10,20 m^2
Wing loading	46,3 kg/m^2
Horizontal tail span	2,72 m
Main landing gear width	1,70 m
Cockpit width	0,78 m

Principal dimensions are reported in Table 3. The considered MTOW (Maximum Take-Off Weight) is 472.5 kg, *i.e.* the maximum admissible for compliance to the airworthiness requirements.

Primary flight controls are mechanically actuated, while longitudinal trim and flap controls are electrical. All controls are doubled, one for each seat. Basic cockpit flight instrumentation includes ASI (AirSpeed Indicator), altimeter, VSI (Vertical Speed Indicator), engine RPM (Revolutions Per Minute) indicator, ball-slip indicator, magnetic compass, fuel level indicator, oil temperature gauge, oil pressure gauge, cylinder head temperature indicator.

Table. 4 "Trial" declared performances

Take-off distance (on grass)	300 m
Landing distance (on grass)	250 m
Stalling speed (full flaps)	60 km/h
Max rate of climb	> 4 m/s
Max rate of climb airspeed	100 m/s
Service ceiling	3000 m
Typical cruising speeds	125 ÷ 170 km/h
Maneuvering speed	125 km/h
Load factor envelope	–2 ÷ +4
Never exceed airspeed	198 km/h

Declared representative performances, resumed in Table 4 (all figures refer to MTOW, ISA conditions, and sea level – when significant), place the "Trial" on a par with several Italian and international competitors. Also, handling qualities are very good, making the "Trial" a suitable candidate for flying school activities.

The specific model subject to the flight test campaign was powered by a Jabiru 2200 cc engine with a three-bladed, composite, fixed-pitch propeller. The Jabiru 2200 cc is a 4-cylinder, 4-stroke, horizontally opposed, air cooled engine capable of delivering a maximum power of 85 HP at 3300 RPM. For safety reasons, the aircraft features also a ballistic recovery system (*i.e.* a rocket-enabled parachute), a mandatory device in some countries.

4 Certification basis

The certification of the "Trial" has been carried out according to the "LTF-UL 2003 airworthiness requirements for three axis standard control Ultra Light Aircraft" [11]. These norms, issued by DULV (Deutsche UltraLeichtflugVerband) and adopted by the German DFS (Deutsche FlugSicherung) authority, encompass typical light aircraft airworthiness requirements, and closely match those included in the CS-VLA regulations [1], that apply to aircraft of a larger size than ULMs. The LTF-UL requirements involve different means of compliance, including calculations, ground tests and, of course, a number of flight tests. Through analysis of the fight test requirements, a test matrix amounting to 397 test points was obtained, spanning

- stalling speed determination: wing-level idle and power-on, turning (applicable paragraphs: LTF-UL 49, 201, 203);
- take-off (LTF-UL 51), landing and balked landing (CS-VLA 75, 77);
- climb performance (LTF-UL 65);
- longitudinal and lateral-directional control, elevator control forces in maneuvers, trim speed range (LTF-UL 143, 145, 147, 155, 161);
- longitudinal and lateral-directional static stability (LTF-UL 171, 173, 177);
- longitudinal and lateral-directional dynamic stability (LTF-UL 181);

- power plant cooling (LTF-UL 1041);
- ASI calibration (LTF-UL 1323).

Further analysis, aimed to reduce the number of test points without affecting safety and reliability, led to a drastic reduction by more than 50%. In fact, by considering only the most demanding conditions in terms of weight (MTOW), CG position (max forward or max aft, depending on the test item) and flap deflection (cruise or landing setting, depending on the test item), a considerably slimmer campaign made out of 183 test points was prepared and approved by DULV. It must be noted that the higher impact of this optimization procedure occurred on trim, static and dynamic stability flight trials, while items such as stalling speed determination and stick force determination were basically unaffected.

5 Flight test campaign

The flight test campaign involved a "Trial" with the "Mnemosine" FTI system permanently installed on board. The various sensors and transducers as per the previous description have been mounted with minimal impact on the internal ergonomics, in close collaboration with the aircraft designer and manufacturer and the pilot to guarantee maximum safety.



Fig. 2 The FTI Pitot boom mounted under the right wing of the "Trial"

Several flight missions were performed with an additional ADS (Air Data System) employing a Rosemount Pitot boom (Fig. 2), providing static pressure, dynamic pressure, angle of attack and angle of sideslip measurements. This massive (and expensive) device has been made available to DIA-PoliMi courtesy of Alenia Aermacchi, but can be substituted by several different models of much lighter and cheaper Pitot booms available on the market.



Fig. 3 Elements of the "Mnemosine" FTI system installed on-board the "Trial"

The impact of the FTI system on the vehicle was indeed very low: the total weight of the "Mnemosine" system, Pitot boom excluded, stands below 5 kg, and the maximum power consumption is less than 9 W. Figure 3 shows some elements of the system installed in the vicinity of the front seat: the biaxial Futek cell knob mounted on the stick head; the AHRS (Attitude and Heading Reference System) placed between the pedals; a rack stacking several acquisition nodes fixed to the frontal bulkhead panel; the data recording unit secured on a diagonal stiffener of the left panel.

The necessary flight test missions have been designed gathering multiple test points, based on similarity in altitude, speed range and CG position. Special attention has been devoted to contain pilot workload, with missions not exceeding one hour flight time, with limited altitude changes.

The resulting campaign consisted of a few dozens flight test missions, accomplished in slightly more than a week work (albeit distributed over a period of a few weeks). Repetition of test points has been contained to a few occurrences, such as for ASI calibration, when a few different locations for the static probe were investigated after the initial placement showed poor performances.

In many missions, the core activities have been preceded and followed by wind speed determination maneuvers, using the GPS cloverleaf method [12]. Also, most of the missions included take-off and landing performance demonstration, allowing for reliable performance estimation by averaging across multiple tests.

Adopted flight test procedures included most of the typical techniques currently employed in testing aircraft in the CS-VLA and CS-23 categories. For example, air data system calibration has been carried out following the NTPS (National Test Pilot School) GPS Horseshoe Heading with averaging method [13]; stall speeds have been determined by averaging across several standard 1-*g* quasi-steady decelerations; take-off and landing were considered in terms of performances, controllability and pilot workload; sawtooth climbs have been carried out to assess SEP (Specific Excess Power) characteristics; wind-up turns have been used to determine 'stick force per g' measures; longitudinal stick and pedal doublets have been used to assess dynamic stability.

Flight data were usually post-processed on a laptop PC in preliminary form immediately after the mission completion, within 10 to 15 minutes, to support debriefing and preliminary assessment. In a few cases, on-the-fly data analysis has shown the need to repeat some test points, or to improve the flight testing technique, suggesting practical hints to the pilot. The system displayed a remarkable level of overall reliability. In fact, data loss occurred in only one mission, due to unnoticed physical disconnection of the data bus on the "Klios" storage node.

A special characteristic of the "Trial" certification flight test process was the synergy with the concurrent flight activities of the "Flight Testing" course scheduled in June, 2010, close to the end of the classes. This allowed to perform didactic missions valuable also for certification purposes. Most of the missions were thus performed with a two-person test crew, the second being a graduate student specifically trained, within the scope of the present activities, to act as a FTE. In double crew missions, the test pilot workload was significantly reduced with respect to single crew conditions, and enhanced situation awareness was achieved.

6 Results

The outcome of the flight test campaign was globally successful. The "Mnemosine" FTI system installed on the "Trial" recorded the time histories of some 30 parameters during all the missions performed, amounting to a considerable number of flight hours. Data were preliminary verified after flight completion and then thoroughly post-processed off-line.

A 'top' (event marker) counter was used to identify single trim points and maneuver durations, easing the task of isolating specific sections from the full data streams. The "top" counter event is included in the acquired parameter list, and as such is time-stamped with the CAFFE procedure and evaluated during the off-line post-processing.

All the certification requirements that needed in-flight demonstration have been assessed, showing compliance with the relevant LTF-UL standards.

As an example of the quality of the performed analyses, Fig. 4 shows the time histories of CAS (Calibrated AirSpeed, diamond mark), pressure altitude (square mark), pitch angle (triangle mark), and roll angle (circle mark) during a stall test in landing configuration, with maximum forward CG

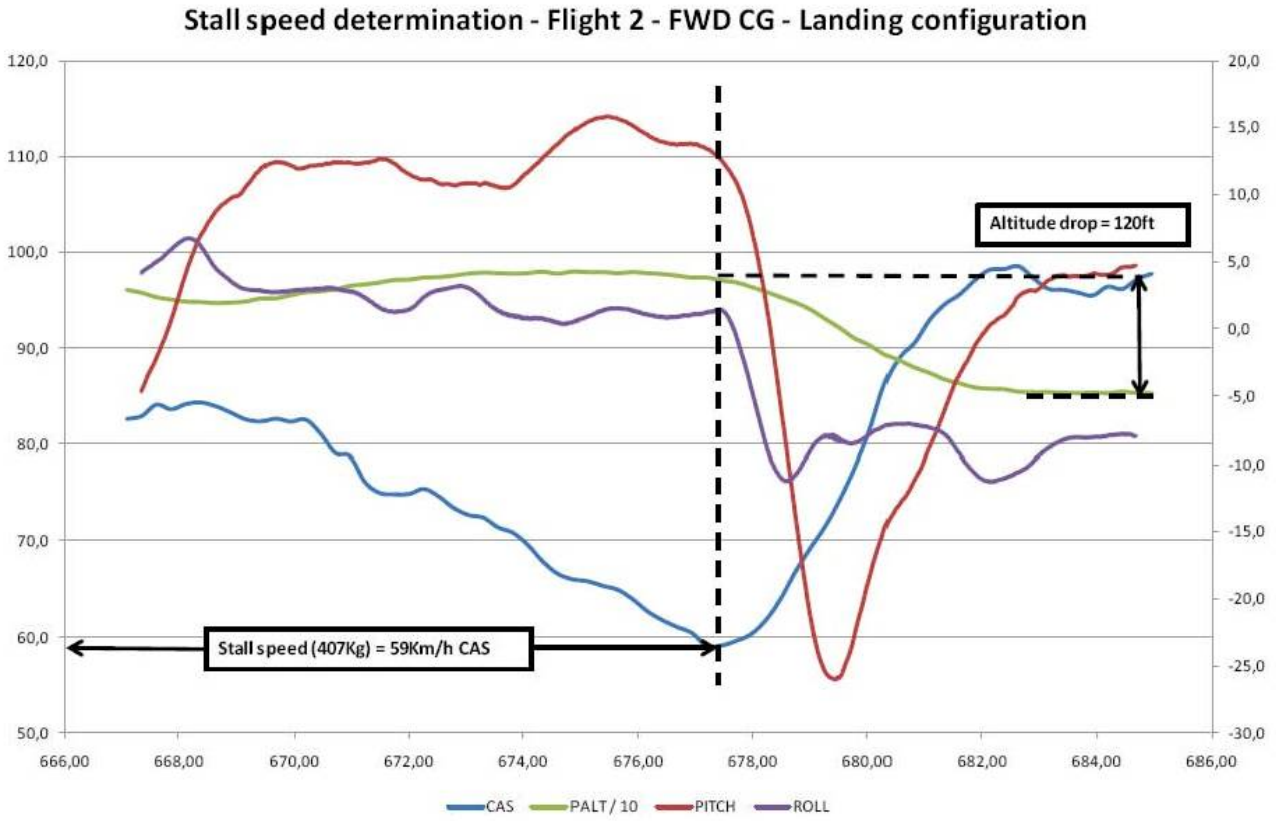


Fig. 4 Time histories during a stall test maneuver (courtesy of Ing. Nando Groppo Srl)

position. CAS values in km/h are given on the left axis, as well as pressure altitude values in feet divided by a factor 10, while pitch and roll angle values in degrees are given on the right axis. Pitch break, roll behavior, altitude loss, entry deceleration and other significant characteristics can be easily observed. After checking against several analogous results, it appears that the behavior of the airplane is symmetric and fairly benign, and elevator control is sufficient. Also, adequate stall warning in the form of slight buffeting has been observed.

Another example is shown in Fig. 5, where height AGL (Above Ground Level, square mark), CAS (diamond mark), and normal load factor (circle mark) time histories are plotted for a typical take-off maneuver. Height values in feet and CAS values in km/h are both given on the left axis, while normal load factor values are given on the right axis. In these trials, the take-off procedures employed were both the recommended "three-point" technique, leaving the aircraft lift off the ground spontaneously at the appropriate speed, and the conventional technique, where the pilot lifts up the tail during the ground course before rotation to lift-off. In both cases the aircraft behavior was judged excellent and fully compliant with the relevant requirements. Further examples dealing with static and dynamic stability characteristics can be found in Ref. [8].

7 Concluding remarks

In the present paper, we strived to illustrate the main points concerning an original approach to flight testing dedicated to small aircraft developed by DIA-PoliMi. Within this safety-inspired, affordable and reliable approach, academic expertise in flight testing tools and procedures was made available to a private company to accomplish a complete type certification process according to rigorous airworthiness requirements. The related flight test campaign involved an extended array of

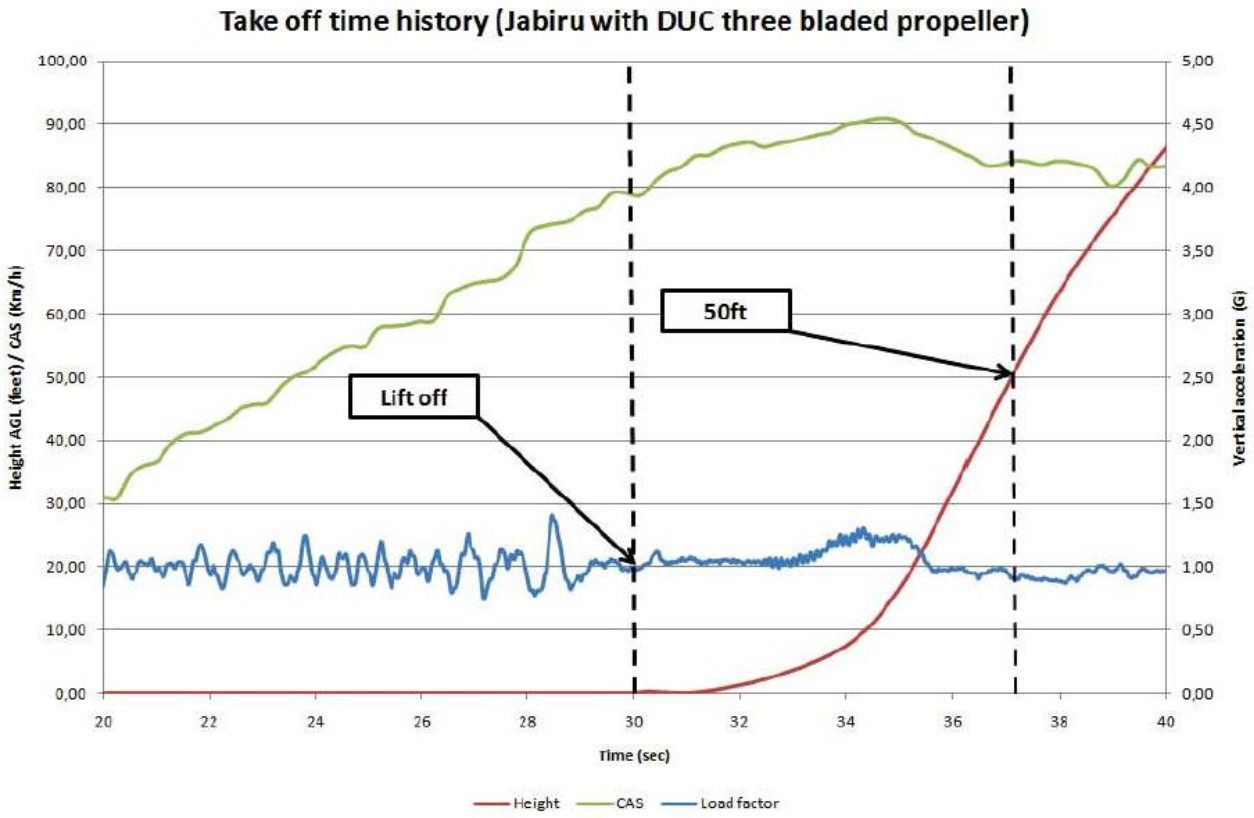


Fig. 5 Time histories during a take-off test maneuver (courtesy of Ing. Nando Groppo Srl)

missions to demonstrate performance, stability, control of the aircraft and of some of its main subsystems (such ad ADS, powerplant, flight controls, etc.).

DIA-PoliMi provided substantial support at all stages, from flight planning to flight operations, data processing and analysis. In addition, academic education effectiveness was enhanced by the synergy established between the certification activities and the "Flight testing" graduate course, having several graduate students directly involved in the flight test missions, experiencing hands-on in the role of Flight Test Engineers.

All the flight missions were performed recording relevant flight data through the "Mnemosine" system, a unique FTI suite designed and developed by DIA-PoliMi. The "Mnemosine" FTI suite has shown full compliance with its design requirements as well as the certification needs in terms of overall reliability, availability, accuracy, and safety. Concerning easiness of operation it must be noted that DIA-PoliMI staff and students, as well as Ing. Nando Groppo staff, interacted promptly and effectively with the instrumentation, in spite of their short training. The system fully supported the analysis necessary to demonstrate compliance with severe airworthiness standards, with actually sustainable global costs. Indeed, a single "Mnemosine" set in the considered configuration has a raw cost below 10 k€, and non recurring production costs stand below 5 k€. This represents a striking depart from the considerable investments traditionally related to flight testing, which have led to the current situation where only major companies (typically involved in bigger aircraft classes than ULMs) can afford this activity on a regular, consistent basis.

Furthermore, the system displays a considerable growth potential in term of number of acquisition nodes and on-board data processing and display, and can be considered fully validated and basically mature for regular use outside the academic environment. Additional capabilities, including telemetry to

allow for real-time mission monitoring on the ground, are currently under development.

Also, migration to light gliders and other small aircraft is under consideration, in view of further demonstrating the potential benefits of sound and cost-effective flight testing tools and procedures in this particular, highly dynamic and economically significant segment of aeronautics.

Acknowledgements

The authors gratefully acknowledge the co-operative support and expertise of Paolo Chimetto of Alenia Aermacchi, Giovanni Bonaita and Michele Riccobono, both of AgustaWestland.

References

[1] "Certification Specifications for Very Light Aeroplanes CS-VLA", European Aviation Safety Agency (EASA), 2003.

[2] Trainelli L., Rolando A., Cardani C., Folchini A., Chimetto P., Bonaita G., "Experiences in Flight Mechanics education: Getting the students hands on the real thing", *SIMAI Congress 2008*, Roma, Italy, 2008.

[3] Trainelli L., Rolando A., Chimetto P., and Bonaita G., "Education in Flight Testing at the Politecnico di Milano", *20th AIDAA National Congress*, Milano, Italy, 2009.

[4] Rolando A., "Development of an Integrated Flight Test Instrumentation System for Ultra Light Machines", Ph.D. Dissertation, Department of Aerospace Engineering, Politecnico di Milano, Italy, 2008.

[5] Cardani C., Folchini A., and Rolando A., "Mnemosine: a Federated Flight Test Instrumentation System for Sport Aviation Aircraft," *19th AIDAA National Congress*, Forlì, Italy, 2007.

[6] Rolando A., "A Telemetry Node for MNEMOSINE, a Flight Test Instrumentation System for Sport Aviation Aircraft", *28th European Telemetry Conference ETC 2008*, Munich, Germany, 2008.

[7] Cardani C., and Rolando A., "Improving a VLA Dedicated FTI System With Flight Control Force Data Acquisition", *20th AIDAA National Congress*, Milano, Italy, 2009.

[8] Cardani C., Rolando A., Trainelli L., Chimetto P., and Bonaita G., "A Cost-effective Approach to Ultralight Machine Flight Testing – An Academic Perspective", *21st SFTE EC Symposium*, Vergiate, Italy, 2010.

[9] Trainelli L., Rolando A., "Reliable and Cost-effective Flight Testing of Ultralight Aircraft", *Journal of Aircraft*, Vol. 49, No. 8, 2011.

[10] Rolando A., Carpanetti M., "A Time Stamping Method for a CAN Based Distributed Flight Test Data Acquisition System", *30th European Telemetry Conference (ETC 2010)*, Hamburg, Germany, 2010.

[11] "Bekanntmachung von Lufttüchtigkeitsforderungen für Aerodynamisch Gesteuerte Ultraleichtflugzeuge LTF-UL", Deutsche Flugsicherung (DFS), Langen, Germany, 2003.

[12] Lewis, G.V., "A Flight Test Technique Using GPS For Position Error Correction Testing", *Cockpit – Quarterly of the Society of Experimental Test Pilots*, Jan/Feb/Mar, 1997.

[13] Olson, W.M., "Pitot-Static Calibration Using a GPS Multi-Track Method", *SFTE Symposium*, Reno, NV, USA, 1998.

A400M ES - Main Airframe Static Test

Angel Manuel Gago Tripero
Airbus Military, Spain

Keywords: *Full Scale, Static Test, A400M*

Abstract

In this paper is briefly presented the Full Scale Static Test of the A400M aircraft performed at Airbus Military in Getafe, Madrid, Spain. It is shown the main test objectives, how the aircraft was supported and loaded, the test cases selection processes and the test sequence. Finally the test systems, load control system and data acquisition system, are described together with their special functions specially developed for this of test.

1 General Introduction

The structural design of the Aero-Structures requires a continuous evaluation to determine if the original requirements have been satisfied. Some of the most important tests, but not the only ones in this process are the static and fatigue tests. It is started at coupon level in order to obtain the mechanical characteristics of the selected materials followed by the elements, details, sub-components, components and finally the complete Aircraft or Full Scale.

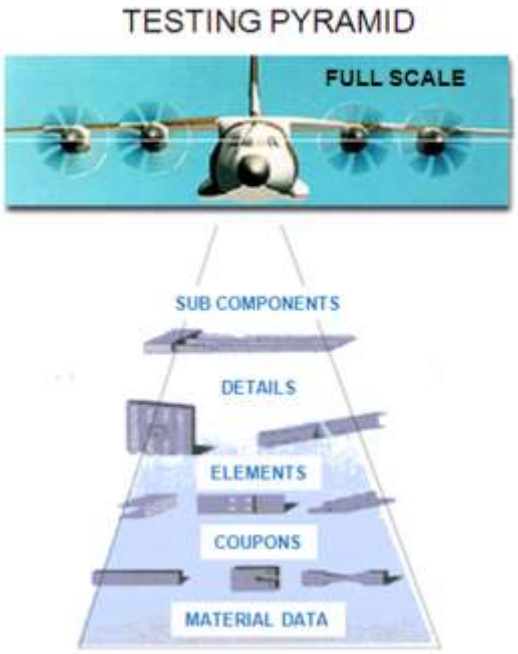


Fig. 1 Testing Pyramid

The Full Scale, due to his size and the organizations involved in its development generally is one of the most important investments in the overall test plan. In the case of the A400M test program, former EADS CASA, now Airbus Military was designed for preparing and performing all the Full Scale Static Tests with the highest level of responsibility. Airbus Military was not only a

Testing Laboratory with all the associated obligations, but also a partner for the success of the overall program. With this additional task the Laboratory was proactive to propose solutions, methods etc. to accomplish the certification program.
The Full Scale Test is one of the primary means to demonstrate how successfully the structure meets the structural performance requirements. Furthermore, in this type of test is where the contour conditions of the critical elements are represented in the most real way, especially for the carbon fiber structures.
The structure to be tested is usually one of the first two fully completed manufactured units. The manufacturing as well as the test itself, usually run on a very tight schedule, so if any unexpected structure failure were to occur, this would result in a possible structure redesign or tooling changes in which would compromise costs, penalties or even contract renegotiations. Another aspect to be considered, unusual but possible, is an unexpected test system failure which can cause accidental structure damage. The overall impact can result in a general delay with the program schedule as well so, special care have to be taken especially with the background, means, procedures, certifications and experience of the people involved in the testing Laboratories.
The Full Scale is the test where all the structural components find their real behavior when the overall structure is subjected to any load case. The stresses obtained in such way are the base for the FEM verification and qualification, for that reason is the base for the structure certification by the corresponding authorities. The certification authority requires a Full Scale Test to fulfill with the standards. In the case of the A400M the certification authority was EASA.

2 Test Objectives

In the case of the A400M was agreed that there would be two complete aircrafts structures, one for Static Tests (ES) and another for Fatigue Tests (EF) in order to achieve between both all the structural parts in static and in fatigue. Some structures parts of the structure are covered with the ES or the fatigue phase of the Static Test (ES), others with the EF. The Full Scale Static Test would be performed in the Airbus Military Structural Test Laboratory at Getafe, Madrid, and the Full Scale Fatigue Test would be performed at IMA Dresden. However, the most important tasks were the certification sceneries accomplished by each of the Full Scale tests.

For the Static Test **ES**:
Test objective is to support:

- The static justification of the main components (wing and fuselage), not only for the metallic parts but also the carbon fiber parts.
- The fatigue & damage tolerance justification of carbon fiber parts.
- Complete justification of the main interfaces, static and fatigue for: Wing-Fuselage, VTP-Fuselage. This interfaces are fully representative in the Full Scale Static Test
- FEM Validation through the test results
- To perform the functional tests for Aileron, spoilers and Ramp
- To perform some local tests which are very difficult or impossible to perform outside the whole structure as for example Nose Fuselage local cases or ramp local cases etc...

3 Full Scale Static Tests “ES”

The test specimen is considered representative of a production structure so it has been manufactured according to production drawings and the pieces and components have been inspected according to the standard quality procedures. The delivered structure to the Laboratory has a complete map of defects. Some of these defects are due to the production process and some others have been created for damage tolerance purposes. Both of them will be followed along the test in order to establish how tolerant to damages the structure is. The inspection processes is one of the most time consuming task, especially in carbon fiber structures.
The components of the structure to be tested were:

- Complete Fuselage
- Wing boxes, Outer Wing Boxes (OWB) & Centre Wing Box (CWB)
- Complete VTP (Vertical Tail Plane)
- For wing functional tests : aileron & spoilers, will be fitted on LH side

Due to the lack of some components (tested alone in dedicated test rigs), and being the ones which transmit the loads to the main components of the structure under test, it was needed to design simulated structures called "dummies" for MLG (Main Landing Gear), NLG (Nose Landing Gear), Flap Tracks, Engine Mounts and HTP (Horizontal Tail Plane).

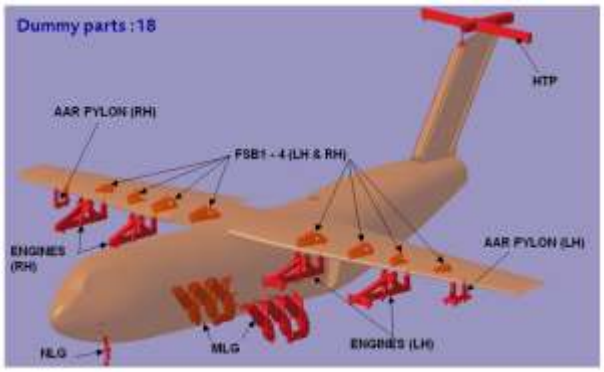


Fig. 2 Dummy Parts

3.1 Aircraft Support

The test specimen has to be supported by 6 points as any rigid body (6 DOF). These points are the reactions of the structure under test when any load case is applied. In the theoretical case of a pure flight test case these 6 reactions would be zero. All of them have load cells to control in any moment that their values are as expected. In the Fig. 3 are shown the reactions P_X, P_{Y1}, P_{Y2}, P_{Z1}, P_{Z2} y P_{Z3}.

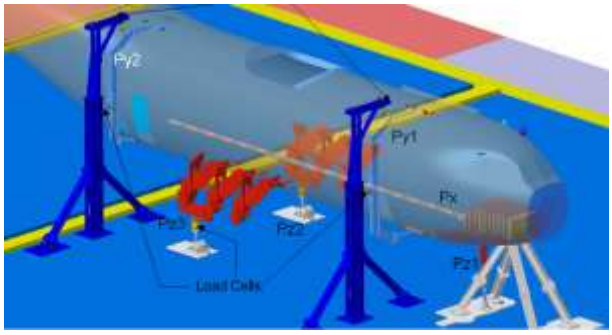


Fig. 3 Reactions

Because the test cases usually are a combination of several cases, the reactions are also used to apply some specific loads to the structure, for example the F_X load. The F_X load due to engine thrust and inertias was necessary to apply the reaction on the floor rails attachments. This load was distributed along 18 frames in order to simulate in a most realistic way the real distribution.

The reactions P_{Zi}, were used in combination with other external loads, for the simulation of landing cases and for aircraft balance in other test cases.

3.2 Test Cases Selection

Airbus Military was in charge of the test cases merging process and the corresponding test actuation loads. More than 120 cases were initially selected for testing. Because such number of cases is not practical, in terms of time consuming and because many cases are covering the same structural areas, it is a common practice to perform Combined Cases. The Combined Cases merge several individual cases which would cover only one structural area. It is important to know through the GFEM (General FEM) that the combined case is not penalizing any area more than the individual case would do. It is also important to make the aircraft balance. In the following figure is shown the typical process.

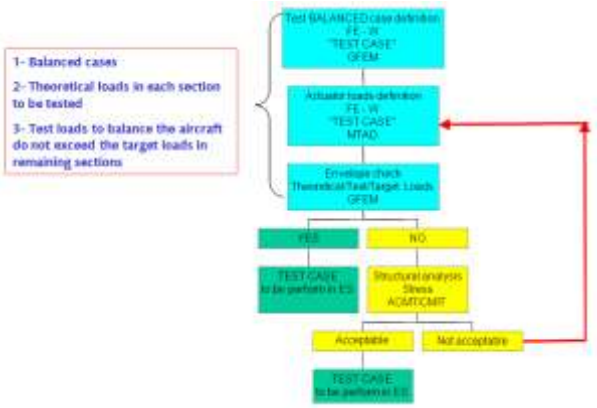


Fig. 4 Test cases Selection Process

In order to do the cases selection it is a common practice to use graphs as shown in the fig. 5, representing, for any structural area, the envelope of loads and Moments, or only moments or any adequate combination of the area to justify.

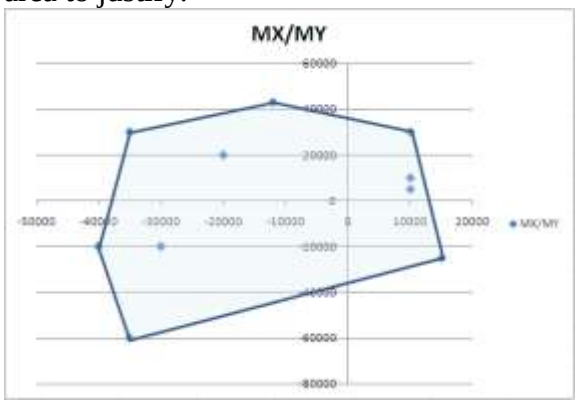


Fig. 5 Loads Envelope

The combination, or merging process, is usually as follows:

1. Individual cases per component (Fuselage & Wing). Original load cases.
2. Envelope case (Fuselage).

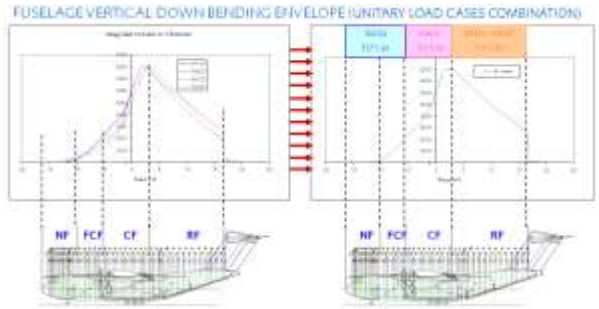


Fig. 6 Fuselage Test Cases Combination

3. Envelope case (Wing)

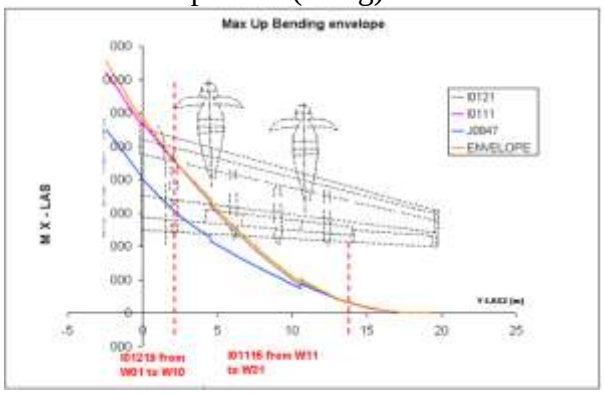


Fig. 7 Wing Test Cases Combination

4. Evaluation of different possibilities of combination (Fus. & Wing).

   Depending of:
   - Max actions (F_X….M_Z) in each a/c station (frames and ribs).
   - Max capability of the rig.
5. Aircraft balancing.
   - Once the initial match of both wing and fuselage is done, the overall balancing errors may be compensated by a/c constraints and/or small modifications in wing or fuselage loads, depending on the nature of the test case.
   - Obviously, for critical cases, the option is to use small loading modifications to correct balancing errors and to maintain the best possible match with theoretical loads.

3.3 Test Cases & Test sequence

The final test cases selection was as follows:

- Wing envelope: 5 (extracted from 13 load cases)
- Fuselage envelope: 5 (extracted from 14 load cases)
- Functional: 3
- Wing specific: 9
- MLG: 3
- NLG: 3
- NF specific: 5
- Pressurisation: 2
- Open ramp: 2
- Pressure & Flight case: 2
- Ramp cases: 4
- Stabilizer struts: 1

After combination between them: 37 Test cases was performed (18 Limit Load &1.25LL, 13 Ultimate Load, 3 Thermal, 3 UL after fatigue) The sequence was as follows:

- Static test
  - LL and UL cases will be applied
  - Wing Thermal cases at LL
- additional CFRP wing requirements
  - fatigue cycling for 2/3 DSG
- Static Tests
  - UL load cases

- Military Cases

Combined with some cases was necessary to apply the fuselage pressurization. The pressurization is typically called Delta P. In some test cases to Limit Load was necessary to apply Delta P and, it was also necessary to perform test cases with only pressurization up to 1.33 Delta P, which is the valve relief tare pressure multiplied by the required factor, according to the standards, in case of pressurization without any other loads (in this situation it corresponds to Limit Load). So, for only pressurization cases at ultimate loads were applied 1.5 x 1.33 Delta P, what means 2 times Delta P.

For the cases with thermal interactions, due to the different materials used in the structure, mainly in the joint of wing with fuselage, it was necessary to apply thermal loads to better simulate the thermal stresses on the area. Depending of the structure this kind of thermal interactions can be covered applying some extra loading factors but in this case the structure was going to be extremely penalized so the best solution was to manufacture a climatic chamber around the investigation area (CWB + Part of Left OWB + Part of Fuselage + all the joints wing - fuselage). The maximum thermal load applied was 70 ºC together with mechanical loads of the corresponding test case.

3.4 Load Distribution Systems

The load distribution Systems are very important to achieve the F_X, F_Y, F_Z, M_X, M_Y & M_Z distribution for each one of the selected load cases of the structure components. As much more external loads are applied (Hydraulic actuators) better combinations of F_X, F_Y, F_Z, M_X, M_Y & M_Z can be simulated. However, it is important to find the minimum necessary number of external loads in order to minimize costs.

The load distribution systems are basically the following:

- Wing
- Fuselage
- VTP (Vertical Tail Plane)
- Dummy Engines
- Dummy MLG & NLG

3.5 Wing Loading System

The loading system for the wing was based in a totally isostatic whiffletrees which are placed in both surfaces intrados and extrados. The basic principle is to push on both surfaces to avoid any surface delamination and be able to inspect both surfaces in any moment because the loading pads are not bonded to the skins. See Figs. 8 & 9. There were used **69** hydraulic actuators (39 on LH and 30 on RH) and 164 compression pads on both upper and lower skin surfaces.



Fig. 8 Wing Loading system

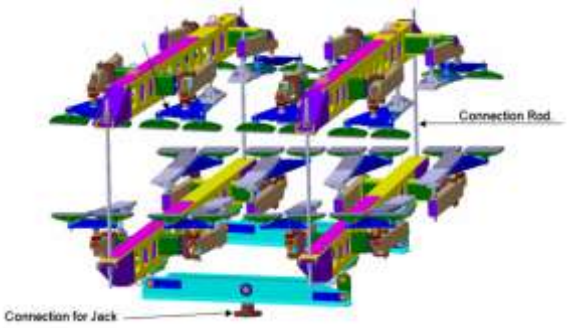


Fig. 9 Wing Whiffletrees example

3.6 Fuselage Loading System

The same basic concept of using whiffletrees isostatically determined was applied to the fuselage. In this case, in most of the places, were used built in fittings instead of rubber pads.

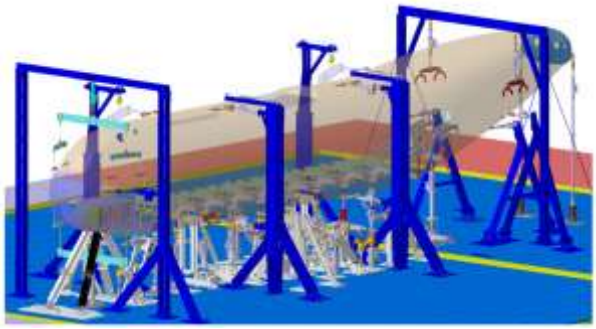


Fig. 10 Fuselage Loading System

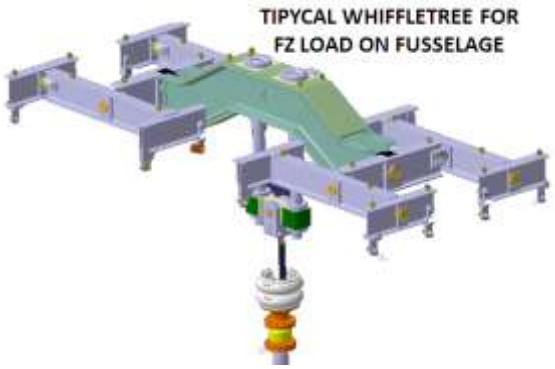


Fig. 11 Whiffletree inside Fuselage

3.7 VTP Loading System

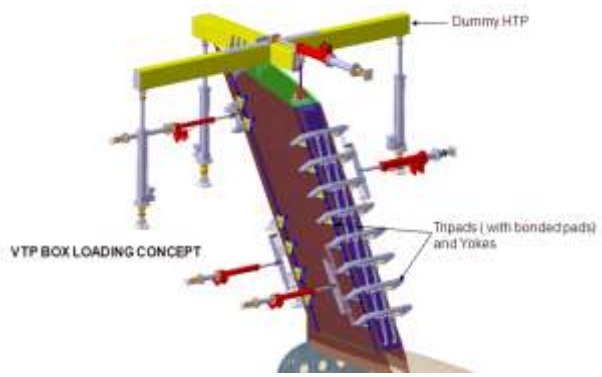


Fig. 12 VTP Loading System

3.8 Dummy Engines Loading system

The actuators were directly attached to the dummy engines applying locally the 3 Forces and 3 Moments. The Dummy engine transmit the actions to the corresponding attachments of the wing

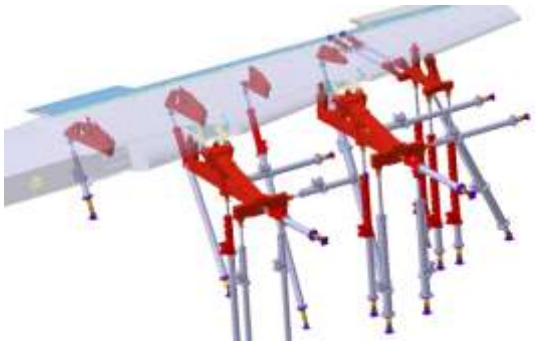


Fig. 13 Dummy Engines Loading System

3.9 Dummy MLG & NLG Loading System

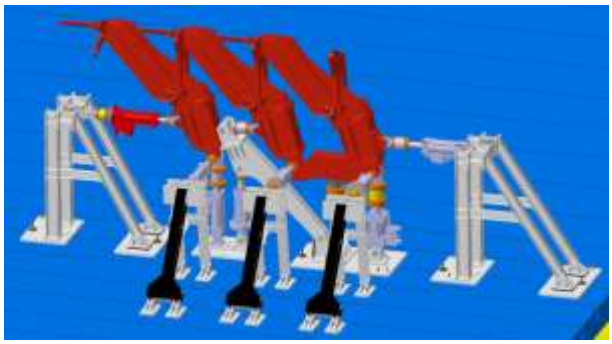


Fig. 14 Dummy MLG Loading System

4 Test Systems

The load control Systems in the test are very important in order to achieve an adequate simulation of the test cases. The Data Acquisition Systems are equally important for correct monitoring of the stresses that occurred in the structure.

In the load control systems is not only important to reproduce with precision the load in the test cases but to make special mention to the safety and integrity of the specimen. As we said before, an unexpected specimen failure will lead to an important economic and schedule impact. If the failure comes from a load control system

failure, the fail would impact not only in the same way but also in the test confidence.

During the tests execution usually appear warning situations caused by settlements in the load introduction Systems, unexpected rig failures, Computers and Electronic Systems failures, Hydraulic/electrical power supply failures or even operator failures. Any of them could lead to an emergency and cause an unexpected test unloading. Test unloading have to be fully under test control and have to be performed in a soft and symmetric way so that the structure never suffer damage due to any asymmetry that leads to an unwanted load case. The structure must be unloaded in the same way it was loaded.

4.1 Test Control System

The Airbus Military Structural Test Laboratory decides that the best system to safeguard the structure integrity would be based in a redundant system to load and unload the structure in case of system warning or failure. It was decided to use MTS equipments series Aero ST.

Two control systems of 124 channels were arranged, one for primary control and another one as an emergency unloading system. The primary system, also, while operative, will monitor the operation of the concurrent system. If the primary system doesn't suffer a critical failure will take on the complete test control, the loading and unloading. A failure in the primary control systems, will lead to a swap control to the other. In the hypothetical case of both systems failing it was foreseen an unloading method by switching the actuators to a pre-adjusted dumping mode.

In addition to the two concurrent systems it was necessary to develop, in conjunction with MTS, a Hydraulic Block (Smart Load Abort, SLA) that was capable to perform the functions required by both systems. The hydraulic block SLA had been developed by MTS for hydraulic actuators of equal areas. In the A400M case, they were hydraulic actuators of different areas and the SLA hardware was needed to be modified and the control functions too. This hydraulic block would operate even in the case of hydraulic power interruption taking advantage of the accumulated power in the test specimen and consequently with pressures remaining inside the actuator. This accumulated energy would be free using Direct Drive Valves (DDV´s) managed by the unloading control system.

The following figure shows schematically the interaction between the control and the unloading system. Also can be observed how the data acquisition system is completely integrated within the general system. The UPS's were used in a redundant way so if a power supply failure where to occur none of the control Systems would be affected.

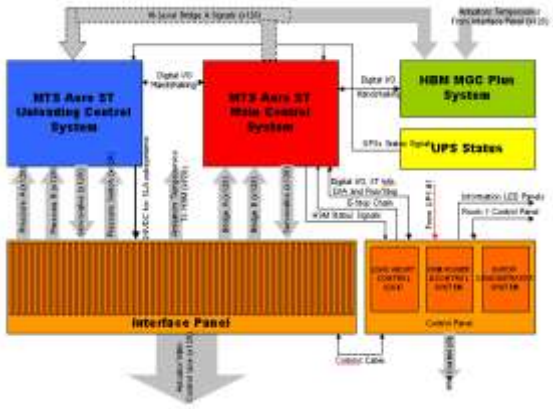


Fig. 15 Dummy MLG Loading System

In the following figure is shown a 201 MTS actuator with an SLA used in the Hydraulic System of the A400M Test.



Fig. 16 Hydraulic Equipments

4.2 Data Acquisition System (DAC)

As usual in these kinds of tests, the strain gauges are the most widely used sensor together with load cells. With the strain gauges are obtained strains and stresses anywhere they are located. In the A400M ES test there were bonded about 10.000 gauges and close to 140 load cells to control the external loads as well as the reactions. However, during the test monitoring it was necessary another more 7.000 virtual channels (calculated channels combining the readings of the physical channels) which were necessary also to be monitored

The DAC system hardware is based in the HBM MGC+ system while the software for data acquisition and monitoring is MTS AERO Pro software with special features specially made for this test. Both together are the ideal combination for the test following and data safety.



Fig. 17 DAC System

Due to the huge amount of data to monitor in real time several protocols were written to provide the clients with the proper service. A couple of days before de tests, the clients are requested to fill up some templates, requiring in some specific manner, the channels to be monitored as spreadsheets, graphs/t, graphs/xy etc… At the same time on the graphs could be represented the theoretical lines (or expected values) for each test case. To configure each one of the more than 40 terminals some macros were made which automatically generate the windows terminal lay-out with the required graphs and data previously requested with mentioned templates. The monitoring effort is simplified due to the easy windows management and the easy comparison with the expected values for the test case.

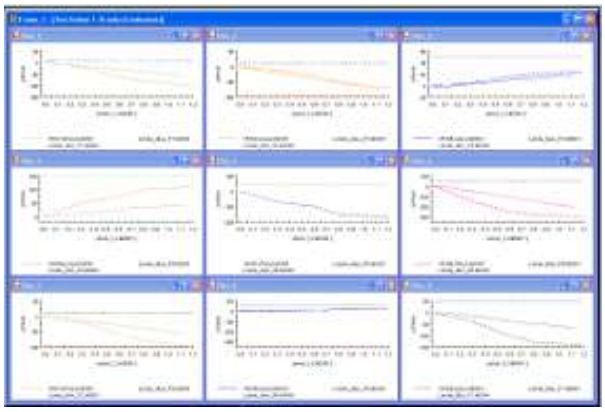


Fig. 18 Typical Monitoring Terminal Lay-out

During the static test loading process the load increments are usually 5 or 10% of the Limit Load. When the desired load step has been achieved the operator takes manually one or several scans after some standard settling time. However the system is running in background a circular buffer at the typical rate of 300 samples per second in order to be capable to investigate any event. This circular buffer can usually save 15 sec. before and after the event. In the case of carbon fiber, due to its brittle nature, the sampling rate should be increased at the maximum possible in order to be able to follow the strain changes, which can occur in less than one or two milliseconds in the case of a sudden structure failure.

5 Test Images

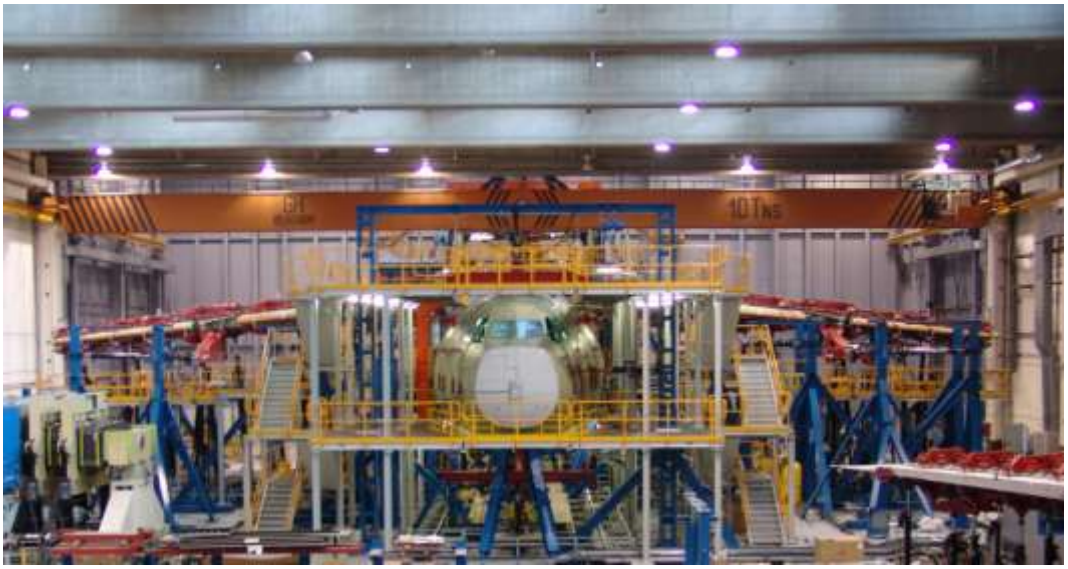


Fig. 19 Front View Photograph



Fig. 20 CATIA 3D Front View

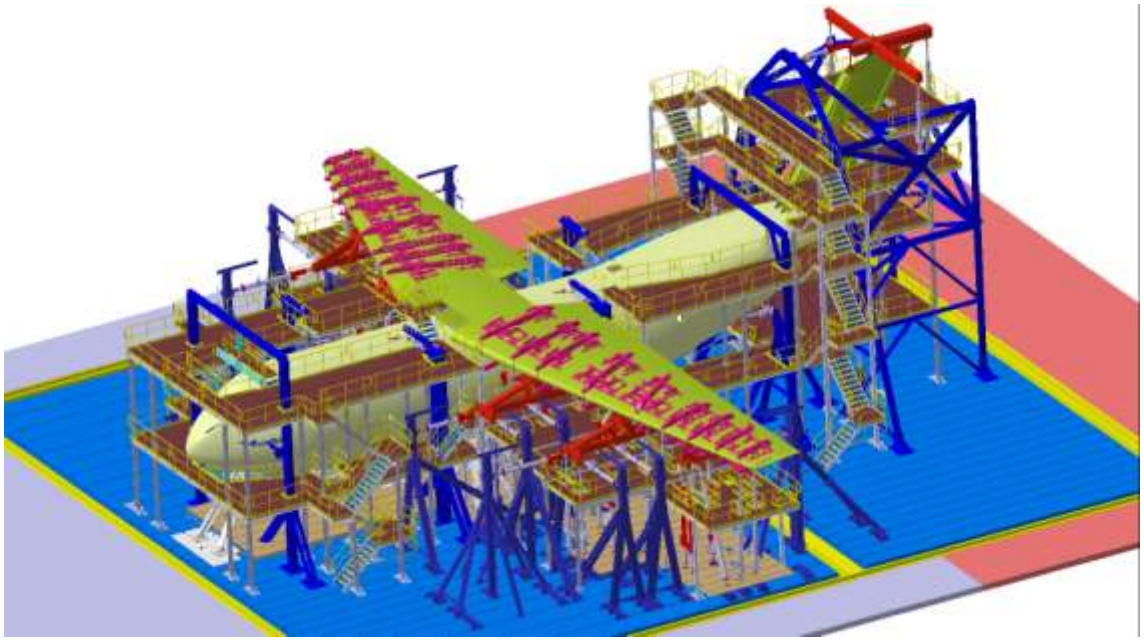


Fig. 21 CATIA 3D General View

6 Current Status

At the moment of writing this paper the test has completed all the test cases at LL and UL. The Fatigue cycling has been also completed together with some of the UL cases after Fatigue.

After completion of the last UL cases the structure will be ready to be certified.

Contact Author Email Address
Email: angel.gago@military.airbus.com

Abstract proposed for the 3rd CEAS Air&Space Conference Venice (Italy), 24-28 October 2011

Full-Scale Static and Fatigue Tests of Aircraft Components from Composite Materials

Maksim Smoljaninov, Kristina Soskoveca,
AVIATEST LNK, Riga, Latvia
Dr. Sc.Ing. Alexader M. Sorokin
AVIATEST LNK, Riga, Latvia
Aleksey Chepusov
AVIATEST LNK, Riga, Latvia
"AVIATEST LNK" Structural Testing Laboratory, Riga, Latvia (http://www.aviatest.lv)

The full-scale tests give the most reliable data on the strength of aircraft components from composite materials. These tests can be conducted as under static loading conditions to determine the ultimate failure load as under cyclic loading to determine the number of poster "flights" to destruction.

When making new plane type, it's appropriate to conduct both of the tests, because area of damage under static and fatigue tests are usually not the same.

A very important point in these tests is flight loads simulations. Usually, customer brings to the laboratory schedules for load distribution along the length of the component and along the chord, also repetition of the continuity of the load is needed (without significant jumps in loads). To realise it AVIATEST laboratory uses load distribution methods, using special fabric load pads.

One loading pad has maximal sizes 90mm x 200mm. It is designed in such a way as to transmit the load evenly on the surface of the component along the four strips that stick to the surface.

Experience in using such devices in AVIATEST shows maximum load on one load pad up to 350-400 kg. The advantage of using pads in the tests compared to the pusher (bottom) on the unit load method or tool trays is that the in-flight upper panel component (for example, the flap as shown in Figure 1) is subjected to suction effect - the pressure on the upper surface is bigger than on the bottom.

Usually in static tests the entire surface of the component is covered by these pads. Actuator load (usually just one actuator is used) is placed so that his line of action passes through the component's center of pressure, then the load through a levers system is sent on the transmitter pad (these instruments are also shown in Figure 1).

AVIATEST has big experience in composite aggregates tests (ailerons, spoilers, air brakes, flaps). AVIATEST has calculation procedures for determining the loads and pads sticking procedures.



Fig.1 Static Tests of the Flap in Structural Testing Laboratory AVIATEST LNK

CEAS AIDAA
3rd CEAS Air&Space Conference
21st AIDAA Congress

Evaluation of vibro-acoustic behavior of composite fuselage structures realized with embedded viscoelastic damping treatments

Amoroso F., De Fenza A., Monaco E., Pecora R., Lecce L.
Aerospace Engineering Department, University of Naples "Federico II" 80125 Naples – Italy
Reference author: Ernesto Monaco – ermonaco@unina.it

Keywords: *Composite Structure, Vibro-Acoustic behavior, Viscoelastic, Damping*

Abstract

In aerospace field, low structural damping, or high oscillations amplitudes, may impact negatively on structural stability and emitted noise. This is much evident if composite structures are referred to. In force of their stiffness-over-weight ratio, composite materials are deeply used in large assemblies and subassemblies belonging to fuselage and/or lifting surfaces (wing, tail empennages, control surfaces); on the other hand, structural advantages related to composite solutions are accompanied by drawbacks mainly due to materials' high vibration and noise permeability. By increasing structural damping it is possible to obtain a considerable noise and vibration reduction. Composite materials' structural damping can be increased through damping treatments. Among the most common damping treatments in aerospace field, viscoelastic materials embedding is widespread. In order to assess a proper evaluation of their induced damping performances, a rational approach must be defined in order to predict and/or experimentally evaluate damping of deeply heterogeneous layered structural arrangements. Damping of complex structures is often dominated by losses at joints etc. and so is very difficult to be evaluated by means of numerical models. So the damping of treated structures is usually evaluated by experimental tests. Two different coupons were analyzed in this work and the vibro-acoustic behavior was carried out by Impulse Response Decay Method (IRDM)[1]. The tests were performed at temperature condition of -35°C for the no-treated panel and at temperature conditions of 15°C, -20°C, -25°C and -35°C for the treated panel. For each temperature condition, and for each measurement point, loss factors trends have been evaluated. The effectiveness of the experimental approach has been proved by the mean of a statistical analysis of the obtained results.

1 Introduction

Structures subjected to dynamic loads, generally show structural damping values which are just slightly capable of reducing oscillation amplitude. In aerospace field, low structural damping, or high oscillations amplitudes, may impact negatively on structural stability and emitted noise. This is much evident if composite structures are referred to. In force of their stiffness-over-weight ratio, composite materials

are deeply used in large assemblies and subassemblies belonging to fuselage and/or lifting surfaces (wing, tail empennages, control surfaces); on the other hand, structural advantages related to composite solutions are accompanied by drawbacks mainly due to materials' high vibration and noise permeability. By increasing structural damping it is possible to obtain a considerable noise and vibration reduction thus augmenting overall composite appeal in aerospace applications. Composite materials' structural damping can be increased through damping treatments. Among the most common damping treatments in aerospace field, viscoelastic materials embedding is widespread; such materials are often used as foils package by thin aluminum sheets acting as constraining layers. In order to asses a proper evaluation of their induced damping performances, a rational approach must be defined in order to predict and/or experimentally evaluate damping of deeply heterogeneous layered structural arrangements.

Damping of combined and complex structures is often dominated by losses at joints etc. and so is very difficult to be evaluated by means of numerical models. Difficulties dramatically increase in case of composite materials realized with viscoelastic damping treatments. Numerical models

More difficult is the damping simulation for composite materials realized with viscoelastic damping treatments. Numerical model available in literature permit to simulate only viscoelastic materials and not the coupled composite materials-viscoelastic. For this reason, in general, the damping of materials and structures must be determined experimentally i.e. measured. In this work the vibro-acoustic behavior was carried out by Impulse Response Decay Method (IRDM)[1]. Two different composite structures were analyzed, the first one is no-treated composite fuselage structure tested at the temperature condition of -35°C; the second one is a treated composite fuselage structure tested at the temperature conditions of 15°C, -20°C, -25°C and -35°C. A sensors position affect and a loss factor results, for each temperature conditions, will be shown in the following.

2 Experimental set-up

The experimental set-up is composed by the excitation system, the climatic room and the acquisition system.

2.1 Excitation system

In this activity an innovative excitation system was been utilized. It allows us to excite the structure up to 3000÷3500Hz without compromise the structural behavior. It generates an impulse signal that approaches to the Dirac signal so allowing a direct measurement of Impulse Response Function (time history)[3]. The excitation system is made up of pressure regulator device, electrical valves, pneumatic drives and force sensor (Fig.1); all these ensure a perfect control in a time impact and in a amplitude load impact. This system can be automatized through a PLC (Programmable Logic Controller) and thanks its compactness and versatility it is possible to use it in a special conditions i.e. in a climatic room.



Fig. 1 Excitation system

2.2 Climatic room

The climatic room (Fig.2) is a device able to reproduce the environmental flight conditions in terms of temperature and humidity. Our tests have been performed in the climatic room of the acoustic laboratory of the Alenia Aeronautica and it allows temperature range variation from 15°C to -35°C and the humidity at 5%.



Fig. 2 Climatic room

2.3 Acquisition system

The acquisition system is composed by four mono-axial accelerometers (fig.5 a-b) located in a different position on the panel surface skin. The acceleration time histories are showed and registered in a real time by the signal analyzer device LMS scadas mobile SCM01 (Fig.3).



Fig.3 LMS scadas mobile

3 Structure lay-up

Two different coupons were analyzed in this work. The first one (fig.4-a) is a treated composite fuselage structure realized with a composite material (carbon-epoxy) developed using embedded viscoelastic damping treatments. The second one (fig.4-b) is the same composite material without viscoelastic damping treatments. This last one is as reference panel used to estimate the viscoelastic effect on the entire structure.

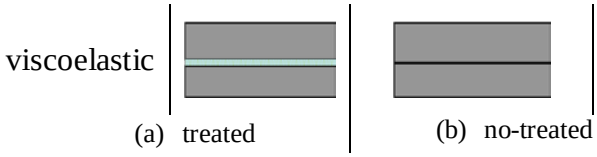
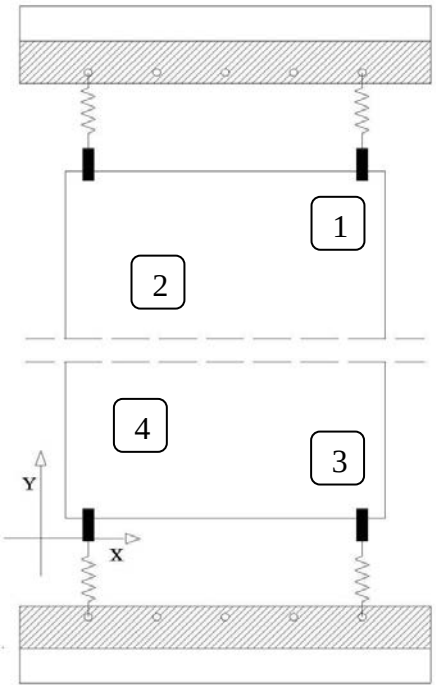
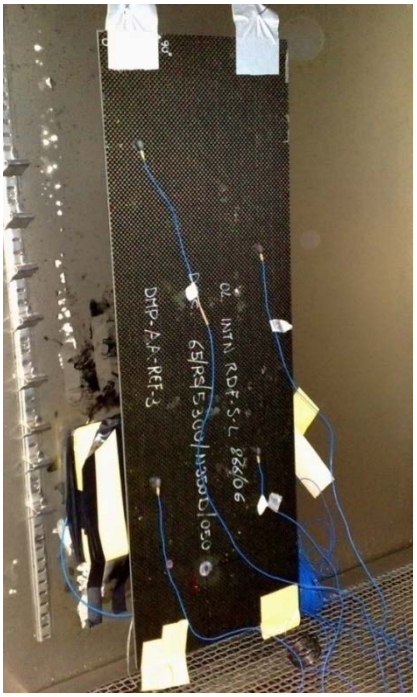


Fig.4 Panels lay-up

About the boundary conditions, the panel is linked to the climatic room through four springs (fig.5) in order to realize a free-free condition. The springs' stiffness are such to avoid the influence on the structural panels behavior of the vibrations due to climatic room. The coupon's dimension are equal to 610x200mm.



(a) CAD view



(b) Test panel

Fig.5 Test Panel and position of accelerometers

4 Analysis method

In literature there are different methods for measuring damping, each of them is based on different types of measurements. In this paper the procedure for damping evaluation based on decay of oscillations [1] was used.

The damping characterization of a structure can be made through different parameters, all related and equivalent to each other. For experimental damping evaluation it is of common use to refer to a dimensionless quantity, called loss factor. It represents the fraction of mechanical energy dissipated (generally as heat) in a vibration cycle [2] [3].

The main point of the procedure is that the impulse response is obtained directly in time domain and not as inverse frequency response of the structure. One of the assumptions underlying the application of this method is that the structural response is an IRF (Impulse Response Function). This claim allows to extract information from signal by applying the Hilbert transform. One of these information is the decay rate D that is related to loss factor (η) as follows:

$$\eta = -\frac{D}{4.34\omega_0} \qquad (1)$$

In the standard IRDM procedure, the Hilbert transform must be applied on each resonance peak and then this procedure evaluates the loss factor at resonance [5]. For highly damped structures, such as skins in advanced composite materials where the modal density is high, it is not possible to separate each resonance peak because the structural response at a resonance is influenced by contribution of the neighboring resonances (heavily coupled modes). In this case, it is assumed to obtain the loss factor in a frequency band as average loss factor in the band. Once acquired the Time History, the first step is to filter the signal in order to isolate the contributions to various bands; then the Hilbert transform is applied to the various bands and the Hilbert's envelope is evaluated. The final step is to evaluate the regression line, and then the decay rate of the response for each analysis band.

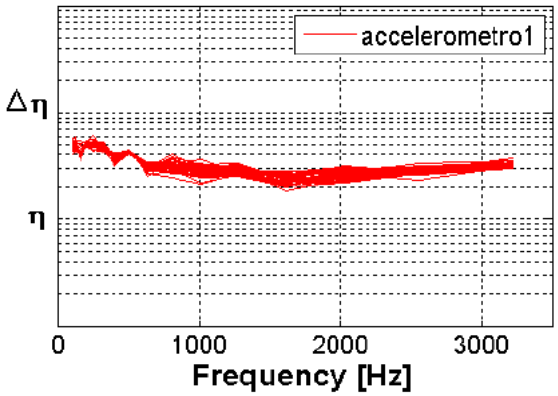
5 Results

Many tests at different temperature conditions have been performed for both no-treated and treated panels as follows:

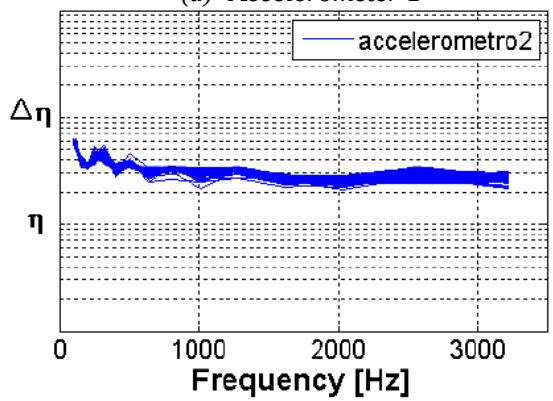
- 212 tests for the no-treated panel at temperature condition of -35°C;
- 168 tests for the treated panel at temperature condition of 15°C;
- 100 tests for the treated panel at temperature condition of -20°C;
- 50 tests for the treated panel at temperature condition of -25°C;
- 50 tests for the treated panel at temperature condition of -35°C.

5.1 Position effect

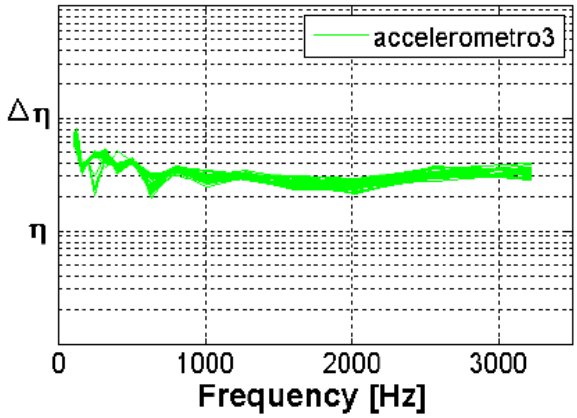
Four accelerometers for each test have been utilized in different positions and a first estimation of the bias error has been evaluated. In a figures 6-7-8, a loss factor versus frequency for each test on the treated panel in a temperature condition respectively of 15, -20, -35°C are shown. The loss factor ratio (Δη) is equal to 10^{-1}.



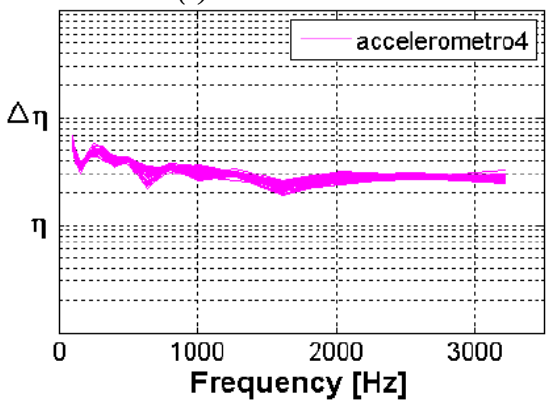
(a) Accelerometer 1



(b) Accelerometer 2

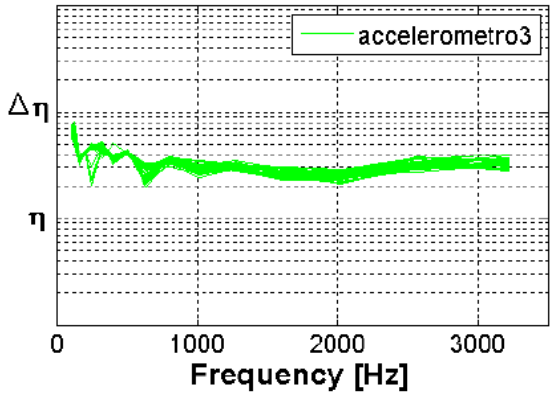


(c) Accelerometer 3

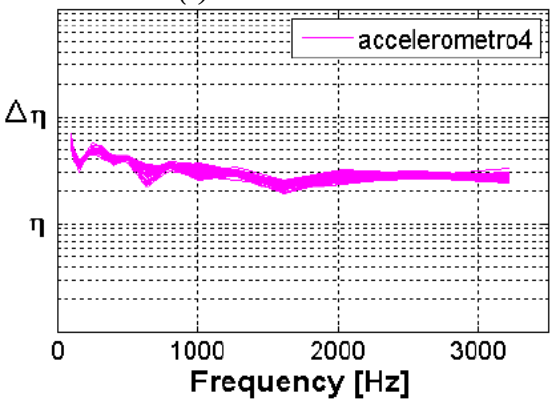


(d) Accelerometer 4

Fig. 6 Loss factor for each test on the treated panel at T=15°C

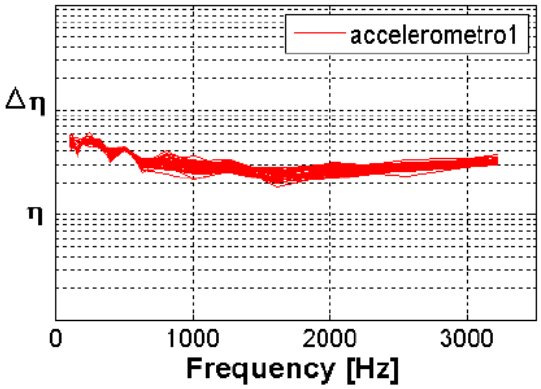


(c) Accelerometer 3

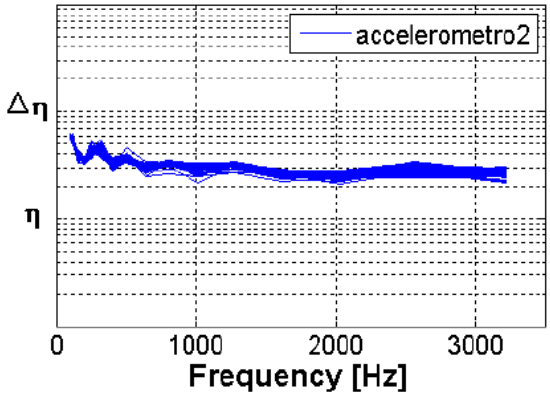


(d) Accelerometer 4

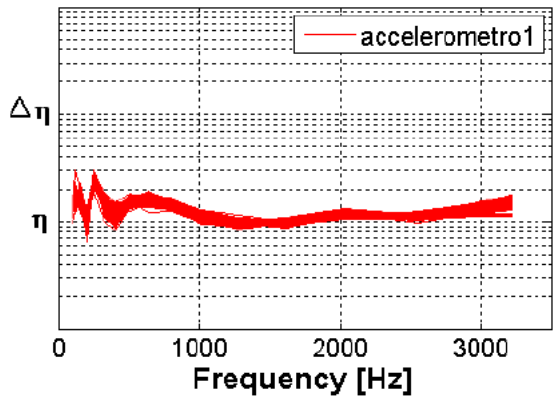
Fig.7 Loss factor for each test on the treated panel at T=-20°C



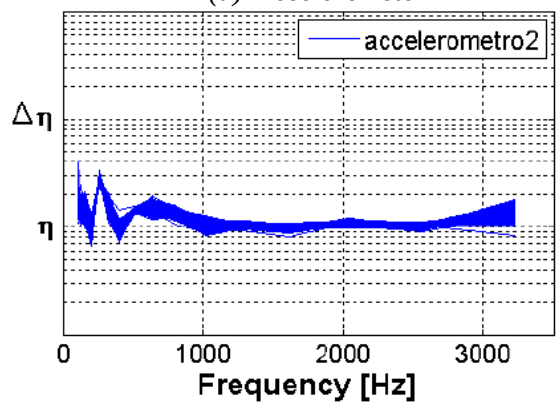
(a) Accelerometer 1



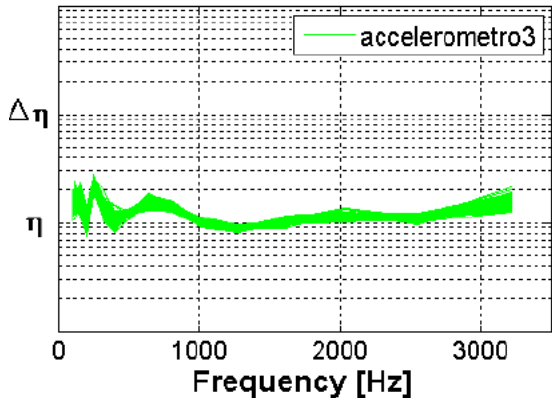
(b) Accelerometer 2



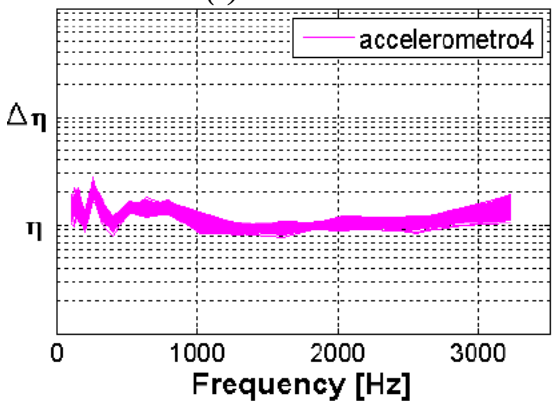
(a) Accelerometer 1



(b) Accelerometer 2



(c) Accelerometer 3



(d) Accelerometer 4

Fig.8 Loss factor for each test on the treated panel at T=-35°C

The last images shown that, at fixed temperature condition, the loss factor have the same curve progress and is not be dependent on accelerometer positions. So the bias error can be excluded. A position error would be gotten if one o more accelerometers were located in modal nodes. In this way the measured acceleration time histories not would have been correct because, as it is known, the displacements and so the accelerations would have been not real.

5.2 Temperature effect

Generally, the viscoelastic materials change their mechanical properties with the frequency and the temperature and these variation are known (fig.9).

In this work, the investigation was focused on the overall damping characteristics of composite plates (carbo-epoxy) realized with embedded viscoelastic damping treatments. Considering that in aerospace application from cruise and climb/ descent conditions the temperature ratio is very high, it is very important to appraise the range of temperature in which the composite structures give an optimal noise and vibration reduction (best loss factor).

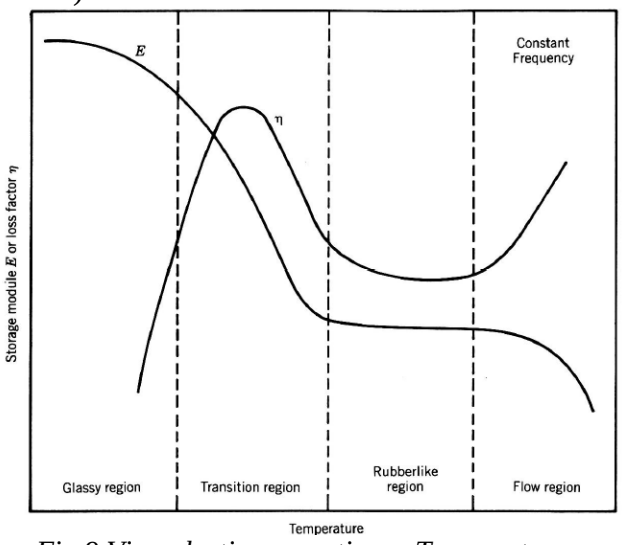


Fig.9 Viscoelastic properties vs Temperature

To get this many tests have been performed.

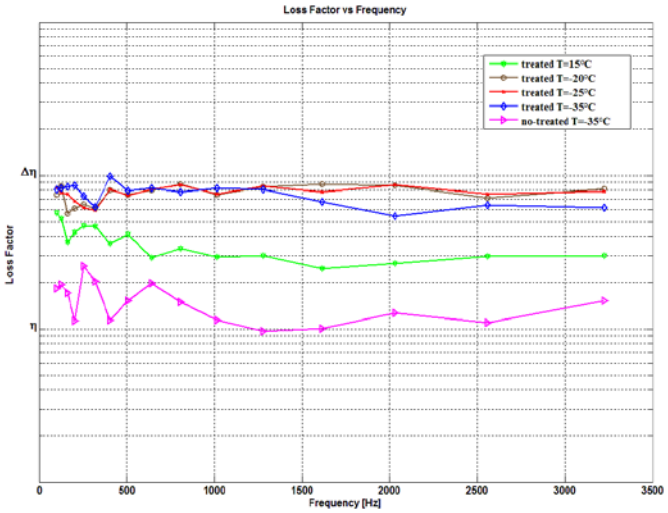


Fig.10 Loss Factor vs Frequency at varying temperatures

The results in terms of a loss factor versus frequency at different temperature conditions are shown in figure 10. At the same temperature (-35°C) the loss factor of the treated panel is almost one order of magnitude higher than that of the no-treated panel. Decreasing the temperature from 15°C to -25°C, the loss factor of the treated panel increases of about 5 logarithmic levels reaching the optimal value. At the temperature of -35°C, the loss factor has

the same curve progress referred to these at -20°C and -25°C up to 1500Hz, but it decreases of about 2 logarithmic levels for the other frequencies. This behavior is coherent with the viscoelastic properties (fig.9) where the optimal temperature range in which the composite structure have a greater damping is about from -20°C to -35°C.

To estimate the uncertainty level in our tests, standard deviations (fig.11) have been calculated as follows:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(\eta_i - \overline{\eta})^2}{N}} \qquad (2)$$

where N is the number of the tests, η_i is the loss factor at i-th test, $\overline{\eta}$ is the mean loss factor.

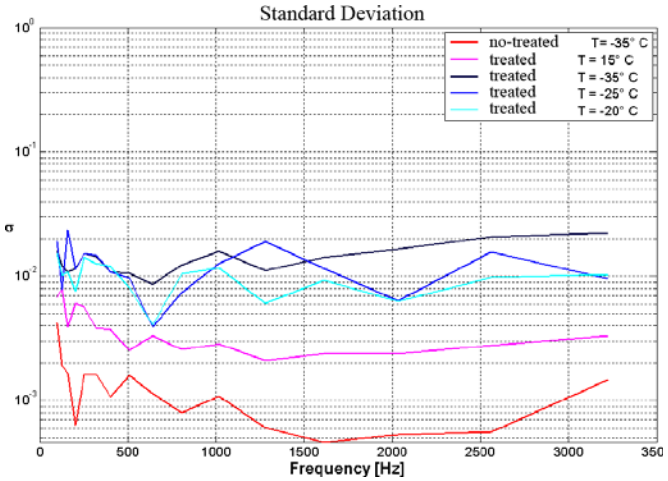


Fig.11 Standard Deviation

The least value of the standard deviation is obtained for the no-treated panel where the higher number of tests (212) were performed. The maximum standard deviation is obtained for the temperature condition of -35°C where only 50 tests were performed. Despite this, its value is about 10^{-2} and so the mean error is about 1% with a peak of 2.5% at the low frequencies.

6 Conclusion

During latest years many studies demonstrated the effectiveness of the structural damping properties of structures in order to reduce both airborne and structure-borne noise into aircrafts, and more in general wherever noise is an annoying factor. Innovative solutions able to increase damping levels with low weight associated increment are under investigation all over the scientific community, especially within the aeronautical world. The authors supported within many research project in the last years aeronautical company in developing numerical and experimental methodologies in order to characterize as precisely as possible innovative materials and damping solutions.

This paper presented an experimental approach based on a dedicated excitation system and a modified IRD method for damping measurement of simple structural components like plates or reinforced plates. The effectiveness of the approach has been proved by the mean of a statistical analysis of the obtained results; these are also in agreement with the typical behavior of damping capacities of viscoelastic materials with frequency and temperature. The standard deviation of the damping measurements, even if higher at lower temperatures due to the increasing measurements errors, is always on order of magnitude lower than the measured value, proving the reliability of the acquisitions and of the analysis process. Speaking of the specific results presented within this paper the damping characterization approach was able to demonstrate the effectiveness of “embedding” viscoelastic damping into a composite lay-up; even without considering the measured absolute values, it is clear that for the same experimental set-up and uncertainties, the damping capacities of the treated panels are clearly enhanced if compared to the bare material. Finally, the obtained results, can be considered a baseline for tuning numerical finite element models and optimization approaches aimed at providing also a validated design tool for the presented damping methodologies. Numerical simulations were performed by the authors in order to characterize the damping behavior for a composite structure realized with embedded viscoelastic treatments. In figure 12 a first result in terms of Frequency Response Function is presented. In accord to experimental one, the numerical FRF show how the simulation at -27°C is more damped that at 0°C. Further descriptions of the numerical methods and more

numerical results will be presented in next future works.

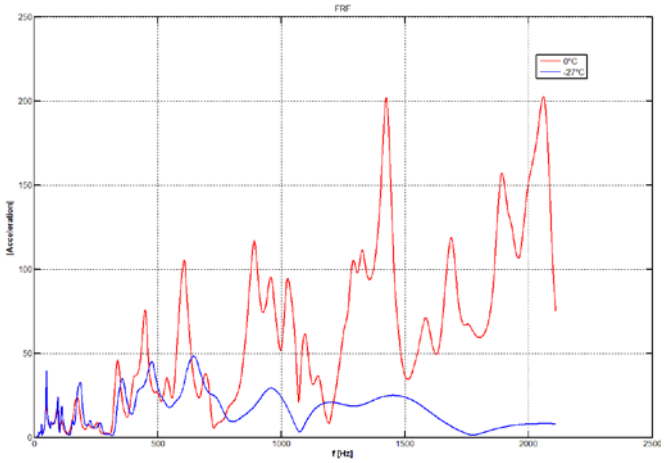


Fig.12 Numerical FRF

References

[1] Amoroso F., De Fenza A., Monaco E., Pecora R., Lecce L., Improved impulse response decay method (IRDM) for structural damping measurements, *The 17th International Congress on Sound & Vibration*, Cairo, Egypt, July 18-22 (2010).

[2] Gade S. , Herlufsen H., Digital Filter Technique vs. FFT Techniques for Damping Measurements (Damping Part I), *Brüel & Kjær, Naerum*, Denmark, 1994.

[3] Hahn S. L. , Hilbert transforms in signal processing, *Artech House*, Inc., Boston, 1996.

[4] Døssing O., Structural Testing, Part 2 – Modal Analysis and Simulation, *Brüel & Kjær*, 1988.

[5] Beranek Leo L., Noise and Vibration Control Engineering, *John Wiley & Sons*, 2006.

[6] K.U. Ingard, "On the theory and design of acoustic resonators", *Journal of the Acoustical Society of America* **25**,1037-1067 (1953).

[7] Johnson C.D., Kienholz D.A., Prediction of damping in structures with viscoelastic materials, *CSA Engineering, Inc. 1983 MSC/Nastran conference*, Los Angeles California, 24-25 March 1983.

[8] Nashif Ahid D., Jones D.I.G., Henderson J.P., *Vibration Damping*, *John Wiley & Sons*, 1985.